The Hunt for Mixed Octonion Algebras

Karsten Naert

Supervisor: Prof. Dr. Tom De Medts

Department of Mathematics
Faculty of Sciences
Ghent University
October 2017

Dissertation submitted in fulfillment of the requirements for the degree of Doctor in science: mathematics
# Table of contents

**General Introduction** .......................... 1

**Acknowledgements** .......................... 7

## Octonions

1 **Introduction** .......................... 11

2 **Linear algebra** .......................... 19  
   2.1 The Grassmann menace .......................... 19  
   2.2 Attack of the Clifford algebra .......................... 25  
   2.3 Revenge of the characteristic .......................... 26  
   2.4 Fabulous multivectors .......................... 32  
   2.5 1-fabulosity .......................... 37  
   2.6 2-fabulosity .......................... 39

3 **Superfabulosity** .......................... 45  
   3.1 Definition and goal .......................... 45  
   3.2 First lemmas .......................... 46  
   3.3 Connection with other structures .......................... 48  
   3.4 Trivectors .......................... 50  
   3.5 Other multivectors .......................... 56  
   3.6 Multiplication tables for octonions .......................... 60

4 **Interlude on orthogonal groups** .......................... 63  
   4.1 Introduction .......................... 63  
   4.2 Anisotropic tori .......................... 64  
   4.3 Generating orthogonal groups .......................... 66
5 Structures of type $G_2$ .................................................. 71
  5.1 The Lie algebra $\mathfrak{g}_2$ ............................................ 72
  5.2 Small characteristic phenomena ..................................... 75
  5.3 Generators for anisotropic $G_2$ .................................... 79
  5.4 Split versions .......................................................... 81
  5.5 The orbits of the mixed group ....................................... 84
  5.6 The mixed hexagon .................................................... 88

Twisting and Mixing

6 Introduction ............................................................... 93

7 Mathematical summary .................................................. 103
  7.1 Overview .............................................................. 104
  7.2 Three classes of exotic groups: problems ......................... 108
  7.3 Three classes of exotic groups: solutions ......................... 112
  7.4 The language of category theory ................................... 116

8 Twisting and mixing objects ........................................... 119
  8.1 The twisted and mixed categories ................................... 119
  8.2 Many functors ....................................................... 122
  8.3 Twisting versus mixing ............................................. 125
  8.4 Categories of presheaves .......................................... 130
  8.5 Mixed objects over a visible base object ......................... 132
  8.6 The concept of a fairy ............................................. 137
  8.7 Restriction to visible objects ....................................... 148

9 Twisting and mixing schemes ......................................... 157
  9.1 Twisted and mixed schemes ......................................... 158
  9.2 Rational points and functor of points ............................ 160
  9.3 Modules and sheaves ................................................ 164
  9.4 Examples ............................................................. 167

10 Twisting and mixing groups .......................................... 179
  10.1 Informal statement of the theorems ............................... 179
  10.2 Mixed groups over visible fields ................................. 181
  10.3 Twisted groups: the Suzuki–Ree groups ......................... 186
  10.4 Mixed groups and buildings of mixed type ...................... 186
  10.5 Exotic pseudo-reductive groups .................................. 191
Appendices

A Historical notes and references .......................... 197
  A.1 1950–1960 ............................................. 197
  A.2 1960–1970 ............................................. 199
  A.3 1970–1990 ............................................. 202
  A.4 1990–2000 ............................................. 204
  A.5 2000–2016 ............................................. 206

B Research notes ................................. 211
  B.1 m-categories ........................................... 211
  B.2 Fields .................................................... 218
    B.2.1 Twisted fields and $p = 3$ ......................... 219
    B.2.2 Mixed fields and $p = 2$ ......................... 221
  B.3 Conjectural taxonomy of $F_4$ ........................ 223

C Crash course in category theory .................. 227
  C.1 Categories, functors, natural transformations ..... 227
  C.2 Diagrams, limits and continuity .................... 232
  C.3 Presheaves and the Yoneda embedding ............. 236
  C.4 Adjunctions ............................................. 238
  C.5 Endomorphisms of the identity functor .......... 244

D Nederlandstalige samenvatting .................... 247
General Introduction

“Master Yoda: The Dark Side clouds everything. Impossible to see the future is.”

From Attack of the Clones

To almost every successful PhD there is a bright side and a dark side. The bright side is the thesis, the dark side the struggle which led to it. The latter story typically remains untold but for the occasion of finishing off this text, let me make an exception.

I started my PhD in 2011 as an assistant at the mathematics department. I had applied for funding at the FWO but that application had been declined—apparently I didn’t score well enough on statistics courses—so I was quite happy I could start a PhD in algebra after all.

My advisor, Tom De Medts, had written a research project which was all about Moufang sets: classifying subclasses of Moufang sets, and investigating in particular Moufang sets of type $^2G_2$, also known as the small Ree groups. I had been given some research notes which were ‘work in progress’ written by Tom De Medts in collaboration with Richard Weiss, dated August 2010. In these notes, they observed that a certain problem raised by Guralnick, Kantor, Kassabov and Lubotzky—and that I was familiar with through my master’s thesis—could be phrased elegantly in the language of Moufang sets (of type $^2G_2$). Their initial hope was that this would also mean that the problem had an elegant solution in the language of Moufang sets and to this end, they made a few interesting observations of computational nature. It was my task to see if I could make any progress on the problem and perhaps find this elegant solution.
General Introduction

To briefly elaborate on this, the language of Moufang sets refers to a computational device which deals with a certain class of groups such as these groups of type $^2G_2$. The description is very minimalistic, in terms of a group $U$ and a map $\omega : U^\times \to U^\times$—subject to a few conditions of course. Just to scare the reader a little, let me give some of the gory details for the case of $^2G_2$. We will need a field $k$ of characteristic 3, together with an endomorphism $\theta : k \to k : x \mapsto x^\theta$ with the property that $x^{\theta\theta} = x^3$. Amongst the finite fields, these are precisely the fields of order $3^{2e+1}$. (1) The group $U$ is then defined as the set $k \times k \times k$ with the operation\(^2\)

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} + \begin{pmatrix} d \\ e \\ f \end{pmatrix} = \begin{pmatrix} a + d \\ b + e + ad^\theta \\ c + f + ae - bd - ad^{\theta+1} \end{pmatrix}.$$ 

The map $\omega$ (3) is given by

$$\omega \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} -T_1(a, b, c)/N(a, b, c) \\ -T_2(a, b, c)/N(a, b, c) \\ -c/N(a, b, c) \end{pmatrix},$$

where

$$T_1(a, b, c) = a^\theta b^\theta - c^\theta + ab^2 + bc - a^{2\theta+3},$$

$$T_2(a, b, c) = a^2 b - ac + b^\theta - a^{\theta+3},$$

$$N(a, b, c) = aT_1(a, b, c) + bT_2(a, b, c) + c^2$$

$$= -ac^\theta + a^{\theta+1}b^\theta - a^{\theta+3}b - a^2 b^2 + b^{\theta+1} + c^2 - a^{2\theta+4}.$$

The description may be short but the computations quickly become horrendous! Tits found these formulas from deeper geometric considerations, but these were far beyond my knowledge at that point. For instance, Tits’s geometric reasoning shows that the denominator

---

1. The Frobenius automorphism generates $\text{Gal}(F_{3^h}/F_3) \cong C_h$, so it will be divisible by 2 if and only if 2 and $h$ are coprime.

2. It is customary to denote the operation with $+$ even though it is non-commutative!

3. Many authors prefer to use $\tau$ for $\omega$ but this map also has the property that it swaps 0 and $\infty$, and I like how the symbol $\omega$ is halfway in between 0 and $\infty$. 
$N(u, v, w)$ can only be 0 if $u = v = w = 0$, but to prove this just from the formula is really quite hard! For me, these were just a bunch of magical formulas that I was doomed to work with. It didn’t take long before I came to the conclusion that even if an elegant solution in this language existed it was still not very likely that it could be found in the same language.

The next problem that I worked on came from Hendrik Van Maldeghem, who wanted to determine the automorphism group of a certain geometrical object known as the *Ree unital*. The ambition was to show that this group is actually the Ree group $2^4 G_2$ (extended with automorphisms of the field) and so it fit nicely in the project of my PhD. This would be the final missing piece in Tits’s grand project of realizing all groups of Lie type as automorphism groups of geometrical objects, which starts with the angelic fundamental theorem of projective geometry and would then end with these devilish Ree unitals.

Initially, I felt that I could really make a contribution to the problem and offer some nice ideas and observations. But every hypothesis or conjecture always seemed to end up getting stuck in a swamp of computations where it slowly sank to the bottom, sinking my enthusiasm with it. One worrisome observation was that all the nice observations in the aforementioned ‘work in progress’ were actually just algebraic restatements of certain nice properties of this Ree unital, and so all these algebraic niceties were trivial from the right point of view. At one point I found an equation and if only I could show that this equation always has a solution, then this would solve the problem. But after a few weeks or perhaps months of aimlessly stomping my way through the swamp, I could only conclude that the situation was hopeless.

To be fair, it is hard to tell if I actually made any progress at all,

---

4See [DW10] for the algebraic proof, [VM98, p. 7.7.15] for the geometric proof.
5In the finite case, this result can be deduced from the classification of finite simple groups, but in the infinite case the result was and is not yet known for the Ree groups.
6In a different sense, Tits’s program was already completed by [HSV11], where it is shown that $2^4 G_2$ can be realised as automorphism group of a different geometry which is naturally associated with it.
because I later learned that Hendrik had been proposing this problem to many PhD-students over the course of the past few decades and two years later one of them proved that sometimes, this equation does not have a solution.

I deeply felt that there was so much about these Ree groups that I could not even begin to fathom, simply because my understanding of them was limited to this handful of magical formulas. I knew Lie theory and Chevalley groups were somehow important, but at the time I didn’t know what they were or how to use them. At one point I tried to read Chevalley’s important Tôhoku paper but I don’t think I even got to page 3.

At the same time, it occurred to me that perhaps the literature was simply inadequately equipped to deal with these groups; for instance nowhere in the existing literature could I find an investigation into the Galois theory of the fields over which these Ree groups can be defined—this is typically the first thing that people wonder about when they invent something new! Even the simple observation that such groups cannot be defined over an algebraically closed field seemed absent from the literature. (See Proposition B.2.1.) The entire situation where I felt I had to deal with these Ree groups with tools that were woefully inadequate, while the proper tools were either far beyond my reach or inexistent, and perhaps both, was a very frustrating experience. The problem with such failure is that you can never know if the failure is your own, and perhaps research in mathematics just isn’t for you, or if it is a sign that you should work on something else while such problems will have to lie dormant until someone finds the right approach. Anyway, I turned my back to the Ree groups in anger and frustration and wondering where I went wrong.

Then I found a welcome distraction. Every 8 years, the Mathematics Programme had to be reevaluated to acquire a teaching accreditation. I knew that the preparations hadn’t been going very smoothly, and my colleague Bert Seghers had convinced me to take up some minor duties here and there. But in a meeting with someone from the central administration of our university it became clear to us we were perfectly on track to not get this accreditation. The main problem was that we had to write a self-evaluation report to present to the evaluation
committee and this process had encountered some delays. Since my research had been slacking a bit, I took some time off my research to become secretary of the Mathematics Program Committee and, together with Bert, we took up the gauntlet. After some of the most hectic months of my life, we succeeded; the verdict of the evaluation committee felt a bit like graduating. But at the same time, almost two years had passed and I was left with this bitter feeling of getting nowhere with my PhD.

Meanwhile, my advisor had taken note of my lack of progress and proposed a few other problems for me to work on. Whether I was not suited for these problems or whether these problems were not suited for me is tough to answer but it certainly wasn’t a perfect match. One of the things that plagued me tremendously was that I seem to have a natural tendency to be distracted by wrong definitions, and the definition of a Moufang set certainly seemed amenable to such criticism.

A problem from those days that I worked on for quite a while was the problem of determining Moufang sets with a Hua subgroup of order 2. I’m not sure what I tried to achieve—ideally a classification, but that seemed a far stretch without a clear approach—but I felt that a careful study of Moufang sets with a Hua subgroup of order 1—these are the sharply 2-transitive groups—would be a good start. At that point, there was still a conjecture around—although it has been disproven a few years thereafter\(^7\)—that every such group has a regular normal subgroup. I really wanted to prove this conjecture but alas, none of my tools, as sharp as marbles, could prove the existence of such a subgroup, except in cases where it was already known.

During the third year of my PhD, desperate for a problem that I could actually solve, I joined my colleague Korneel Debaene to the Erdős centennial conference in Budapest. There we started thinking about a cute problem in finite geometry that we had heard from Bert. I ended up thinking about it for many months, sometimes together with Korneel or Bert. Although personally, I found it a great form of progress not being stuck on my own, we were now stuck with the three

\(^7\)https://arxiv.org/pdf/1406.0382.pdf
of us because the cute problem turned out to be quite resilient and
in hindsight all of our progress was quite superficial. In an attempt
to finally use some mathematics that is more clever than mulching
formulas, I ended up reading small portions of the book on additive
combinatorics by Tao and Vu, and I tried to apply these techniques
to our cute problem. I think I had some clever ideas, but of course,
nothing worked.

So by the end of year three, my motivation was at zero Kelvin and
my progress bar at zero percent; this latter fact hadn’t escaped my
advisor who could also read it in my annual progress report—seems
like no-one is immune to the tentacles of the administration—and he
invited me to his office to propose a new project to work on. We
met July 7th 2014 (I looked up the date) and he proposed me the
following project. There exist objects, he said, called twin SPO spaces;
these are related to Jordan pairs; Jordan pairs are related to algebraic
groups; algebraic groups are related to buildings. Therefore twin SPO
spaces are related to buildings, and perhaps you should find out how
that works in detail. The good news is that on paper, it was a fail
safe project. All the necessary ingredients were readily available in
the literature. The bad news is that I didn’t know what a twin SPO
space was, or a building. I had a vague idea what a Jordan pair or
algebraic group was supposed to be but the former of these I didn’t
like very much because of my earlier trauma caused by overexposure
to horrible formulas. He also mentioned a little side project, about
mixed groups of type $G_2$—not something that he seriously wanted me
to consider, but it had been on his mind lately. As it always goes in
such stories, I started working on one thing, and ended up doing the
other thing.

...to be continued in Chapter 1
I would like to thank:

- My advisor, for his advice;
- My parents, grandparents and siblings, for their support and encouragement;
- The people with whom I shared the office throughout the last six years, for their company and friendship;
- The people from the algebra group for the seminars and lively discussions;
- And a couple of old friends for being there, even when they were far away.

My thanks to Erik Rijcken for letting me use the LaTeX style files from his PhD thesis.
Part I

Octonions
1  

Introduction

"All of us have clear understanding of a few things and murky concepts of many more. There is no way to run out of ideas in need of clarification. The question of who is the first person to ever set foot on some square meter of land is really secondary."  

Bill Thurston

This first part of this thesis arose out of a study of algebraic $k$-groups of type $G_2$, which we will in this part mainly study through the group of $k$-rational points $G_2(k)$.\footnote{https://mathoverflow.net/a/44213/44668} It is quite well known that over fields of characteristic 3 such groups exhibit curious behaviour, as witnessed by the existence of a related Ree group $G_2(k, \theta)$ over fields $k$ with a Tits endomorphism $\theta : k \to k$, and the existence of a related mixed group $MG_2(k, \ell)$ for a field $\ell$ which satisfies $k^3 \subseteq \ell \subseteq k$.

Every such group $G_2(k)$ arises as the group of automorphisms of an octonion $k$-algebra $\mathbb{O}$: $G_2(k) = \text{Aut}_k(\mathbb{O})$. The question that Tom De Medts originally proposed on that July 7th 2014 was to investigate whether there exists a variant on the theme of an octonion algebra, a hypothetical mixed octonion algebra $M\mathbb{O}$, such that

$$\text{Aut}_? (M\mathbb{O}) = MG_2(k, \ell).$$

The question mark here signifies that it is not clear that $M\mathbb{O}$ should be an algebra over a field, and in fact one rather expects this not to be the case. The issue is that in contrast with $G_2(k)$, the group $MG_2(k, \ell)$ is usually not a group of rational points of an algebraic $k$-group in

\footnote{In this part, I will use $k$ to denote a field, which frees up the letter $k$ that will denote a natural number later on.}
a natural manner so it is not even clear what sort of algebras one
should be investigating.

The inspiration for this question grew out of the work [CD15] where
Tom De Medts and Elizabeth Callens investigate another type of
mixed groups, related to the diagram $F_4$ and in characteristic $2$. In
this work they note that that certain spaces of the type $k^6 \oplus \ell$ play an
important role.\footnote{For the specialist: it concerns a Moufang set $F_{4,1}$ with a $B_3$ diagram as its
anisotropic kernel. It is determined by the traceless part of the norm of an octonion
division algebra. In the mixed case this invariant is then a mixed (anisotropic)
quadratic form.} Although one can of course see $\ell$ as a vector space
over $k$, the dual roles played by the fields $k$ and $\ell$ have long been
recognized by incidence geometers—starting with Tits—and so each
time such a structure pops up in the study of some kind of mixed
group, the commonly accepted wisdom is that this space is really a
vector space over two fields at the same time.

But at that point, I had no ambition to think about the problem in
a generality beyond $G_2$ in characteristic $3$, which has the nice feature
decoupling the problems related to quadratic forms which arise in
characteristic $2$ and problems related to having a $2$-fold or $3$-fold edge
in an associated Dynkin diagram. My plan of attack was as follows: It
is known that the group $MG_2(k, \ell)$ is a subgroup of $G_2(\ell)$ in a natural
manner and thus acts on an octonion $\ell$-algebra $O_\ell = O \otimes_k \ell$. So
if there is a thing that deserves the name mixed octonion algebra,
one expects to find it perhaps by studying the orbits of the action of
$MG_2(k, \ell)$ on $O_\ell$.

But I didn’t find it so easy to compute these orbits at first. The
reason is that although the group $MG_2(k, \ell)$ has been defined by
Tits very explicitly as a subgroup of $G_2(\ell)$ with certain generators,
I didn’t understand the action of $G_2(\ell)$ on $O_\ell$, or equivalently, the
action of $G_2(k)$ on $O$, in terms of these generators. So I looked up
the proof (in [SV00]) that $G_2(k) = \text{Aut}_k(O)$ and to my surprise the
proof simply showed that $\text{Aut}_k(O)$ was a group of rational points of a
14-dimensional algebraic group with certain properties, from which it
could be concluded that it had to be $G_2(k)$ by means of a classification
result for semi-simple algebraic groups. I found this proof unsatisfying
and decided that before I could continue my mixed path, I had to understand this action a little better.

An important observation is that, provided the characteristic is not 2, one can always decompose the algebra as $\mathbb{O} = \langle 1 \rangle \perp V$, since $\text{Aut}_k(\mathbb{O})$ preserves the unit and inner product which live on the octonion algebra, and therefore also the 7-dimensional orthogonal complement $1^\perp = V$. The action of $G_2(k)$ on the line $\langle 1 \rangle$ is trivial anyway, so we should focus on the other part $V$ where we expect $G_2(k)$ to act more or less transitively, i.e. with few orbits. (The action of $G_2(k)$ on $V$ is called the standard representation but I did not know this then; it is also not very relevant for now.) Now it turns out that the space $V$ is naturally endowed with a product $\times : V \times V \to V$, in addition to the quadratic form $q : V \to k$. In fact, characterizing this product is just as good as characterizing the octonion algebra because the group of automorphisms of $(V, \times)$ is also $G_2(k)$. In a way the product $\times$ is even better, because it is anti-symmetric in the sense that $a \times b = -b \times a$ for all $a, b \in V$, and so it can be seen as a linear map $\wedge^2 V \to V$.

A few years earlier, Kris Coolsaet had given me an old set of notes of his. In his PhD, he had investigated the groups of type $F_4$ and two related classes: the groups of type $E_6$ and also the large Ree groups, i.e. the groups of type $^2F_4$. He told me that one day he intended to do a similar investigation into groups of type $G_2$ and $^2G_2$. For this purpose, he had a set of notes in which he investigated properties of this cross product $\times$. His approach was to choose a basis $e_0, \ldots, e_6$ for a 7-dimensional space $V$ and define a trilinear operation $V \times V \times V \to k : a, b, c \mapsto \langle a, b, c \rangle$ such that $\langle e_i, e_{i+1}, e_{i+3} \rangle = \varepsilon_i$ for a certain set of (what he called) structure constants $\varepsilon_0, \ldots, \varepsilon_6$. From there, and with hard computational work, he deduced many properties of the cross product $\times$ and dot product $\cdot$ which are related to the trilinear operation by $(a \times b) \cdot c = \langle a, b, c \rangle$.

I quickly realised that the approach would be relevant in characteristic $\neq 2$ only and that the structure constants should be replaced by an arbitrary quadratic form $q : V \to k$, and so I set for myself the goal to work through his computations in a computation-free manner. On
Wikipedia, I learned there was a way to express the product $x \times y$ as

$$x \times y = (x \wedge y) \downarrow F,$$

where $F$ is a certain trivector $F \in \wedge^3 V$ and $\downarrow$ an operation

$$\wedge^i V \otimes \wedge^{i+j} V \to \wedge^j V$$

which is called contraction. (I suspect this is a way of looking at $G_2$-structures that is very familiar to differential geometers, but I am still not very acquainted with this side of the literature.) Nonetheless, I recognized that the trilinear operator $\langle \cdot, \cdot, \cdot \rangle$ should also satisfy

$$\langle x, y, z \rangle = (x \wedge y \wedge z) \downarrow F,$$

and moreover in an article by Wilson [Wil10], where he provides a minimalistic description of groups of type $2G_2$, he finds—quite ad hoc—a map $V \rightarrow \wedge^2 V$ which can also be interpreted as a map $x \mapsto x \downarrow F$. Of course, the trivector $F$ is in each case the same and so this single trivector neatly packages all there is to know about the octonion product. So this led me to the conclusion that the real object of study should somehow be the Grassman algebra $\wedge^* V$, and all maps that I described are just pieces of a map

$$\wedge^* V \to \wedge^* V : u \mapsto u \downarrow F.$$

(In hindsight, I don’t think I did a very good job in making use of this algebra. In practice, I still think of its graded pieces separately even though there ought to be a way to avoid this. See Problem 2.3.6.)

So the idea was to characterize this trivector $F$ by a certain property, from which it should be easy to deduce the properties of the trilinear form and the cross product that Kris Coolsaet found in a computation-intensive manner. I found the property that I was looking for in an article by Brown and Gray\footnote{[BG67], written 25 years before Reservoir Dogs!} where they study cross product algebras. In the language that I have been utilizing, it looks like this:

$$q(x) = q(x \downarrow F).$$
An issue here is that the formula is not valid for arbitrary $x$, but instead $x$ must have a specified degree and be pure.\textsuperscript{5} I also inserted a scalar $\lambda$ into the equation because although one can typically assume that $\lambda = 1$, it is precisely the observation that sometimes $\lambda = 3$, and thus $\lambda = 0$ in characteristic 3, which leads to the interesting behaviour in characteristic 3. Another issue is that $F$ must be homogeneous of a certain degree and although I wouldn’t mind if this were simply a consequence of a (better) axiom, as an ad hoc demand, I am not too fond of it.

Nonetheless that is the setup that I started working with. These \emph{fabulous multivectors} are already quite a wild generalization of the original research problem which occurs when the degree of $F$ is 3 and the degree of $x$ in the above formula is equal to 2. In general, the case where the degrees of $x$ and $F$ differ by 1 is the case of a \emph{superfabulous multivector} to which Chapter 3 was devoted. This situation also corresponds to the cross product algebras that were studied and classified by Gray and Brown, see Section 3.3. Along the way I noted that there is more to the relation between the Fano plane (which is a very small combinatorial structure and in particular a block design) and octonion algebras (which are a particular case of \emph{fabulous multivectors}) than meets the eye. There are many properties for block designs that seem to have a natural generalization to the world of fabulous multivectors, as will be discussed in Section 2.4. Also the other cross product algebras in Gray and Brown’s classification are naturally associated to block designs. So I was a bit tempted to investigate this deeper but there are two issues.

The first issue is a bit embarrassing really: I don’t know any other examples. There are the ones that have been found by Gray and Brown, which are all $t$-fabulous $k$-multivectors when $k = t + 1$, but I don’t know any others. Perhaps because there aren’t any? This would make the whole philosophy that ‘fabulous multivectors correspond to block designs’ a bit hollow!

The second issue is that I was really still trying to understand mixed groups of type $G_2$ so there was no point in getting too far off track. I

\textsuperscript{5}Proposition 2.6.2 removes the purity assumption in the most important case by adding a term which vanishes on pure vectors.
have had this problem quite often throughout my work, and the moral
seems to be that one cannot follow just about any rabbit through the
hole—one should focus on the odd-looking rabbit that wears a watch!
So instead I have inserted little ‘Problem’ sections throughout the
text to indicate the locations of rabbits and it is up to the interested
reader to keep an eye out for suspicious rabbits with watches.

So here is the content of Part I:

- Chapter 2 sets up the general framework that we will work with;
- Chapter 3 is mainly concerned with providing a different proof
  of the classification of Gray and Brown by exploiting the con-
  nection with block designs, and deriving some consequences for
  octonion algebras;
- Chapter 4 is an interlude on orthogonal groups where we ex-
  plain how to obtain a nice set of generators from within our
  framework;
- Chapter 5 provides a description of the Lie algebra $\mathfrak{g}_2$ and
  from there a description of $G_2$ in its action on $V$ and therefore
  on the octonions. We also (finally!) manage to describe the
  (most relevant) orbits of the split $G_2(\ell, k)$ in its action on $V$, for
  $k^3 \leq \ell \leq k$.

Let me summarize the outcome of all this. It is important that the
action of the split group $G_2(k)$ on $V$ preserves the quadratic form.
This implies that the subsets of $V$ of the form \{ $v \in V \mid q(v) = a$ \},
i.e. the level surfaces of the quadratic form $q$, must be invariant under
$G_2(k)$. In fact, they are also orbits under $G_2(k)$ and, as it turns out,
if $a \neq 0$, also under $G_2(k^3, k)$—and thus in particular for the mixed
group $G_2(\ell, k)$ which is contained between $G_2(k^3, k)$ and $G_2(k)$. At the
time, this was a huge disappointment for me, since I had expected
these orbits to split into multiple orbits, which I would then try to
combine into the mythical $\textit{mixed octonion algebra}$.

The quadric $Q$ though—given by the equation $(q = 0)$—which is also
an orbit under $G_2(k)$, is no longer a full orbit under $G_2(\ell, k)$. It turns
out that on $Q$, one can define an invariant under the action of the
short root groups which takes values in the group $(k, +)/(\ell, +)$. By
taking the set of vectors for which this invariant vanishes, and still
vanishes under the action of the long root groups, one obtains an interesting orbit $\mathcal{H} \subseteq \mathcal{Q}$ of $G_2(\ell,k)$. A mixed octonion algebra, at last? Alas, this orbit too was quite foreseeable, since the vector lines of $\mathcal{Q}$ are really the set of points of a generalized hexagon on which the group $G_2(k)$ acts. Within this hexagon lives a mixed hexagon on which the subgroup $G_2(\ell,k)$ acts and it is really this orbit that I found.

When I noticed this, I abruptly quit all research in this direction, and I only started to look back at it in some detail a few weeks ago for the occasion of writing down this thesis. That is why Part I is really a gilded set of research notes rather than a wrapped up article with a clear outline.

So it would appear this research project ended with another failure. I remember having a discussion with Koen Struyve on the subject somewhere in November 2015, after which I could only conclude that there probably was no mixed octonion algebra in sight after which I angrily left my office at the department, walked around the block and decided it was time to try something new.

... to be continued in Chapter 6
The applications of the theory of exterior algebras are very wide, e.g.: theory of determinants, representation of linear variety in projective space using Plücker coordinates, and the theory of differential forms and their applications to many branches of analysis. But I am sorry not to be able to describe them in detail, because of the limitation of time.

Claude Chevalley

2.1 The Grassmann menace

We must first recall a number of standard notions from linear algebra. Although the exposition and notations are our own, the results mentioned in the first three sections of this chapter are far from new and go back on Chevalley, Cartan and ultimately Grassmann. Good references of a general nature are [Che97] and [Bou98, Chapter III].

Let $k$ be a field and $V$ a vector space over $k$ of finite dimension. To $V$ corresponds a Grassmann algebra\(^2\) $(\wedge V, \iota)$, which is pair consisting of a unital associative graded $k$-algebra $\wedge^* V$, with product denoted by $\wedge$, and a morphism of vector spaces $\iota : V \to \wedge^* V$. The pair is

\(^1\)[Che97, Preface]
\(^2\)It is remarkable that in the mathematics literature, we speak of the exterior and Clifford algebra, whereas in more physics oriented literature the terminology Grassmann and geometric algebra is preferred instead. As a compromise, and to acknowledge Grassmann’s contributions to mathematics, we will use Grassmann and Clifford.
characterized by a universal property, and made explicit by a well
known construction that we shall not repeat here.

The Grassmann algebra is endowed with a natural \( \mathbb{N} \)-grading

\[
\bigwedge^i V = \bigoplus_{i \in \mathbb{N}} \bigwedge^i V.
\]

We denote the projection onto the grade \( i \) component of \( x \in \bigwedge^i V \) by

\[
\bigwedge^i V \to \bigwedge^i V : x \mapsto \langle x \rangle^i.
\]

A non-zero element \( x \in \bigwedge^i V \) is homogeneous of degree \( \ell \) if

\[
x = \langle x \rangle^\ell
\]

for some \( \ell \); in this case we write \( \ell = \delta x \). A pure (or decomposable) element of the Grassmann algebra is one of the form \( x = x_1 \wedge \ldots \wedge x_\ell \) for some \( \ell \), where all \( x_i \in V \).

We will also consider the opposite Grassmann algebra \( (\bigwedge^i V)^{\text{op}} \) which is as a set equal to \( \bigwedge^i V \) but has a product given by \( v \wedge^{\text{op}} w = w \wedge v \); by the universal property the map \( V \to (\bigwedge^i V)^{\text{op}} \) extends to a \( k \)-algebra isomorphism \( \text{op} : \bigwedge^i V \to (\bigwedge^i V)^{\text{op}} : (x_1 \wedge \ldots \wedge x_\ell)^{\text{op}} = x_\ell \wedge^{\text{op}} \ldots \wedge^{\text{op}} x_1 \).

A central object in our considerations will be the algebra \( E(V) = \text{End}(\bigwedge^i V) \), which inherits a \( \mathbb{Z} \)-grading from \( \bigwedge^i V \) where the elements of grade \( i \) are those endomorphisms that map \( \bigwedge^j V \) into \( \bigwedge^{i+j} V \) for all \( j \). The subspace of elements of grade \( i \) will be denoted \( E_i(V) \).

Left multiplication by \( x \in \bigwedge^i V \) will be denoted by \( a_x^{\dagger} \in E(V) \), i.e.

\[
a_x^{\dagger}(y) = x \wedge y;
\]

so we obtain a map

\[
a^{\dagger} : \bigwedge^i V \to E(V) : x \mapsto a_x^{\dagger}
\]

which is a \( k \)-algebra homomorphism which preserves the grading.

To obtain elements of negative grade, we will assume that \( V \) is endowed with a \( k \)-bilinear form \( \beta : V \times V \to k \) and we will use the notation \( \beta^b : V \to V^* : x \mapsto \beta(x, \cdot) \) for the corresponding musical morphism. We call \( \beta \) non-degenerate if \( \beta^b \) is bijective. There is an opposite bilinear form \( \beta^o \) given by \( \beta^o(x, y) = \beta(y, x) \) which is non-degenerate if and only if \( \beta \) is non-degenerate. We say that \( \beta \) is symmetric if \( \beta = \beta^o \). Finally, we will use the abbreviation \( q(x) = \beta(x, x) \).
2.1. The Grassmann Menace

For an arbitrary $v \in V$, we may define a linear map $a_v \in E_{-1}(V)$ by extension of

$$ a_v(x_1 \wedge \ldots \wedge x_\ell) = \sum_i (-1)^i \beta(v, x_i) x_1 \wedge \ldots \wedge \hat{x}_i \wedge \ldots \wedge x_\ell. $$

Since $(a_v)^2 = 0$, the universal property of $(\wedge V)^{\text{op}}$ implies that $a : V \to E(V)$ extends uniquely to a $k$-algebra homomorphism

$$ a : (\wedge V)^{\text{op}} \to E(V) : v \mapsto a_v. $$

The operator $a$ attached to the form $\beta^\circ$ will be denoted by $a^\circ$.

(We choose to let $a$ be an anti-homomorphism to minimise the number of minus signs that must be written later on. The idea is that we would like to let $a$ act on the right and $a^\dagger$ on the left, but since we want to compose both operators this is not always possible.)

The following lemma provides a number of useful formulas.

**Lemma 2.1.1.** Let $x, y, z \in V$ and $h \in \wedge^\bullet V$ be homogeneous. Then we have the following properties:

1. $a_x a_h^\dagger - (-1)^{\delta h} a_h^\dagger a_x = a_{a_x h}^\dagger$.
2. $a_x a_y^\dagger + a_y^\dagger a_x = \beta(y, x) 1$.
3. $a_x a_{a_y h}^\dagger + a_{a_y h}^\dagger a_x = q(x) 1$
4. Let $D = a_{a_y h}^\dagger a_x^\dagger$, then $D = a_{D(1)}^\dagger$.
5. $a_y a_x a_{a_y h}^\dagger = \beta(y, x) a_y$ and $a_{a_y h}^\dagger a_x a_y^\dagger = \beta(x, y) a_y^\dagger$.
6. $a_y a_x^\dagger a_z + a_z^\dagger a_y = \beta(y, x) a_z + \beta(z, x) a_y$ and $a_{a_y h} a_x a_{a_y h}^\dagger + a_{a_y h} a_x^\dagger a_y = \beta(x, z) a_y^\dagger + \beta(x, y) a_z^\dagger$
7. $a_y a_x^\dagger a_z + a_z^\dagger a_y = \beta(y, x) a_z + q(x) a_y$ and $a_{a_y h} a_x a_z^\dagger + a_{a_y h} a_x^\dagger a_y = q(x) a_y^\dagger + \beta(x, y) a_z^\dagger$
8. $[a_y a_x, a_{a_y h}^\dagger a_y] = \beta(x, y) a_{a_y h}^\dagger - \beta(y, x) a_y a_{a_y h}^\dagger$
9. $[a_x a_y, a_{a_y h}^\dagger a_y] = \beta(x, y) a_{a_y h}^\dagger a_x - \beta(y, x) a_x a_{a_y h}^\dagger + \beta(y, x)^2 - \beta(x, y)^2$.

**Proof.**

1. For any $w \in \wedge^\bullet V$, we have

$$ a_x a_h^\dagger w = a_x h \wedge w + (-1)^{\delta h} h \wedge a_x w = a_{a_x h}^\dagger w + (-1)^{\delta h} a_{a_x h}^\dagger a_x w. $$
2. Special case of (1) when \( h = y \).
3. Special case of (2) when \( y = x \).
4. Multiply (1) by \( a_h^\dagger \) on the left.
5. Multiply (2) by either \( a_x \) or \( a_y^\dagger \).
6. By linearizing (5) in \( y \).
7. Special case of (6) when \( z = x \).
8. Multiply the first equation from (7) by \( a_y^\dagger \) on the left, the second by \( a_y \) on the right, and subtract.
9. Combine (8) and (2).

Lemma 2.1.2. For any \( x, y \in V \) such that \( \beta(x, y) \neq 0 \) we have

\[
\wedge^\bullet V = \text{im } a_x \oplus \text{im } a_y^\dagger.
\]

Furthermore \( \text{im } a_x = \ker a_x \) and \( \text{im } a_y^\dagger = \ker a_y^\dagger \).

Proof. Consider any \( z \in \wedge^\bullet V \); then \( \beta(x, y)z = a_x a_y^\dagger z + a_y a_x z \in \text{im } a_x + \text{im } a_y^\dagger \) by Lemma 2.1.1 (2). Now assume \( u \in \text{im } a_x \cap \ker a_y^\dagger \).

Then \( u = a_x w \) for some \( w \) and \( a_y^\dagger u = 0 \). But then \( 0 = a_y^\dagger a_x w = a_x a_y^\dagger a_x w = \beta(x, y) a_x w = \beta(x, y) u \) by Lemma 2.1.1 (5), thus \( u = 0 \).

From \( \wedge^\bullet V = \text{im } a_x + \text{im } a_y^\dagger \), \( \text{im } a_y^\dagger \leq \ker a_y^\dagger \) and \( \text{im } a_x \cap \ker a_y^\dagger = 0 \) we conclude that \( \wedge^\bullet V = \text{im } a_x \oplus \text{im } a_y^\dagger \) and \( \ker a_y^\dagger = \text{im } a_y^\dagger \). Next, assume \( u \in \ker a_x \cap \text{im } a_y^\dagger \). Then \( u = a_y^\dagger w \) for some \( w \) and \( a_x u = 0 \). But then \( 0 = a_y^\dagger a_x a_y^\dagger w = \beta(x, y) a_y^\dagger w = \beta(x, y) u \) by Lemma 2.1.1 (5), thus \( u = 0 \). And similarly as before, this implies \( \text{im } a_x = \ker a_x \). \[ \square \]

Remark 2.1.3. The equality \( \text{im } a_y^\dagger = \ker a_x^\dagger \) is independent of \( \beta \) and therefore always valid. The equality \( \text{im } a_x = \ker a_x \) is valid if and only if \( \beta(x, y) \neq 0 \) for some \( y \); i.e. when \( \beta \) is non-degenerate. In fact \( \beta(x, y) = 0 \) for all \( y \) implies \( a_x = 0 \).

Definition 2.1.4. The Dirac operator \( \partial \in E(V) \) is defined by \( \partial(h) = (\delta h) \cdot h \) for all homogeneous \( h \) of degree \( \delta h \). The operator \( \partial \) is a derivation of degree 0, because

\[
\partial(g \wedge h) = (\delta g + \delta h)(g \wedge h)
= (\delta g)(g \wedge h) + (\delta h)(g \wedge h)
= \partial g \wedge h + g \wedge \partial h.
\]
The next two lemmas show that we have access to it in terms of $a^\dagger$ and $a$, provided $\beta$ is non-degenerate.

**Lemma 2.1.5.** For all $v, w \in V$, the operator $a^\dagger v a_w \in E(V)_0$ is a derivation of degree 0. In particular, it is completely determined by its restriction to $V$.

**Proof.** Consider arbitrary homogeneous $g, h \in \wedge V$. Then
\[
(a^\dagger v a_w)(g \wedge h) = a^\dagger v (a_w g \wedge h) + (-1)^\delta g a^\dagger v (g \wedge a_w h) = a^\dagger v a_w g \wedge h + g \wedge a^\dagger v a_w h.
\]

**Lemma 2.1.6.** Let $B = (b_1, \ldots, b_n)$ be an ordered basis for $V$. Then there exist constants $c_{ij}$ such that $\partial = \sum_{ij} c_{ij} a_i^\dagger a_j$ if and only if $\beta$ is non-degenerate. If $B$ is an orthogonal basis then $\partial = \sum_i \frac{1}{q(b_i)} a_i^\dagger a_i$.

**Proof.** For given constants $c_{ij}$ the operator $d = \sum_{ij} c_{ij} a_i^\dagger a_j$ is entirely determined by its values on $V$ by Lemma 2.1.5. Since $\partial$ is a derivation as well, we have $d = \partial$ if and only if $d(b) = b$ for all $b \in B$. If we denote $\beta_{ij} = \beta(b_i, b_j)$ this is equivalent to the equations
\[
\sum_{ij} b_i c_{ij} \beta_j \ell = b_\ell.
\]
In other words the constants $c_{ij}$ must be solutions to the system $\sum_j c_{ij} \beta_j \ell = \delta_{\ell j}$. This happens precisely when the matrix $(\beta_{ij})$ is invertible with inverse given by $(c_{ij})$. The last statement follows immediately from this observation.

If $e \in E(V)$ is an operator, its *expected value* is defined as $\langle e \rangle = \langle e(1) \rangle_0$. We extend $\beta(\cdot, \cdot)$ to all of $\wedge V$ by setting for all $x, y \in \wedge V$
\[
\tilde{\beta}(x, y) = \langle a_x a_y^\dagger \rangle = \langle a_x y \rangle_0.
\]

**Proposition 2.1.7.** The inner product $\tilde{\beta}$ satisfies
\[
\tilde{\beta}(v_1 \wedge \ldots \wedge v_\ell, w_1 \wedge \ldots \wedge w_m) = \begin{cases} 
0 & \text{if } \ell \neq m \\
\det(\beta(v_i, w_j))_{i,j=1}^\ell & \text{if } \ell = m.
\end{cases}
\]
Chapter 2. Linear algebra

Proof. We can expand $a_{v_\ell} \ldots a_{v_1} a^\dagger_{w_1} \ldots a^\dagger_{w_m}$ by repeatedly using the commutation relation Lemma 2.1.1 (1) until all the $a^\dagger$ operators are in front. Since $a_x(1) = 0$, the expected value is non-zero only when $\ell = m$ and must come from all terms of the form

$$
\prod_{i=1}^{n} (-1)^{\sigma} \beta(v_i, w_{\sigma(i)})
$$

for every $\sigma \in S_n$ and the sum of these is one way to define the determinant $\det(\beta(v_i, w_j)_{i,j=1}^\ell)$.

Corollary 2.1.8. $\hat{\beta} = \hat{\beta}^\circ$, in particular $\hat{\beta}$ is symmetric if and only if it is symmetric on $V$; if $\beta$ admits an orthogonal basis then $\hat{\beta}$ is non-degenerate on $\wedge^i V$ if and only if $\beta$ is non-degenerate.

Proof. The first assertion is trivial. For the second, let $b_1, \ldots, b_v$ be the orthogonal basis and denote $\beta(b_j, b_k) = \beta_{jk}$ so that $\beta$ is degenerate if (an only if) $\beta_{jj} = 0$ for some $j$. Then the set of vectors $b_{j_1} \wedge \ldots \wedge b_{j_i}$ where $\{j_1, \ldots, j_i\}$ runs through subsets of $\{1, \ldots, \dim V\}$ of order $i$ is an orthogonal basis for $\wedge^i V$ and $\hat{\beta}$ is degenerate on $\wedge^i V$ if $\beta_{j_1j_1} \ldots \beta_{j_ij_i} = 0$ for some subset and thus $\beta_{ii} = 0$ for some $i$.

Problem 2.1.9. Is the second assertion still true if $\beta$ does not admit an orthogonal basis?

So not only does $\hat{\beta}$ extend $\beta$, it also shares its most important properties. This allows us to suppress the notation $\hat{\beta}$ and simply use $\beta$ for both.

We will now observe that with respect to $\beta$, the operators $a$ and $a^\dagger$ are adjoint, which justifies the notation. In other words, the next proposition implies that $(a_x)^\dagger = a^\dagger_x = (a^\dagger)_x$ may be interpreted either as the adjoint operator of the contraction $a_x$ or as the operator which is left multiplication by $x$, without ambiguity.

Proposition 2.1.10. Whenever $v, w, z \in \wedge^* V$, we have $\beta(v, a_w z) = \beta(a^\dagger_w v, z)$

Proof.

$$
\beta(v, a_w z) = \langle a_v a_w a^\dagger_z \rangle = \langle a_w \wedge v a^\dagger_z \rangle
$$
\[ = \beta(w \wedge v, z) = \beta(a^\dagger_w v, z). \]

2.2 Attack of the Clifford algebra

A bilinear form \( \beta \) gives rise to a quadratic form \( q_\beta \), denoted simply by \( q \) if there is no confusion, as follows:

\[ q_\beta : V \to k : x \mapsto q_\beta(x) = \beta(x, x). \]

Every quadratic form \( q \) arises this way.

To the vector space \( V \) endowed with the quadratic form \( q \) corresponds a Clifford algebra \( (C\ell(V, q), \iota) \), where \( C\ell(V, q) \) is a unital associative algebra and \( \iota : V \to C\ell(V, q) \) is an embedding of vector spaces which satisfies \( \iota(x)^2 = q(x)1 \). We will not repeat the construction and universal property here, but we recall that on the Clifford algebra there are also a natural \( \mathbb{N} \)-filtration, a natural \( \mathbb{Z}/2\mathbb{Z} \)-grading and an opposition \( x \mapsto x^{\text{op}} \). We now recall the following important fact.

**Proposition 2.2.1.** The map \( m : V \to \text{End}(\wedge^\bullet V) : x \mapsto m_x = a^\dagger_x + a_x \) extends to an algebra morphism \( m : C\ell(V, q) \to \text{End}(\wedge^\bullet V) \). The map \( \psi : C\ell(V, q) \to \wedge^\bullet V \) given by \( \psi(x) = m_x(1) \) is an isomorphism of vector spaces; in particular \( m \) is injective.

**Proof.** The first statement is a consequence of the universal property of \( C\ell(V, q) \) and

\[ (a^\dagger_x + a_x)^2 = a^\dagger_x a_x + (a^\dagger_x a_x + a_x a^\dagger_x) + a_x a_x = 0 + \beta(x, x)1 + 0. \]

For the next statement, we will show by induction on \( i \) that the induced maps \( \psi_i : C\ell^{\leq i}(V, q) \to \wedge^{\leq i}V \) are bijective. The case \( i = 0 \) is obvious. If we take an arbitrary homogeneous element \( x = x_1 \wedge \ldots \wedge x_i \in \wedge^i V \) and let \( \rho : V \to C\ell(V, q) \) be the canonical map, we observe that:

\[ \psi_i(\rho(x_1) \ldots \rho(x_i)) = m_{x_1} m_{x_2} \ldots m_{x_i}(1) \equiv x_1 \wedge x_2 \wedge \ldots \wedge x_i \pmod{\wedge^{i-1} V}, \]
as can be seen by expanding the product of the $m_{x_i} = \alpha_{x_i}^* + \alpha_{x_i}$ and observing that only the term $\alpha_{x_1}^* \ldots \alpha_{x_i}^* 1$ will not necessarily end up in $\land^{<i}(V)$. By induction $\psi_i(x_1 \ldots x_i) - x \in \text{im } \psi_{i-1}$, thus $x \in \text{im } \psi_i$ and $\psi_i$ is surjective. Furthermore, if $\psi_i(u) = 0$, for some $u \in \mathcal{C}^\leq i V$, then by expanding $u$ as a polynomial in $\rho(b_i)$ for $b_i$ in some ordered basis $B$, it is clear that the terms of degree $i$ will have non-zero image if they are present. Thus $u \in \mathcal{C}^\leq i - 1$, so that $u = 0$ by the induction hypothesis. Therefore $\psi_i$ is an isomorphism and so is $\psi$.}

There are far better proofs available in the literature, for instance we accidentally stumbled upon [Bas74, (2.4)] where Bass proves there is a natural isomorphism of functors

$$\mathcal{C} \circ \mathbb{H} \simeq \text{END} \circ \land^*,$$

which implies Proposition 2.2.1 via the standard embedding of quadratic spaces

$$(V, q) \to \mathbb{H}(V, q) = (V \perp V, q \perp (-q)) : v \mapsto (v, 0).$$

(Care is needed in the definition of the categories under consideration and in the definition of END as a functor, see [Bas74, (1.2)].)

### 2.3 Revenge of the characteristic

From here until the end of Part I we assume $\text{char } k \neq 2$.

Recall that $\beta$ was a bilinear form and $q$ a quadratic form such that $q(x) = \beta(x, x)$, in particular

$$q(x + y) = q(x) + q(y) + \beta(x, y) + \beta(y, x).$$

But there is now a canonical choice available for $\beta$, which is the symmetric bilinear form $\frac{1}{2}b_q(x, y) = \frac{1}{2}(q(x + y) - q(x) - q(y))$. In this case, we may extend $q$ to all of $\land^* V$ by setting $\tilde{q}(u) = \tilde{\beta}(u, u)$ whenever $u \in \land^* V$.

From here until the end of Part I, $\beta$ is a non-degenerate symmetric bilinear form and Notation 2.3.1 applies.
2.3. Revenge of the characteristic

Notation 2.3.1. We introduce the notations $x \mathbin{\triangledown} y := a_x y$ for the contraction seen as a product
\[
\wedge^i V \otimes \wedge^{i+j} V \to \wedge^j V,
\]
and $x \cdot y = \beta(x, y) = \langle x \cdot y \rangle_0$ for the inner product whenever $x, y \in \wedge^\bullet V$.

We let the contraction associate on the left: $x \mathbin{\triangledown} y \mathbin{\triangledown} z = (x \mathbin{\triangledown} y) \mathbin{\triangledown} z$. For instance, we have:
\[
(x \wedge y) \mathbin{\triangledown} z = y \mathbin{\triangledown} (x \mathbin{\triangledown} z),
\]
\[
(x \wedge y) \cdot z = y \cdot (x \mathbin{\triangledown} z).
\]
These notations will make some computations more transparent. (And others more obscure.)

For instance, an important notion in the theory of quadratic forms is the Lie algebra $\mathfrak{so}(q)$ of matrices $D \in \text{End}(V)$ that are skew-symmetric with respect to the symmetric bilinear form $\beta$:

Definition 2.3.2. The Lie algebra $\mathfrak{so}(q)$ is the vector space
\[
\{ D \in \text{End}(V) \mid \beta(Dx, y) + \beta(x, Dy) = 0 \},
\]
endowed with the commutator bracket $[D, D'] = DD' - D'D$.

One of the merits of our notation is that we can provide a canonical isomorphism between $\wedge^2 V$ and $\mathfrak{so}(q)$ without too much trouble.

Proposition 2.3.3. There is an isomorphism $\wedge^2 V \to \mathfrak{so}(q)$, given by $t \mapsto D_t$, where $D_t : V \to V$ is given by $D_t(z) = a_z t$.

Proof. We first show that indeed $D_t \in \mathfrak{so}(q)$, i.e. that $\beta(x, D_t y) + \beta(D_t x, y) = 0$. But since $\beta(x, a_y t) + \beta(y, a_x t) = \beta(y \wedge x, z) + \beta(x \wedge y, z) = 0$ this is clear. Furthermore, injectivity is trivial and therefore the map is an isomorphism by a dimension count.

Remark 2.3.4. This gives $\wedge^2 V$ the structure of a Lie algebra which satisfies
\[
x \mathbin{\triangledown} u \mathbin{\triangledown} v - x \mathbin{\triangledown} v \mathbin{\triangledown} u = x \mathbin{\triangledown} [u, v], \forall x \in V.
\]
In the case where $\text{char } k = 2$, the proposition becomes false because we had to rely on the symmetry of the bilinear form. One can still define a
Lie algebra structure on $\wedge^2 V$, by identifying it with $\mathcal{O}^\leq 2 V/\mathcal{O}^\leq 1 V = \mathcal{O}^\leq 2 V/\langle 1 \rangle$, where $\mathcal{O}_0$ denotes the even part of the Clifford algebra. One can verify that this space is closed under the commutation bracket $[xy] = xy - yx$ of the Clifford algebra, and that the linear subspace generated by 1 is an ideal of square 0. This is related to Lemma 2.1.1 (9), where we saw that the space of operators $a_+_a$ is closed under the commutator bracket if $\beta$ is symmetric, but otherwise we must include 1 in the space to close it under the commutator.

**Problem 2.3.5.** The reader can try to verify that both products coincide by proving the formula

$$x \cdot u \cdot v - x \cdot \frac{uv}{2} = \frac{1}{4}(u^{op}xv + v^{op}xu), \quad x \in V, u, v \in \wedge^2 V,$$

and noting that the right hand side is symmetric in $u$ and $v$. (I just discovered this equation in my old notes but it was stated without proof...)

**Problem 2.3.6.** A major problem the theory as we develop it in Part I is that it does not work in characteristic 2. On the one hand, octonion algebras exist in characteristic 2 and the theory of composition algebras works more or less uniformly in all characteristics. On the other hand, even in characteristic 2 there are some oddities. Let us briefly recall what these are.

In characteristic $\neq 2$, Hurwitz’s theorem states that the dimension of a unital composition algebra $C$ is always 1, 2, 4 or 8 over the ground field. This is proven by relying on a construction, often called the *Cayley-Dickson (doubling) procedure*. This procedure allows you to take a subalgebra $C'$ of an arbitrary composition algebra $C$ and an element $d \in C \setminus C'$ and construct a larger subalgebra $\langle C', d \rangle$ with dimension twice that of $C'$. (Hence the name *doubling procedure.*) One can then start from the trivial subalgebra spanned by the identity $k \cdot 1_C$ and repeatedly apply the doubling procedure to find subalgebras of dimension 2, 4, 8,... But there is a catch: one can also prove that in each step the algebra will lose a nice property. In the first step, the involution that the composition algebra is endowed with becomes non-trivial—it becomes the Galois involution of a separable quadratic extension. In the second step, the algebra becomes non-commutative—think of Hamilton’s quaternions. In the third step,
the algebra becomes non-associative—as for Cayley’s octonions. But
then, the process must come to a halt, because the candidate algebras
of dimension 16 that are produced by the Cayley-Dickson procedure
are never composition algebras if one starts from a non-associative
algebra. This shows that we must have exhausted the algebra before
it came to that and so the dimension must be 1, 2, 4 or 8. For more
details, we recommend [McC04], I.2.8–I.2.12 and II.2.

In characteristic 2, this does not work so well. The trouble is that
as we go from the trivial subalgebra $k \cdot 1_C$ to a two-dimensional sub-
algebra, the Cayley-Dickson procedure will produce an inseparable
quadratic extension—with trivial involution. As a result, the process
can continue indefinitely, producing ever larger inseparable field ex-
tensions of degree 2, 4, 8, 16,... All these composition algebras have
trivial involution and the quadratic form is just the squaring operator
which takes values in the field $k$, identified with $k \cdot 1_C$. (The associated
quadratic form will be regular, but degenerate.) To rule out these
degenerate cases, we must bootstrap the Cayley-Dickson procedure
in dimension 2: if we assume there exists a subalgebra of dimension
2 with non-trivial involution then we can start doubling from there
and the conclusions are the same as in the case of characteristic $\neq 2$.

Another closely related issue with the case of characteristic 2 is that,
to the best of our knowledge, there is also nothing even close to the
‘Fano plane’-mnemonic for the product of octonions in characteristic 2.
Let us also explain this in some greater detail. Recall how Hamilton’s
quaternions $\mathbb{H}$ can be described by the equations $i^2 = j^2 = k^2 = i j k =
-1$, which Hamilton famously carved into Broom Bridge in Dublin,
as he happened to be in the neighbourhood without his notebook
when he made this discovery. A graphical depiction of this discovery
would be as an ordered line with three points $i$, $j$ and $k$ on it. The
convention is that $x^2 = -1$ for every point $x$ on the line and the
product of two points is the third point, with the sign provided by
the ordering.
We can read off from this diagram that \( ij = k, \; jk = i, \; ki = j; \) but \( ji = -k, \; ik = -j \) and \( kj = -i. \) A similar picture will also provide a mnemonic for the multiplication rules for Cayley’s octonions \( \mathbb{O}, \) but then we need to take 7 points, with a unique line through every 2 of them: the Fano plane.

From this, we can read off that \( e_0e_2 = e_6, \) but \( e_2e_0 = -e_6 \) because that goes against the direction of the arrow. There are many interesting insights to be gained from this way of representing an octonion algebra. For instance, every line will correspond to a quaternion subalgebra; since the plane is the union of three lines through a point, there ought to be a way to understand an octonion algebra as some kind of amalgam of three quaternion algebras which share a quadratic extension of the ground field. This is achieved by the representation of the octonion algebra as a Zorn matrix-vector algebra, which is a way to represent an octonion algebra as a \( 2 \times 2 \)-matrix:

\[
\begin{bmatrix}
x & \mathbf{v} \\
\mathbf{w} & y
\end{bmatrix}.
\]

Here \( \mathbf{v} \) and \( \mathbf{w} \) are themselves 3-dimensional vectors, with a specific way to multiply a pair of such vector-matrices. If one takes vectors \( \mathbf{v}, \; \mathbf{w} \) with a 0 in two out of three components, this reduces to a quaternion algebra. See [McC04, p. II.2.4] for details.
Unfortunately, the Fano plane approach to octonions dramatically fails if the characteristic is 2! In fact the whole approach is a disaster right from the start, since the points are supposed to represent an orthogonal basis for the orthogonal complement of the identity in the octonion algebra; but if the characteristic is 2, there is no orthogonal basis and everything fails.

The trouble starts already when one considers a single point, which is supposed to represent a quadratic extension of the ground field; in other words, the culprit is the absence of a canonical basis for separable field extensions $k[x]/(x^2 + ax + b)$ in characteristic 2, whereas one can always take $1, \sqrt{b^2 - 4a}$ if the characteristic is different from 2.

Nonetheless, we spent a great amount of time on trying to make the theory work uniformly in all characteristics. The issue is that one must untangle many of the notions that come into play.

- We use a non-degenerate bilinear form $\beta$ in two ways: once to define a quadratic form $q(x) = \beta(x, x)$ from Section 2.2 on, and once to identify $V$ with $V^*$ right off the bat. For instance, our contraction operators $a x y$ would be better off if they were maps $\wedge^i V^* \otimes \wedge^j V \to \wedge^{j-i} V$. So we should either untangle our use of $V$ (into $V^*$’s and $V$’s) or choose a fixed bijection $\beta^{\flat} : V \to V^*$ and make a choice of $q$ which is a priori not related to $\beta^{\flat}$.

- Although we don’t often use the identification $\mathcal{C}(V, q) \cong \wedge^\bullet V$ of vector spaces (Proposition 2.2.1), it suggests that we may often be using $\wedge^\bullet V$ where, morally speaking, we should be using $\mathcal{C}(V, q)$. It appears we can still make the identification in characteristic 2, by choosing an arbitrary (non-symmetric!) bilinear form $\beta$ such that $q(x) = \beta(x, x)$, and we can still extend $\beta$ to $\wedge^\bullet V$, but the trouble is that $\hat{\beta}(u, u)$ will then depend on our choice of $\beta$ and not only on $q$ so this approach seems to meet a dead end. The proper way around it is probably to replace the quadratic form on $\wedge^\bullet V$ by the map

$$\mathcal{C}(V, q) \to \mathcal{C}(V, q) : x \mapsto xx^{\text{op}}.$$ 

If the characteristic is $\neq 2$, thanks to the canonical identification $\wedge^\bullet V \cong \mathcal{C}(V, q)$ we may compose this with the projection onto
the degree 0-vectors $\wedge^0 V \to \langle 1 \rangle$ to recover the quadratic form; but in characteristic 2 this is impossible.

### 2.4 Fabulous multivectors

We will define the notion of *fabulosity* of a multivector $F \in \wedge^t V$ with respect to the quadratic form $q$. The main goal of this section is to point out many similarities between fabulous multivectors and block designs—see Remark 2.4.6 for a summary.

**Definition 2.4.1.** A multivector $F \in \wedge^t V$ is *$t$-fabulous* (with respect to $q$), with multiplier $\lambda \in k$, if $F$ is homogeneous and the condition $(C_t)$ holds:

$$\forall \text{ pure } x \in \wedge^t V : q(x \downarrow F) = \lambda q(x).$$

$(C_t)$

We will abbreviate the sentence “$F$ is fabulous with respect to $q$, with multiplier $\lambda$” by $C_t(q, F, \lambda)$.

**Notation 2.4.2.** If we study fabulous multivectors, we will always use the notations $v = \dim V$, $k = \delta F$, $\lambda_0 = q(F)$. For a fixed $t$-fabulous $F$, we will use the following notations, where all $a_i \in V$:

$$X(a_1, \ldots, a_t) = (a_1 \wedge \ldots \wedge a_t) \downarrow F \in V,$$

$$(a_1 \mid \ldots \mid a_{t+1}) = (a_1 \wedge \ldots \wedge a_{t+1}) \downarrow F \in k.$$

Both operations are related by

$$(a_1 \mid \ldots \mid a_{t+1}) = a_{t+1} \downarrow X(a_1, \ldots, a_t).$$

If $t = 2$, we will also write $a \times b = X(a, b)$, for $a, b \in V$.

**Remark 2.4.3.**

1. It is very important that the condition $(C_t)$ is only specified on pure vectors. In particular, if $x$ and $y$ are pure, the equation is in general not satisfied for $x + y$ and therefore cannot be linearized in a straightforward fashion. (Although $(C_t)$ can sometimes be linearized in a non-trivial manner, see Proposition 2.6.2.)
2. The fabulosity relation can also be written as follows:

$$\forall \text{ pure } x \in \land^t V : \langle a_F a_x^\dagger a_x a_F^\dagger \rangle = \lambda \langle a_x a_F^\dagger \rangle$$

In fact, our notations were to some extent inspired by the bra-ket notation $$\langle F | a_x^\dagger a_x | F \rangle = \lambda \langle x | x \rangle$$ used in physics.

3. The notation $$C_t(q, F, \lambda)$$ is loosely modeled after the notation $$S(t, k, v)$$ for a Steiner system although the order of the parameters is closer to the notation $$t - (v, k, \lambda)$$ for the parameters of a block design, see Remark 2.4.6.

4. Note that every multivector $$F$$ is 0-fabulous, with multiplier $$\lambda_0 = q(F)$$, since

$$q(a \downarrow F) = q(aF) = a^2 q(F) = q(a)q(F),$$

whenever $$a \in \land^0 V \simeq k$$.

Let us first describe the behaviour of fabulosity under rescaling of $$q$$ and $$F$$.

**Proposition 2.4.4.** Whenever $$\mu, \nu \in k^\times$$, we have

$$C_t(q, F, \lambda) \iff C_t(\mu q, \nu F, \nu^2 \mu^{\delta F} \lambda)$$

**Proof.** Let $$\dot{q} = \mu q$$ for some $$\mu \in k$$. Then we have, for homogeneous elements the equalities

$$\dot{q}(u) = \mu^{\delta u} q(u) \quad \text{and} \quad \dot{a}_u v = \mu^{\delta u} a_u v.$$ 

Therefore if we also denote $$\dot{F} = \nu F$$, then

$$\dot{q}(\dot{a}_x \dot{F}) = \mu^{\delta_F - \delta x} \nu^2 \mu^{2\delta x} q(a_x F)$$

$$= \mu^{\delta_F + \delta x} \nu^2 \lambda q(x)$$

$$= \mu^{\delta F} \nu^2 \lambda \dot{q}(x),$$

which implies that if $$\dot{\lambda} = \mu^{\delta F} \nu^2 \lambda$$, then $$C_t(\dot{q}, \dot{F}, \dot{\lambda}).$$

**Remark 2.4.5.**

1. We would like to know whether for given values of $$v = \dim V$$, $$k = \delta F$$ and $$t$$, there exists a fabulous multivector $$F$$. With that in mind, it makes sense to assume that $$k$$ is algebraically closed and by the previous proposition, we may as well assume that $$\lambda = 1$$ or $$\lambda = 0$$. 


2. In fact, even if $k$ is not algebraically closed but $\delta F$ is odd, one can still freely rescale $\lambda$; but when $\delta F$ is even, $\lambda$ can only be rescaled by a square.

3. It may very well happen that a multivector $F$ is $t$-fabulous and $t'$-fabulous for $t \neq t'$; in that case the corresponding multipliers are denoted $\lambda_t$, $\lambda_{t'}$. One can of course not rescale both variables at the same time.

4. Furthermore, upon rescaling as in the previous proposition, $\lambda_0$ scales to $\dot{q}(\dot{F}) = v^2 \mu \delta F q(F)$, so the value of $\lambda_t/\lambda_0$ and more generally the proportions of $\lambda_t$ to $\lambda_{t'}$ (whenever both are defined), are invariant under rescaling.

5. Our main case of interest we will a 2-fabulous trivector, i.e. $t = 2$ and $k = 3$. We will also restrict ourselves to the case $\lambda_2 \neq 0$, which allows us to assume that $\lambda_2 = 1$. But in fact, $F$ will also be 1-fabulous by Proposition 2.4.7 and moreover $\lambda_1/\lambda_2 = 3$. It is this 3 which eventually causes groups of type $G_2$ to behave differently in characteristic 3, see Remark 2.5.5.

**Remark 2.4.6.** The following propositions will explore the analogy between fabulous multivectors and block designs. Recall that a $t-(v, k, \lambda)$ block design $(\mathcal{P}, \mathcal{B}, I)$ consists of a set $\mathcal{P}$ of points, a set $\mathcal{B}$ of blocks and a relation $I \subseteq \mathcal{P} \times \mathcal{B}$ which declares which point lies on which blocks with the following properties: (i) there are $v$ points; (ii) every block contains exactly $k$ points; (iii) through every $t$ points there are exactly $\lambda$ blocks. One also defines the constant $b$, which is the number of blocks and the constant $r$ which is the number of blocks through a point. The analogy between the numerical constants that appear for fabulous multivectors is summarized in the following table.

<table>
<thead>
<tr>
<th>multivector</th>
<th>block design</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t$</td>
<td>$t$</td>
</tr>
<tr>
<td>$v = \dim V$</td>
<td>$v$</td>
</tr>
<tr>
<td>$k = \delta F$</td>
<td>$k$</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>$\lambda$</td>
</tr>
<tr>
<td>$q(F)$</td>
<td>$b = \lambda_0$</td>
</tr>
<tr>
<td>$\lambda_1$</td>
<td>$r = \lambda_1$</td>
</tr>
</tbody>
</table>
One can easily show that a \((t + 1) - (v, k, \lambda)\)-design is always also a \(t - (v, k, \lambda')\)-design, with \(\lambda'(k - t) = \lambda(v - t)\). We will show in Proposition 2.4.7 that this property is typically also true for fabulous multivectors. (Unless something really strange happens for which we know no examples—although we also did not look for them.)

In particular every block design is also a \(1\)-design, so the constant \(r\) is always well defined and satisfies the relation \(bk = rv\). In Corollary 2.4.8, we will see this equation also holds for \(1\)-fabulous multivectors.

Furthermore, for a given \(t - (v, k, \lambda)\)-design and a set with \(s \leq t\) points, there is a derived \((t - s) - (v - s, k - s, \lambda)\) block design of blocks containing this set. In Proposition 2.4.9, we will show that there is an analogous operation for fabulous multivectors, in fact also given by a derivation but in the algebraic sense of the word.

**Proposition 2.4.7.** Let \(C_{t+1}(q, F, \lambda_{t+1})\)—see Definition 2.4.1. Then (at least) one of the following is true.

- \(C_t(q, F, \lambda_{t+1}\frac{(v-t)}{(k-t)})\);
- \(k - t = \lambda_{t+1}(v - t) = 0 \) (in \(k\)).

**Proof.** Recall that fabulosity states that

\[
\forall \text{pure } x \in \wedge^{t+1}V : \langle a_F a_x a_x a_F^\dagger \rangle = \lambda_{t+1}\langle a_x a_x^\dagger \rangle.
\]

So taking an orthogonal basis \(b_1, \ldots, b_n\) for \(V\) and \(y \in \wedge^t V\) pure, we have:

\[
\langle a_F a_y a_y^\dagger \frac{1}{q(b_i)} a_{b_i} a_y a_F^\dagger \rangle = \lambda_{t+1}\langle a_y a_{b_i} a_y a_{b_i} a_y a_F^\dagger \rangle - \lambda_{t+1}\langle a_y a_{b_i} a_y a_{b_i} a_y \rangle.
\]

Taking the sum over the \(b_i\), we recognise the Dirac operator from Definition 2.1.4 and Lemma 2.1.6 and get, with \(v = \dim V\)

\[
\langle a_F a_y^\dagger \partial a_y a_F^\dagger \rangle = v\lambda_{t+1}\langle a_y a_y^\dagger \rangle - \lambda_{t+1}\langle a_y \partial a_y^\dagger \rangle,
\]

and thus, with \(k = \delta F\):

\[
\forall \text{pure } y \in \wedge^t V : (k - t) q(y \cup F) = \lambda_{t+1}(v - t) q(y). \quad (X)
\]
So either $k - t$ is non-zero and we are in the first case, or $k - t = 0$ and then we can certainly choose $y$ such that $q(y) \neq 0$ to conclude that we are in the second case.

**Corollary 2.4.8.** If $C_1(q, F, \lambda)$, then $q(F)k = \lambda_1 v$.  

*Proof.* Apply Equation X from the previous proposition to the case $t = 0$.  

**Proposition 2.4.9.** Let $C_t(q, F, \lambda)$ with $\delta F = k$. Let $x = e_1 \wedge \ldots \wedge e_s \in \wedge^s V$ be pure and define $V' = \langle e_1, \ldots, e_s \rangle^\perp$. Then 

$$C_{t-s}(q|_{V'}, a_x F, \lambda q(x)).$$

*Proof.* Let $G = a_x F$. Take any pure $y \in \wedge^{t-s} V'$. Then $y \wedge x \in \wedge^s V$ is pure and $q(x \wedge y) = q(x)q(y)$. So we compute 

$$q(a_y G) = q(a_y a_x F) = q(a_{x \wedge y} F) = \lambda q(x \wedge y) = \lambda q(x)q(y).$$

**Problem 2.4.10.** We would really like to make the link between block designs and fabulous multivectors more explicit. For instance, we would certainly be interested in results of the type 

block design + good property $\implies$ n-fabulous.

Proposition Blueprint 2.4.11 is a meager attempt at characterizing 1-fabulosity with respect to the standard inner product. A better result could for instance start from the Fano plane together with a cyclic group of automorphisms which acts regularly on the blocks and produce a multivector which is 2-fabulous with respect to a standard inner product.

**Proposition Blueprint 2.4.11.** Let $D = (\mathcal{P}, \mathcal{B})$ be a block design with a group of automorphisms $\Gamma \leq \text{Aut}(D)$ such that 

1. $\Gamma$ acts transitively on the blocks.
2. The induced action of the stabilizer $\Gamma_B$ on the block $B$ is contained in the alternating group.
3. For every $k - 1$ points there is at most one block containing them
4. *(Some Unknown Condition)* holds.

Then this determines a 1-fabulous multivector.

**Proof.** Consider a vectorspace $V$ over a field $k$, with a basis $e_p, p \in P$. Choose a block $B$ and order the points somehow: $(b_1, b_2, \ldots)$. Define

$$F = \sum_{\gamma \in \Gamma/\Gamma_B} \bigwedge_{i=1}^{k} e_{\gamma(b_i)}.$$  

The sum is well defined by condition (2) and, moreover $q(F)$ is the number of blocks by condition (1). For every vector $b_i$, it is clear that $b_i \cdot F \cdot F = c_i b_i$ for some scalar $c_i$ by condition (3); moreover $c_i = c_j$ for all $i$ and $j$ by Some Unknown Condition (4) which implies that $F$ is 1-fabulous by Proposition 2.5.2 (5). \hfill \Box

### 2.5 1-fabulosity

1-fabulosity, or ordinary fabulosity, is a linear condition and therefore quite easy to understand: it says that the map $x \mapsto x \cdot F$ preserves orthogonality. We will prove this, and a number of other characterizations, in Proposition 2.5.2, relying on the following simple lemma.

**Lemma 2.5.1.** For homogeneous $x, y, F, G$:

$$(x \cdot F) \cdot (y \cdot G) = (-1)^{\delta x(\delta G - \delta y)} x \cdot (y \cdot G \cdot F).$$

**Proof.**

$$
(x \cdot F) \cdot (y \cdot G) = F \cdot (x \wedge (y \cdot G)) \\
= (-1)^{\delta x(\delta G - \delta y)} F \cdot ((y \cdot G) \wedge x) \\
= (-1)^{\delta x(\delta G - \delta y)} (y \cdot G \cdot F) \cdot x \quad \square
$$

**Proposition 2.5.2.** Let $F \in \wedge^k V$, then the following are equivalent.

1. $C_1(q, F, \lambda_1)$
2. $\exists \lambda \in k : \forall x, y \in V : (x \cdot F) \cdot (y \cdot F) = \lambda x \cdot y$
3. $\forall x, y \in V : x \perp y \implies (x \cdot F) \perp (y \cdot F)$
4. \( \forall x \in V : \exists \lambda_x \in k : x \perp F \perp F = \lambda_xx \)

5. \( \exists \lambda \in k : \forall x \in V : x \perp F \perp F = \lambda x \)

**Proof.** (1 \( \implies \) 2) By linearization. (2 \( \implies \) 3) Immediate. (3 \( \implies \) 4) Take \( y \in \wedge^2 V \) arbitrary and set \( z = a_x y \). Then \( z \cdot x = y \cdot a_x^\dagger x = 0 \), i.e. \( z \perp x \). Then by Lemma 2.5.1

\[
(x \wedge (x \perp F \perp F)) \cdot y = (x \perp F \perp F) \cdot z = \pm (x \perp F) \cdot (z \perp F) = 0.
\]

Since \( y \in \wedge^2 V \) was arbitrary, \( x \wedge ((x \perp F) \perp F) = 0 \), so \( x \) and \( (x \perp F) \perp F \) are proportional. (4 \( \implies \) 5) If \( x \) and \( y \) are linearly dependent, say \( y = \mu x \) with \( \mu \) non-zero, then

\[
\lambda_y \mu x = y \perp F \perp F = \mu \lambda_x x,
\]

and thus \( \lambda_x = \lambda_y \). If \( x \) and \( y \) are linearly independent then

\[
\lambda_x x + \lambda_y y = (x + y) \perp F \perp F = \lambda_{x+y}(x + y),
\]

and thus \( \lambda_x = \lambda_{x+y} = \lambda_y \). (5) \( \implies \) (1) This follows by taking the inner product with \( x \) and applying Lemma 2.5.1 again. \( \Box \)

There is also the following formulation, which we pay special attention to because it does not require the \( a \)-operator in a direct way, only the extension of \( q \) to \( \wedge^* V \) is required. (This could be relevant in the context of Problem 2.3.6.)

**Proposition 2.5.3.**

\[
C_1(q, F, \lambda_1) \iff \forall \text{pure } x \in \wedge^t V : q(x \wedge F) = \tilde{\lambda}_1 q(x),
\]

where \( \lambda_1 + \tilde{\lambda}_1 = \lambda_0 \).

**Proof.** This follows from

\[
q(x \wedge F) + q(x \perp F) = \langle a_F(a_x a_x^\dagger + a_x^\dagger a_x) a_F^\dagger \rangle = q(x)q(F) = \lambda_0 q(x).
\]

Finally, we want to draw attention to the fundamental dichotomy between the case where \( \lambda = 0 \) and \( \lambda \neq 0 \).
Proposition 2.5.4 (Fundamental dichotomy). Let $F$ be 1-fabulous with $k = \delta F$, and denote
\[
\alpha : \wedge^{k-1}V \to V : x \mapsto x \updownarrow F,
\]
\[
\beta : V \to \wedge^{k-1}V : x \mapsto x \updownarrow F.
\]

Then:
1. If $\lambda_1 \neq 0$: $\wedge^{k-1}V = \text{im } \beta \oplus \ker \alpha$; $\frac{1}{\lambda_1} \alpha \circ \beta = \text{id}_V$ and $\frac{1}{\lambda_1} \beta \circ \alpha$ is a projection operator $\wedge^{k-1}V \to \text{im } \beta$;
2. If $\lambda_1 = 0$: $\text{im } \beta \leq \ker \alpha$; $\alpha \circ \beta = 0$ and $(\beta \circ \alpha)^2 = 0$.

Proof. Note that by Proposition 2.5.2 (5), $\alpha \circ \beta = \lambda_1 \text{id}_V$. Moreover if $\lambda \neq 0$, then by Proposition 2.5.2 (2) $\beta$ is injective. The rest is straightforward to check. \hfill \square

Remark 2.5.5. The name fundamental dichotomy is borrowed from Wilson [Wil10], who uses an instance of it to give a 2-page construction for the Ree groups of type $^2G_2$. Explaining this dichotomy on a deeper level has been a main source of motivation for this work, so what this shows is that the 3 which is responsible for the constructions specific to $G_2$ in characteristic 3 can be traced back to the fact that $\lambda_1 = 3$, because in the Fano plane there are 3 lines through every point. In most characteristics, $\wedge^2V = \text{im } \beta \oplus \ker \alpha$ corresponds to the unique decomposition of the Lie algebra $\mathfrak{so}_7 = \mathfrak{g}_2 \oplus V$ into irreducible representations for $G_2 \subseteq \text{SO}_7$ (see also [FH91, §22.3, p.353]) but in characteristic 3 there is an inclusion $V \leq \mathfrak{g}_2$. More on this in Section 5.2.

2.6 2-fabulosity

Notation 2.6.1. Every 2-fabulous vector $F \in \wedge^kV$ induces a product $V \wedge V \to \wedge^{k-2}V$, given by
\[
x \times y = (x \wedge y) \updownarrow F, \quad x, y \in V.
\]
This product is anti-symmetric, and has the property that
\[
q(x \times y) = \lambda_2 q(x \wedge y) = \lambda_2 \begin{vmatrix} x \cdot x & x \cdot y \\ y \cdot x & y \cdot y \end{vmatrix}.
\]
Recall also the notations from 2.4.2; with in particular for $k = 3$ the identity $(x \mid y \mid z) = (x \times y) \cdot z$—this immediately proves that this last product is alternating, which is a familiar property of the cross product.

Recall that the condition of 2-fabulosity was non-linear: it applies only to the pure multivectors of degree 2. It is a bit remarkable that this shortcoming can be repaired by adding a term which vanishes on the pure vectors, as we will now show.

**Proposition 2.6.2.** Let $F \in \wedge^k V$ be 2-fabulous. Then there is a unique $T \in \wedge^4 V$ such that

\[
\forall u, v \in \wedge^2 V : (u \wedge F) \cdot (v \wedge F) - \lambda_2 (u \cdot v) = (u \wedge v) \cdot T,
\]

\[
\forall u \in \wedge^2 V : q(u \wedge F) = \lambda_2 q(u) + (u \wedge u) \cdot T
\]

\[
\forall u \in \wedge^2 V : u \wedge F \wedge F - \lambda_2 u = u \wedge T.
\]

If $u \in \wedge^2 V$ is pure, then $u \wedge u = 0$, so the second condition immediately generalizes the fabulosity condition ($C_t$).

**Proof.** Consider the map

\[
T : V^4 \to k
\]

\[
: (x_1, x_2, y_1, y_2) \mapsto (x_1 \times x_2) \cdot (y_1 \times y_2) - \lambda_2 (x_1 \wedge x_2) \cdot (y_1 \wedge y_2).
\]

We will show that this map is anti-symmetric. Note that $T(x, y, x, y) = 0$ because $F$ is 2-fabulous by assumption. So

\[
T(x_1, y, x_2, y) = -T(x_2, y, x_1, y).
\]

But by the symmetry of the bilinear form, this is again equal to

\[
-T(x_1, y, x_2, y)
\]

which implies $T(x_1, y, x_2, y) = 0$, i.e. $T$ is anti-symmetric under the permutation $(2\,4)$. But since we also have that

\[
T(x_1, x_2, y, y) = T(x, x, y_1, y_2) = 0,
\]

and since the permutations $(1\,2)$, $(2\,4)$ and $(3\,4)$ generate the symmetric group $S_4$, $T$ is anti-symmetric as claimed. Thus, there is a unique $T \in \wedge^4 V$ such that

\[
T(x_1, x_2, y_1, y_2) = (x_1 \wedge x_2 \wedge y_1 \wedge y_2) \cdot T,
\]
i.e. the first property holds whenever \( u \) and \( v \) are pure. Since the property is linear in \( u \) and \( v \), it holds for all \( u,v \in \wedge^2 V \).

The second property follows by setting \( u = v \).

To deduce the third property from the first, we take \( v \in \wedge^2 V \) arbitrary, and we compute

\[
v \cdot (u \wedge F \wedge F - \lambda_2 u) = ((u \wedge F) \wedge v) \cdot F - \lambda_2 v \cdot u = (v \wedge (u \wedge F)) \cdot F - \lambda_2 v \cdot u = (u \wedge F) \cdot (v \wedge F) - \lambda_2 v \cdot u = (u \wedge v) \cdot T = v \cdot (u \wedge T)
\]

\( \square \)

**Proposition 2.6.3.** Let \( F \) be 1-fabulous. Then \( F \) is 2-fabulous if and only if

\[
\forall \text{ pure } x \in \wedge^2 V : q(x \wedge F) = \tilde{\lambda}_2 q(x), \text{ equivalently, }
\forall x, y \in \wedge^2 V : (x \wedge F) \cdot (y \wedge F) = \tilde{\lambda}_2 q(x, y) + (x \wedge y) \cdot T,
\]

where we have defined \( \tilde{\lambda}_2 = \lambda_2 - 2\lambda_1 + \lambda_0 \).

**Proof.** The proof is a bit a nasty computation which comes down to comparing \( a_x^\dagger a_x' \) and \( a_{x'} a_x^\dagger \) where \( x = x_1 \wedge x_2 \) and \( x' = x_1' \wedge x_2' \) are two pure bivectors. For brevity, we will denote \( a_{x_i} = a_i^\dagger \) and \( a_{x_i'} = a_i' \) and \( q(i,i') = q(x_i, x_i') \). We compute

\[
a_{x'} a_x = a_i^\dagger a_{x_i} a_{x_i'} a_1^\dagger
= q(2,2') a_i^\dagger a_{x_i} a_{x_i'} - a_i^\dagger a_{x_i} a_{x_i'} a_1^\dagger
= q(2,2') a_i^\dagger a_{x_i} a_{x_i'} - q(2,1') a_i^\dagger a_{x_i} a_{x_i'} a_2 a_1^\dagger + a_i^\dagger a_{x_i} a_{x_i'} a_2 a_1^\dagger
= q(2,2') a_i^\dagger a_{x_i} a_{x_i'} - q(2,1') a_i^\dagger a_{x_i} a_{x_i'} + q(1,2') a_i^\dagger a_{x_i} a_{x_i'} - a_i^\dagger a_{x_i} a_{x_i'} a_1^\dagger
= q(2,2') a_i^\dagger a_{x_i} a_{x_i'} - q(2,1') a_i^\dagger a_{x_i} a_{x_i'} a_2 a_1^\dagger + q(1,2') a_i^\dagger a_{x_i} a_{x_i'} - q(1,1') a_i^\dagger a_{x_i} a_{x_i'} a_2 a_1^\dagger + a_i^\dagger a_{x_i} a_{x_i'} a_2 a_1^\dagger
\]

We now compute \( \langle F | a_{x'} a_x^\dagger | F \rangle = (x' \wedge F) \cdot (x \wedge F) \). Recall that, since \( F \) is 1-fabulous:

\[
\langle F | a_i^\dagger a_i' | F \rangle = \lambda_1 q(i,i')
\]

\[
\langle F | a_i' a_i^\dagger | F \rangle = (\lambda_0 - \lambda_1) q(i,i'),
\]
where $\lambda_0 = q(F)$. So we get

$$\langle F | a_x^\dagger a_x^\dagger | F \rangle = \lambda_1 (q(2, 2')q(1, 1') - q(2, 1')q(1, 2')) + (\lambda_0 - \lambda_1)(q(1, 2')q(1', 2) - q(1, 1')q(2, 2')) + \langle F | a_x^\dagger a_x^\dagger | F \rangle$$

$$= (2\lambda_1 - \lambda_0)\langle x' | x \rangle + \langle F | a_x^\dagger a_x^\dagger | F \rangle$$

From here, the claim easily follows. \qed

**Problem 2.6.4.** Here are a number of problems to work on.

1. There is probably a generalization of Proposition 2.5.3 and Proposition 2.6.3 for $n$-fabulous multivectors which are also $i$-fabulous for $i < n$. This seems easy enough to formulate—e.g. $\tilde{\lambda}_3 = -\lambda_3 + 3\lambda_2 - 3\lambda_1 + \lambda_0$—but not so straightforward to prove elegantly. Perhaps Wicks’s theorem is relevant.

2. The parameters $\tilde{\lambda}_1 = \lambda_0 - \lambda_1$ and $\tilde{\lambda}_2 = \lambda_2 - 2\lambda_1 + \lambda_0$ have a meaning in the theory of block designs. They correspond to the complementary design where the new blocks are the complements of the old blocks. This is an application of the inclusion-exclusion principle; for instance: the number $\tilde{\lambda}_2$ of blocks not containing 2 points $a$ and $b$ is the total number $\lambda_0$ of blocks, minus the number of blocks $\lambda_1$ containing $a$, minus the number of blocks $\lambda_1$ containing $b$, plus the number $\lambda_2$ of blocks containing both $a$ and $b$. We expect that this is related to Hodge duality: if $\omega \in \wedge^p V$ is a multivector of highest degree then $\wedge^i V \to (\wedge^{n-i} V)^*: u \mapsto \star u$, where $\star u(x) = \langle a_x a_u \omega \rangle = (x \wedge u) \cdot \omega$ interchanges the operators $a$ and $a^\dagger$ somehow. So our fabulous multivector gives rise to a dual multivector with complementary parameters.

3. We have not explored many examples but it is remarkable that in the interesting examples that we know—the Fano plane $\text{PG}(2, 2)$ with points and lines and $\text{AG}(3, 2)$ with points and planes—the multivector $T$ from Proposition 2.6.2 also plays the role of the complement $F \perp \omega$. If this were true in general, it would imply that

$$v - k = \delta \omega - \delta F = \delta T = 4$$
which is a very severe restriction and perhaps enough to conclude that there are not many 2-fabulous multivectors.
3.1 Definition and goal

**Definition 3.1.1.** We will call a multivector \( F \in \wedge^k V \) *superfabulous* if it is \( k-1 \)-fabulous.

The goal for the rest of this chapter is to classify superfabulous multivectors with a multiplier \( \lambda_t \neq 0 \), where \( t = k-1 \). Our motivation behind this quest is that this is equivalent to classifying cross product algebras and, in the case \( k = 3 \), also to classifying composition algebras, as we will explain in Section 3.3. In fact, Brown and Gray already completed this classification in [BG67], so the result is not new.

However, Brown and Gray classify these by relying on the classification of (unital) composition algebras, which is traditionally done via the Cayley-Dickson doubling process. We don’t really like this approach because it bypasses the explanatory value of the Fano plane, as we briefly talked about in Problem 2.3.6.

So we will follow a different path; and we will give the essential ideas behind it right away in Proposition 3.2.1 and Proposition 3.2.3 to make a direct connection between Steiner systems and superfabulous multivectors. In the case \( k = 3 \), we will rely on Lemma 3.4.4 which provides a combinatorial substitute for the the Cayley-Dickson doubling process. This process gives a basis as required in Proposition 3.2.1, and therefore shows there is a Steiner system \( S(t, t+1, n) \) which underlies the superfabulous multivector. This Steiner system must then in fact be a projective space \( \text{PG}(n, \mathbb{F}_2) \) and in a final step we exclude the possibility of disjoint lines so that \( n = 2 \). For our
Chapter 3. Superfabulosity

later purposes—the investigation of groups of type $G_2$—we will need a rather precise formulation of the result, i.e. a precise description of octonions in general, and this will be given in Theorem 3.4.8.

From there we can tackle the case of $k > 3$ rather easily by relying on Proposition 3.2.3 to give a more intuitive treatment of the case of superfabulous 4-vectors and show non-existence for superfabulous 5-vectors unless $\dim V = 5$. The block designs that are responsible for the existing superfabulous multivectors are then given by

<table>
<thead>
<tr>
<th>$t$</th>
<th>$k$</th>
<th>$v$</th>
<th>block design</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>$2m$</td>
<td>perfect matching (‘1-factor’)</td>
</tr>
<tr>
<td>$v-2$</td>
<td>$v-1$</td>
<td>$v$</td>
<td>single block containing all points</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>7</td>
<td>Fano plane $\text{PG}(2,2)$</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>8</td>
<td>affine space $\text{AG}(3,2)$</td>
</tr>
</tbody>
</table>

Although we think that our methods provide a new angle on these cross product algebras, we must note that the eventual proof for the case $k = 3$ is actually quite close to proofs published by [Dar09] and [Gor+94].

3.2 First lemmas

Recall Notation 2.3.1 and Notation 2.4.2 which we will use frequently in this chapter.

Furthermore we will identify the orbits of $\mathbb{A}^1(k) = (k, +)$ under the group $\text{GL}_1(k) = (k^\times, \cdot)$ with the set $\mathbb{B} = \{0, 1\}$ and as usual we define $\mathbb{P}V$ to be the set of vector lines, i.e. non-trivial orbits of $V$ under the action of $\text{GL}_1$. Then there is a map

$$\pi : (k, +) \to \{0, 1\} : x \mapsto x \mod \text{GL}_1(k)$$

which sends 0 to itself and every other element to 1 and we will define the bracket $[\ldots]$ by

$$[a_1 | \ldots | a_k] = (a_1 | \ldots | a_k) \mod \text{GL}_1(k).$$

Proposition 3.2.1. Let $S \subseteq \mathbb{P}V$ be a set of anisotropic, mutually orthogonal vector lines, closed under $X$. Define blocks to be the subsets
3.2. First lemmas

\{b_1, \ldots, b_k\} such that \([b_1 | \ldots | b_k] = 1\). Then this is a Steiner system, i.e. a \(t-(v, k, 1)\) block design with \(k = t + 1\).

**Proof.** Note that within the set \(S\), we have\(^1\) that

\[ x \cdot y \equiv 1 \mod \text{GL}_1 \iff x = y. \]

Now take any \(t\) elements \(b_1, \ldots, b_t\) from the set \(S\). Then an element \(u \in V\) will complete this to a block if the following condition holds:

\[ [b_1 | \ldots | b_t | u] = 1 \iff X(b_1, \ldots, b_t) \cdot u \equiv 1 \mod \text{GL}_1 \]

\[ \iff u = X(b_1, \ldots, b_t). \]

Therefore there is a unique such \(u\), and thus a unique block through every \(t\) distinct elements. \(\square\)

**Problem 3.2.2.** Proposition 3.2.1 sketches the main idea behind our approach, but we cannot use it directly because we do not know how to construct such a set \(S\) with good properties \textit{a priori}. So we will use an approach which builds up such a set slowly by adding more vectors. The good news is that this will not only give us a Steiner system, but at the same time allow us to describe its structure.

**Proposition 3.2.3.** If a multivector \(F\) is superbulous, then every contraction \(a_xF\) is also superbulous on the space \(x^\perp\).

**Proof.** For every \(x \in V\) and every \(y \in \wedge^*(x^\perp)\) which is pure of degree \(t - 1\), we may compute

\[ q(y \cdot a_x F) = q((x \wedge y) \cdot F) = \lambda_t q(x \wedge y) = \lambda_t q(x) q(y). \]

\(\square\)

**Lemma 3.2.4** (Exchange condition.). Let \(v = v_1, \ldots, v_{t-1}\) be an ordered set of \(t - 1\) elements of \(V\). Then

\[ X(\overline{v}, a) \cdot b = (\overline{v} | a | b) = -X(\overline{v}, b) \cdot a. \]

\(^1\)This is sometimes called the \textit{defining property} of 24, since the condition “\(xy \equiv 1 \mod n \implies x \equiv y \mod n\)” implies that \(n | 24\).
Chapter 3. Superfabulosity

Proof. If \( v = v_1 \land \ldots \land v_{t-1} \) then we may compute

\[
X(\bar{v}, a) \cdot b = ((v \land a) \bot F) \cdot b \\
= b \bot ((v \land a) \bot F) \\
= (v \land a \land b) \bot F \\
= (\overline{v} \mid a \mid b),
\]

and this expression is anti-symmetric in \( a \) and \( b \).

Remark 3.2.5. The map \( \land^{t-1}V \to so(q) : v \mapsto X(v, \cdot) \) is also the map \( \land^{k-2}V \to \land^2V : x \mapsto x \bot F \) under the identification of Proposition 2.3.3. If we could show that this map is injective, e.g. by showing that is an isometry, then it follows immediately that \( \binom{k-2}{v} \leq \binom{2}{v} \), i.e \( k-2 \leq 2 \) or \( k-2 \geq v-2 \). This would prove immediately that \( k \leq 4 \) or \( k = v \), which is something that we will show but only after a considerable effort.

Proposition 3.2.6. Let \( \bar{v} = v_1, \ldots, v_{t-1} \) be a set of \( t-1 \) elements of \( V \). Then

\[
X(\bar{v}, X(\bar{v}, a)) = -\lambda_t a_{\bar{v}} a_{\bar{v}}^\dagger a
\]

Proof. This is equivalent with

\[
\forall x \in V : x \cdot X(\bar{v}, X(\bar{v}, a)) = -\lambda_t a_{\bar{v}} a_{\bar{v}}^\dagger a \cdot x.
\]

Equality now follows directly by applying the exchange condition and \( t \)-fabulosity to the left hand side.

3.3 Connection with other structures

Let us first recall the definition of a cross product algebra as stated in [BG67].

Definition 3.3.1. A cross product algebra is a vector space \( A \) of some finite dimension \( n \) over a field of characteristic not two, together with an \( r \)-fold product

\[
X : \underbrace{A \times A \times \ldots \times A}_{r \text{ copies}} \to A
\]
where $1 \leq r \leq n$, such that the following two conditions hold. (We copy the numbering from [BG67].)

(1.1) $X(a_1,\ldots,a_r) \cdot a_i = 0$ for all $a_1,\ldots,a_r \in A$.

(1.2) $X(a_1,\ldots,a_r) \cdot X(a_1,\ldots,a_r) = \det(a_i \cdot a_j)$.

We can now prove that the notion of a superfabulous multivector specializes to the notion of a cross product algebra.

**Proposition 3.3.2.** Superfabulous multivectors with $\lambda_t = 1$ correspond to cross product algebras.

**Proof.** For a superfabulous $F$, the product $X : V^t \to V$ immediately satisfies the properties (1.1) and (1.2) from [BG67]. Conversely, let $X : V^t \to V$ be a map satisfying the conditions (1.1) and (1.2). Then we may define a map

$$
\chi : V^{t+1} \to V : (a_1,\ldots,a_{t+1}) \mapsto X(a_1,\ldots,a_t) \cdot a_{t+1}.
$$

By linearization of (1.1) this map is alternating, thus by non-degeneracy of the bilinear form there is an $F \in \wedge^{t+1}V$ such that

$$
X(a_1,\ldots,a_t) \cdot a_{t+1} = (a_1 \wedge \ldots \wedge a_{t+1}) \lrcorner F = a_{t+1} \cdot ((a_1 \wedge \ldots \wedge a_t) \lrcorner F).
$$

Since $a_{t+1}$ is arbitrary, it can be erased in both sides and then the condition (1.2) says precisely that $F$ is $t$-fabulous with $\lambda_t = 1$. \hfill \Box

Our main goal for the rest of this chapter is to prove the following theorem, while staying true to the philosophy sketched in Proposition 3.2.1 that there should be a block design which is responsible for the superfabulous vector.

**Theorem 3.3.3** (Gray–Brown,[BG67]). Cross product algebras exist only if:

$$
v \text{ even } 7 8 \text{ arbitrary} \\
r \text{ 1 2 3 } v - 1
$$

**Proposition 3.3.4.** Superfabulous trivectors correspond to composition algebras.
Chapter 3. Superfabulosity

Proof. Let $V$ be a vector space with quadratic form $q$ and a trivector $F$ which is superfabulous with respect to $q$. Define a quadratic form on $C = 1 \oplus V$ by

$$N(s, u) = s^2 + q(u)$$

and a product given by

$$(s, u)(t, v) = (st - u \cdot v) + (su + tv + u \times v).$$

It is then a simple computation that $N(x)N(y) = N(xy)$ for all $x, y \in C$, relying on the fact that $u \times v$ is orthogonal to $u$ and $v$, and also on the fact that $q(a + b) = q(a) + q(b) + 2a \cdot b$.

Conversely, if $C$ is a composition algebra, one defines $V = 1^\perp$ and $a \times b = \frac{1}{2}(ab - ba)$. Standard properties of composition algebras guarantee that this is a vector cross product and so it defines a 2-fabulous trivector as in Proposition 3.3.2.

3.4 Trivectors

Let us now assume that $F$ is a superfabulous trivector with $\lambda = \lambda_2 \neq 0$. Our goal will be to show that $v = \dim V = 3$ or 7 and in particular that $F$ is related to a projective geometry $\mathbb{P}G(1, 2)$ or $\mathbb{P}G(2, 2)$, seen as a block design. To proceed in the spirit of Proposition 3.2.1, we will in a first step construct a basis of vector lines which is closed under $\times$, and show that if we let the triples $a,b,a \times b$ determine the lines, we retrieve a projective geometry $\mathbb{P}G(n, 2)$ for some $n$. In the last step, we will want to exclude $n \geq 3$.

Lemma 3.4.1. Let $F$ be a superfabulous trivector. Then

$$b \times (a \times b) = \lambda a_b a_b^\dagger a$$

$$b \times (a \times c) + c \times (a \times b) = \lambda (a_b a_c^\dagger + a_c a_b^\dagger)(a).$$

Proof. This is a special case of Proposition 3.2.6 and a linearization.
Remark 3.4.2. Fully written out, Lemma 3.4.1 gives the familiar formulas:

\[ b \times (a \times b) = \lambda ((b \cdot b)a - (a \cdot b)b) \]
\[ b \times (a \times c) + c \times (a \times b) = \lambda (2(b \cdot c)a - (a \cdot c)b - (a \cdot b)c) \]
\[ (c \times a) \times (a \times b) = \lambda (2(a|b|c)a + (a \cdot c)(b \times a) \ldots \]
\[ \ldots - (a \cdot b)(c \times a) + (a \cdot a)(c \times b)) \]

The last equation follows by substituting \( c \rightarrow c \times a \) in the second equation and working this out. Since it is not required for our purposes, we will not bother with it too much, but we do need a special case where most of the occuring vectors are orthogonal. Let us show this in detail.

**Lemma 3.4.3.** If \( c \perp a \perp b \perp (a \times c) \) then

\[ (a \times c) \times (a \times b) = \lambda q(a) b \times c. \]

If \( c \perp a_1 \perp b \perp (a_1 \times c) \) and \( c \perp a_2 \perp b \perp (a_2 \times c) \) and \( a_1 \perp a_2 \),

\[ (a_1 \times c) \times (a_2 \times b) = -(a_2 \times c) \times (a_1 \times b). \]

**Proof.** From Lemma 3.4.1, under the assumption that \( a, b, c \) are orthogonal and anisotropic, we have:

\[ b \times (b \times a) = -\lambda q(b) a \]
\[ b \times (a \times c) = (a \times b) \times c, \]

Therefore, under the assumption that \( c \perp a \perp b \perp (a \times c) \):

\[ (a \times c) \times (a \times b) = (a \times (a \times c)) \times b = \lambda q(a) b \times c. \]

The second equation follows by linearizing this expression in \( a \). \( \square \)

These lemmas make it easy to prove the following analogon to the Cayley-Dickson doubling process.

**Lemma 3.4.4.** Let \( S \subset \mathbb{P}V \) be a set of anisotropic, orthogonal vector lines, closed under \( \times \). Let \( x \in \langle S \rangle ^\perp \) be anisotropic. Then

\[ S' = \{ x \} \cup S \cup \{ x \times s \mid s \in S \} \]

is an anisotropic set of orthogonal vector lines, closed under \( \times \).
Proof. It is immediately clear from 2-fabulosity that for any \( s, t \in S \), we have
\[
(x \times s) \cdot (x \times t) \equiv s \cdot t \mod \text{GL}_1,
\]
and \((x \times s) \cdot t = (s \times t) \cdot x = 0\), therefore all vector lines in \( S' \) are anisotropic and orthogonal. So we must only show that \( S' \) is closed under \( \times \). But this is also immediately clear from the following multiplication table, which holds for arbitrary \( s, t \in S \) up to \( \text{GL}_1 \), and can be computed with the help of Lemma 3.4.1 and Lemma 3.4.3.

<table>
<thead>
<tr>
<th></th>
<th>( x )</th>
<th>( s \times s )</th>
</tr>
</thead>
</table>
| \( x \) | 0 | \( x \times s \)
| \( t \) | \( t \times s \) | \( x \times (s \times t) \)
| \( x \times t \) | | \( t' \times s \)

Lemma 3.4.5. Any superfabulous trivector originates from a projective geometry over \( \mathbb{F}_2 \).

Proof. By inductively applying Lemma 3.4.4 starting with a single vector until the entire vector space is exhausted, we may construct an orthogonal basis which is closed under \( \times \). Define lines in this set to be subsets of the form \( \{a, b, a \times b\} \). Then every 3 distinct elements \( a, b, c \) are either collinear or \( c \perp a \times b \). From there one can immediately check that the Veblen-Young axioms for a projective geometry hold: (i) every line contains 3 points (ii) every two points \( a \) and \( b \) lie on a unique line \( (a, b, a \times b) \) (iii) if distinct lines \( ab \) and \( cd \) intersect, say in \( a \times c \equiv x \equiv d \), then
\[
\begin{align*}
a \times c & \equiv a \times (x \times d) \\
& \equiv (a \times x) \times d \\
& \equiv b \times d
\end{align*}
\]
and so the lines \( ac \) and \( bd \) intersect. (The other cases, such as where \( c = a \times b \) are immediately verified.)

Lemma 3.4.6. The dimension of the projective geometry is at most 2, therefore \( \dim V \in \{3, 7\} \).
Proof. If the dimension is at least 3, then there are at least 2 disjoint lines; this means that we have a set $S = \{a, b, a \times b, u, v, u \times v\}$ of 6 anisotropic, mutually orthogonal vectors. Then by Lemma 3.4.3

$$(a \times u) \times (b \times v) = -(a \times v) \times (b \times u)$$

$$= +(b \times v) \times (a \times u)$$

$$= -(a \times u) \times (b \times v).$$

Where we used that $b \perp a \times u$, $a \perp b \times v$, $u \perp a \times v$ and $v \perp b \times u$, because otherwise the lines would intersect. So we obtain $(a \times u) \times (b \times v) = 0$, which means that lines $(a, u, a \times u)$ and $(b, v, b \times v)$ intersect, but then the points $a, b, u, v$ lie in a plane and the lines $a, b, a \times b$ and $u, v, u \times v$ must intersect after all. \qed

The case where $\dim V = 3$ is of course that of a top-vector as we will see in Proposition 3.5.4, so we focus on $\dim V = 7$.

**Lemma 3.4.7.** Let $\dim V = 7$. Consider $a, b, c$ orthogonal and anisotropic, such that $c \perp a \times b$. Use the notation $a' = \frac{b \times c}{q(b)q(c)}$ and cyclic permutations, and $e = a \times (b \times c)/\lambda$. Then

$$\mathcal{B} = \{a, b, c, a', b', c', e\}$$

is an orthogonal basis and

$$F = a \wedge b \wedge c' + b \wedge c \wedge a' + c \wedge a \wedge b' +$$

$$+ \frac{a' \wedge e \wedge a}{q(a)} + \frac{b' \wedge e \wedge b}{q(b)} + \frac{c' \wedge e \wedge c}{q(c)} +$$

$$+ \frac{q(a)q(b)q(c)}{\lambda} (c' \wedge b' \wedge a').$$

Proof. If we apply Lemma 3.4.4 to the set $S = \{a, b, a \times b\}$ with $c$ in the role of $x$, we get the set $\mathcal{B}$ as depicted in the following diagram. Here $a'$ represents the vector line $b \times c$ (and cyclic permutations) and $e$ the vector line $a \times (b \times c)$. 
This is clearly the Fano plane. Note that it is easy to compute the actual products (not up to $\text{GL}_1$): for instance
\[ a \times a' = \frac{a \times (b \times c)}{q(b)q(c)} = \frac{\lambda}{q(b)q(c)} e. \]
Note that $q(c')q(a)q(b) = \lambda$ and cyclic permutations; and $q(e) = q(a)q(b)q(c)$. Therefore
\[ q(a)q(b)q(c') = \lambda, \]
\[ q(a)q(e)q(a') = \lambda q(a)^2, \]
\[ q(a')q(b')q(c') = \lambda \frac{\lambda^2}{(q(a)q(b)q(c))^2}. \]
(This says that each of the terms occurring in $F$ has norm $\lambda$.) It is now straightforward to verify that the expression for $F$ from the statement of the theorem satisfies $u \times v = (u \wedge v) \cdot F$ for any $u, v \in B$ and therefore defines $F$. For instance
\[ (a \wedge a') \cdot F = \frac{1}{q(a)} (a \wedge a') \cdot (a' \wedge e \wedge a) \]
\[ = q(a') e = a \times a'. \]

**Theorem 3.4.8.** If a quadratic form $q$ admits a superfabulous trivec-
tor with multiplier $\lambda \neq 0$, then with respect to some basis $a, b, c, a', b', c', e$, we have
\[ q = \langle \lambda^2 q_a, \lambda^2 q_b, \lambda^2 q_c, \lambda q_b q_c, \lambda q_c q_a, \lambda q_a q_b, q_a q_b q_c \rangle. \]
in particular $\text{disc}(q) = \lambda \in k^\times/(k^\times)^2$, and
\[ \lambda q_a q_b q_c F = \frac{q_c}{\lambda} a \wedge b \wedge c' + \frac{q_a}{\lambda} b \wedge c \wedge a' + \frac{q_b}{\lambda} c \wedge a \wedge b' + \]
\[ + a' \wedge e \wedge a + b' \wedge e \wedge b + c' \wedge e \wedge c + a' \wedge b' \wedge c'. \]
3.4. Trivectors

Proof. This simpler form is obtained by using the new variables \( a \leftarrow \lambda^2 a, \ a' \leftarrow q_b q_c a' \) in Lemma 3.4.7. (We have changed the notation from \( q(a) \) to \( q_a \) because \( q_a \) is no longer the norm of \( a \) in this form, instead \( q(a) = \lambda^2 q_a \) as can be read off from the description of \( q \).) \( \square \)

Remark 3.4.9. Note that

\[
\langle \lambda^3 \rangle \oplus q \simeq \langle \lambda, q_a \rangle \otimes \langle \lambda, q_b \rangle \otimes \langle \lambda, q_c \rangle,
\]

which means that the quadratic forms \( q \) which admit a superfabulous 3-vector with multiplier \( \lambda \) are precisely the ones such that \( q \oplus \lambda \) is a Pfister 3-form.

Remark 3.4.10. For later use in our later study of the Lie algebra \( g_2 \), we will now record a simple observation. First note that in the notation from above, we have

\[
a_c F = -\lambda(a_u^a a_b a_c + a_v^a a_c a_a + a_w^a a_a a_b) F.
\]

Since every anisotropic \( x \in V \) can be written as \( x = u \times (v \times w) \), this implies \( \cap_{u,v,w \in V} ((a_u^a a_v a_w + a_v^a a_w a_u + a_w^a a_u a_v) F) \perp \subseteq \cap_{x \in V} (a_x F) \perp \).

Finally, here is a strange observation that won’t be used in what follows but looks like it could be useful.

Lemma 3.4.11. For all \( u, v, w \in V \):

\[
(a_u \times (v \times w) + a_v \times (w \times u) + a_w \times (u \times v)) F = -3\lambda(a_u^a a_v a_w + a_v^a a_w a_u + a_w^a a_u a_v) F
\]

In other words, the following bracket :

\[
\{uvw\} = (a_u \times (v \times w) + 3\lambda a_u^a a_v a_w) F \in \wedge^2 V,
\]

satisfies \( \{uvw\} + \{uwv\} = 0 \) and \( \{uvw\} + \{vwu\} + \{wuv\} = 0 \).

Proof. Clearly it is sufficient to verify this for \( u, v, w \) members of an orthogonal basis as in Theorem 3.4.8. Whenever 2 of them are equal, or if \( u, v, w \) are collinear, the claim is trivial since both hands evaluate to 0. So we can without loss of generality assume \( u = e_1, v = e_2 \) and \( w = e_3 \) and then this follows from Remark 3.4.10. \( \square \)
3.5 Other multivectors

As we mentioned a few times, the main goal of this work was to understand groups of type $G_2$, and these are related to the case $(k,v) = (3,7)$. So our treatment of a classification of superfabulous multivectors in the case $k \neq 3$ will be rather sketchy in comparison, but on the other hand it is also more tedious than difficult with what we know at this point.

Let us first rapidly handle the case $k = 2$.

**Proposition 3.5.1.** Superfabulous bivectors exist in even dimensions only; they are all of the vorm

$$F = e_1 \wedge e_2 + e_3 \wedge e_4 + \ldots + e_{2n-1} \wedge e_{2n},$$

with respect to some orthogonal basis for which $\lambda = q(e_1)q(e_2) = \ldots = q(e_{2n-1})q(e_{2n})$. In particular $q(F) = \lambda n$.

**Proof.** Let $F$ be a superfabulous bivector, i.e. a 1-fabulous 2-vector. Take any $e_1 \in V$ anisotropic and note that

$$e_1 \cdot (e_1 \downarrow F) = (e_1 \wedge e_1) \cdot F = 0 \text{ and } q(e_1 \downarrow F) = \lambda \cdot q(e_1).$$

So if we define $f_1 = e_1 \downarrow F$, then $e_1 \perp f_1$ and $\lambda q(e_1) = q(f_1)$. In particular, $F$ is non-degenerate on $\langle e_1, f_1 \rangle$ and we may proceed inductively on the orthogonal complement to obtain a basis

$$e_1, \ldots, e_n, f_1, \ldots, f_n$$

such that $e_i \downarrow F = f_i$ and $f_i \downarrow F = \lambda e_i$ by Proposition 2.5.2. Next, observe that the multivector

$$F_1 = \frac{1}{q(e_1)}(e_1 \wedge f_1)$$

has the property that $e_1 \downarrow F_1 = f_1$ and $f_1 \downarrow F_1 = -\frac{q(f_1)}{q(e_1)} e_1 = -\lambda e_1$. Therefore, if we redefine $e'_1 = e_1/q(e_1)$ then

$$F = e'_1 \wedge f_1 + \ldots + e'_n \wedge f_n,$$
where $\lambda = q(f_1)/q(e_1) = q(f_1)q(e'_1)$. Although the last part is immediately clear, we can also use Corollary 2.4.8 to obtain

$$\lambda \cdot 2n = q(F) \cdot 2.$$ 

We saw in 3.4 that if $k = 3$, then $v = 4$ or $v = 7$. Therefore if $k = 4$, then by Proposition 3.2.3 $v = 5$ or $v = 8$. The former case is not very interesting (see Proposition 3.5.4), so let us focus on the latter case. Although it is possible to provide a more detailed analysis and compute such a multivector in as much detail as we did for trivectors in Theorem 3.4.8, we will only prove a weak version, which says that if it exists, there is also a block design responsible for it.

**Proposition 3.5.2.** Let $\Gamma$ be a 3-fabulous 4-vector. Then there is a block design $\text{AG}(3,2)$, where the planes are the blocks, underlying $\Gamma$ as in Proposition 3.2.1.

**Proof.** Let $\Gamma \in \wedge^4 V$ be 3-fabulous. For any anisotropic $x \in V$, $x \perp \Gamma$ is a 2-fabulous 3-vector on $x^\perp$ by Proposition 3.2.3 and thus either $v = \dim V = 4$ or $\dim V = 8$ and the situation from Theorem 3.4.8 applies to $F = x \perp \Gamma$. We ignore the case $v = 4$, which is dealt with by Proposition 3.5.4 and focus on $v = 8$.

We then obtain a basis

$$B = \{x, a, b, c, a \times b, b \times c, c \times a, (a \times b) \times c\} \subseteq \mathbb{P}V,$$

where the binary operation $\times$ comes from $F$, for instance

$$a \times b = (a \wedge b) \perp F = (x \wedge a \wedge b) \perp \Gamma.$$

The vectors in this set are mutually orthogonal and anisotropic so we are close to the situation from Proposition 3.2.1, if we can show that the set is also closed under the operation $X$. (Recall that closed means: closed in $\mathbb{P}V$, i.e. up to the action of $\text{GL}_1(k)$.) So we show that for every triple $a_1, a_2, a_3 \in B$, their product $X(a_1, a_2, a_3)$ is also in $B$.

Let us first notice that if one of these vector lines $a_1, a_2$ or $a_3$ is equal to $x$, then we may use

$$(x \wedge a_1 \wedge a_2) \perp \Gamma = (a_1 \wedge a_2) \perp F = a_1 \times a_2 \in B.$$
By Proposition 3.2.6, we have that
\[ X(a_1, a_2, x) \equiv a_1 \times a_2 \implies X(a_1, a_2, a_1 \times a_2) \equiv x \pmod{\text{GL}_1}. \]
So the problem is reduced to determining \( X(a_1, a_2, a_3) \) where no 3 of these vector lines are collinear in the Fano plane which underlies \( F \). We can take without loss of generality \( a, b, c \) in the description from Theorem 3.4.8. Let us determine \( u = X(a, b, c) \) First, note that
\[ u \cdot x \equiv X(a, b, x) \cdot c = (a \times b) \cdot c = 0, \]
but also
\[ u \cdot a = u \cdot b = u \cdot c = 0. \]
Furthermore
\[ u \cdot (a \times b) = X(a, b, a \times b) \cdot c = X(a, b, c) \cdot (a \times b) = X(a, b, c) \cdot X(a, b, x) = c \cdot x = 0. \]
So \( u \) is orthogonal to all vectors from \( B \), except for possibly \( e = (a \times b) \times c \). But since \( q(u) = \lambda_3 q(a)q(b)q(c) \), this component must be non-zero and \( \{a, b, c, e\} \) forms a block.

So we found that \( \Gamma \) is actually determined by a Steiner system in the sense of Proposition 3.2.1 and this Steiner system can be thought of as a Fano plane plus a point \( x \), with 7 blocks given by the lines of the Fano plane together with \( x \) and 7 blocks given by the complements of lines in the Fano plane. This is the affine geometry \( AG(3, 2) \) with planes as blocks.

It seems this multivector has been studied in the literature from two different directions. On the one hand, it is studied by differential geometers as the *Cayley 4-form*, see [KS10] and the references therein, in particular [HL82] where this form was first defined. The other direction from which this has been studied is the more algebraic study of ternary composition algebras by McCrimmon [McC83], who has investigated trilinear maps \( : V \times V \times V \to V \) such that \( q(\{x, y, z\}) = q(x)q(y)q(z) \) for every \( x, y, z \in V \); see also [Sha00, 16.B] for an overview and additional references.

But with 4 variables, the story ends, as we will now show.
Proposition 3.5.3. There is no superfabulous 5-vector unless $v = 5$.

Proof. Let $Q$ be a superfabulous 5-vector with $v \neq 5$. By Proposition 3.5.2 and Proposition 3.3.2 we know that $v = 9$. We apply the same procedure as in the proof of Proposition 3.5.2: consider an anisotropic vector line $x$, and consider the contraction $\Gamma = x \cdot Q$. This gives a set $S$ of 8 vector lines in the orthogonal complement, with the structure of an AG(3, 2) block design, and we have a set $\{x\} \cup S$ of 9 anisotropic mutually orthogonal vector lines.

Every 3 vector lines $u$, $v$, $w$ in $S$ uniquely determine a fourth one $t$ which is coplanar in AG(3, 2), so we have that

$$X(x, u, v, w) = t.$$  

Next, take any 4 points $a$, $b$, $c$, $d$ in $S$ which are not coplanar. Then each of the four remaining points lies in a plane with three out of these four points $a$, $b$, $c$ and $d$. For instance, say $u, a, b, c$ form a plane. But then

$$X(a, b, c, d) \cdot u \equiv X(a, b, c, u) \cdot d \equiv x \cdot d \equiv 0.$$  

So we see that $X(a, b, c, d)$ must be orthogonal to all points in $S$. But on the other hand, if

$$X(a, b, c, d) \equiv x \implies X(a, b, c, x) \equiv d,$$  

and then $a, b, c$ and $d$ must be coplanar after all. So we conclude that $X(a, b, c, d) = 0$ and this is a contradiction because its norm is $\lambda_4 q(a)q(b)q(c)q(d) \neq 0$. 

Finally, a top-vector in a $v$-dimensional space is a non-zero vector in the 1-dimensional space $\wedge^v V$. These are always superfabulous.

Proposition 3.5.4. For any quadratic form, any top-vector is superfabulous.

Proof. Choose an orthogonal basis $e_1, \ldots, e_v$ such that $\omega = e_1 \wedge \ldots \wedge e_v$ is the top-vector. Then a basis for $\wedge^{v-1} V$ is given by the vectors

$$f_i = q(e_i)e_1 \wedge \ldots \wedge \hat{e_i} \wedge \ldots \wedge e_v,$$
and thus \( f_i \cdot \omega = q(\omega) e_i \) and \( f_i \cdot f_j = q(\omega)(e_i \cdot e_j) \). Thus

\[
(f_i \cdot \omega) \cdot (f_j \cdot \omega) = q(\omega)^2 (e_i \cdot e_j) = q(\omega) f_i \cdot f_j.
\]

This bilinear condition holds for every pair of vectors from the basis \((f_i)\), so it holds for arbitrary vectors and thus for any \( v \in \wedge^{v-1}V \), we have

\[
q(v \cdot \omega) = q(\omega)q(v).
\]

### 3.6 Multiplication tables for octonions

We have worked out some of the multiplication tables. Starting from

\[
\begin{align*}
e_1 \times e_2 &= q_1 q_2 e'_3 	ext{ and } e_1 \times e'_1 = \frac{\lambda}{q_2 q_3} e_0 \\
e_1 \times e'_2 &= \frac{\lambda}{q_3} e_3 \text{ and } e'_1 \times e'_2 = -\frac{\lambda}{q_3} e'_3,
\end{align*}
\]

where \( q_0 = q_1 q_2 q_3 \), we obtain:

<table>
<thead>
<tr>
<th>( \times )</th>
<th>( e'_3 )</th>
<th>( e'_2 )</th>
<th>( e'_1 )</th>
<th>( e_0 )</th>
<th>( e_1 )</th>
<th>( e_2 )</th>
<th>( e_3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( e'_3 )</td>
<td>0</td>
<td>( \lambda/q_1 e'_1 )</td>
<td>( -\lambda/q_2 e'_2 )</td>
<td>( \lambda e_3 )</td>
<td>( \lambda/q_2 e_2 )</td>
<td>( -\lambda/q_1 e_1 )</td>
<td>( -\lambda/q_2 q_3 e_0 )</td>
</tr>
<tr>
<td>( e'_2 )</td>
<td>( -\lambda/q_1 e'_1 )</td>
<td>0</td>
<td>( \lambda/q_3 e'_3 )</td>
<td>( \lambda e_2 )</td>
<td>( -\lambda/q_3 e_3 )</td>
<td>( -\lambda/q_1 q_3 e_0 )</td>
<td>( \lambda/q_1 e_1 )</td>
</tr>
<tr>
<td>( e'_1 )</td>
<td>( \lambda/q_2 e'_2 )</td>
<td>( -\lambda/q_3 e'_3 )</td>
<td>0</td>
<td>( \lambda e_1 )</td>
<td>( -\lambda/q_2 q_3 e_0 )</td>
<td>( \lambda/q_3 e_3 )</td>
<td>( -\lambda/q_2 e_2 )</td>
</tr>
<tr>
<td>( e_0 )</td>
<td>( -\lambda e_3 )</td>
<td>( -\lambda e_2 )</td>
<td>( -\lambda e_1 )</td>
<td>0</td>
<td>( q_0 e'_1 )</td>
<td>( q_0 e'_2 )</td>
<td>( q_0 e'_3 )</td>
</tr>
<tr>
<td>( e_1 )</td>
<td>( -\lambda/q_2 e_2 )</td>
<td>( \lambda/q_3 e_3 )</td>
<td>( \lambda/q_2 q_3 e_0 )</td>
<td>( -q_0 e'_1 )</td>
<td>0</td>
<td>( q_1 q_2 e'_3 )</td>
<td>( -q_1 q_3 e'_2 )</td>
</tr>
<tr>
<td>( e_2 )</td>
<td>( \lambda/q_1 e_1 )</td>
<td>( \lambda/q_1 q_3 e_0 )</td>
<td>( -\lambda/q_3 e_3 )</td>
<td>( -q_0 e'_2 )</td>
<td>( -q_1 q_3 e'_4 )</td>
<td>0</td>
<td>( q_2 q_3 e'_1 )</td>
</tr>
<tr>
<td>( e_3 )</td>
<td>( \lambda/q_1 q_2 e_0 )</td>
<td>( -\lambda/q_1 e_1 )</td>
<td>( \lambda/q_2 e_2 )</td>
<td>( -q_0 e'_3 )</td>
<td>( q_1 q_3 e'_2 )</td>
<td>( -q_2 q_3 e'_1 )</td>
<td>0</td>
</tr>
</tbody>
</table>

Since the quadratic form determines the cross product uniquely, there is a unique version of the octonions that corresponds to hyperbolic quadratic forms. These are the split octonions. In this case, we may assume that \( q_1 = q_2 = q_3 = -1 \). This leads to the multiplication table


Another typical choice of basis is to introduce \(e_i^\pm = e_i \pm e_i'\). Note that \(q(e_i^\pm) = 0\), while \(q(e_i^+, e_i^-) = 2\).
4 Interlude on orthogonal groups

“The idea is elementary (once Chevalley thought of it).”

Jim Humphreys

4.1 Introduction

The purpose of this interlude is to provide a set of generators for the special orthogonal groups $\text{SO}(V, q) = \text{SO}_q$, starting from the Lie-algebra $\mathfrak{so}(V, q)$. Typically one creates generators for a split group by Chevalley’s construction. This method relies on occurrence of many nilpotent elements but in a non-split group there are not enough nilpotents to make this work. Since our construction aims to deal with orthogonal groups in general and not only with the split forms, we propose a way to generate the group with non-split tori instead. The goal that we had in mind was to provide explicit generators of anisotropic $G_2$ and from there generators for anisotropic mixed groups of type $G_2$, which is something that is not present in the literature. We never looked into whether this works or not. To be more precise, we will obtain a candidate set of generators, by adapting the methods in this chapter to the Lie algebra $\mathfrak{g}_2$, but we never looked into whether the result is something that deserves to be called a mixed anisotropic group of type $G_2$. (See Section 5.3.) Even worse, we will not even prove that the generators that we find in this chapter indeed generate the full orthogonal group, so we leave it as Problem 4.3.3 instead.

1https://mathoverflow.net/q/156963
Let us first give an informal account of the idea behind the construction by reasoning in characteristic 0. We then know that a nilpotent element $X$ of the Lie algebra $\mathfrak{so}$ gives rise to an element of $SO$ by exponentiating and observing that the series $\exp(tX) = 1 + tX + \ldots$ terminates. If the group is non-split, there are not enough nilpotent elements to perform this construction, but there are still many elements $X$ such that $X^2 = a$ is a scalar matrix. (They essentially come from the pure bivectors in the identification of $\mathfrak{so}(V, q)$ with $\wedge^2 V$; and more precisely the equality $X^2 = a$ holds only on a subspace, while $X = 0$ on the orthogonal complement.) In this case, one obtains a series which looks like
\[
\exp(tX) = (1 + \frac{t^2a}{2!} + \frac{t^4a^2}{4!} + \ldots) + (1 + \frac{t^3a}{3!} + \ldots)X.
\]
For $a = -1$, we recognize the (circular) sine and cosine series, for $a = 1$ we recognize the hyperbolic sine and cosine series. Of course sine and cosine are transcendental functions which are not defined over an arbitrary field but the only important fact about them here is that they form a parametrization of the curve $x^2 - ay^2 = 1$ which can be given the group-structure of a $k$-torus. So we can just take a rational parametrization and use this to define
\[
\exp(tX) = r\cos_a(t) + r\sin_a(t)X.
\]
This will provide an explicit embedding
\[
T(k) \to SO_q(k) : t \mapsto \exp(tX)
\]
of the group of $k$-rational points of a non-split torus into the orthogonal group.

### 4.2 Anisotropic tori

Let us first study anisotropic one-dimensional tori associated to a fixed element $a \in k^\times$. Consider the set
\[
T(k) = k \cup \{\infty\} \setminus \{t \in k \mid t^2 = a\}.
\]
This is a group under the operation\(^2\)
\[
u \cdot v = \frac{au + av}{a + uv},
\]
with the expected rules for \(\infty\) that \(\infty \cdot v = \frac{a}{v}\) and \(\infty \cdot \infty = 0\). Note
that 0 is the identity element and \(u \cdot (-u) = 0\). Note the distinction
between the case where \(a\) is a square or not.

First, if \(\sqrt{a} \in k\), the map
\[
f : \mathbb{P}^1(k) \to \mathbb{P}^1(k) : z \mapsto \frac{\sqrt{a} - z}{\sqrt{a} + z},
\]
satisfies \(f(x) \cdot f(y) = f(x \cdot y)\), which shows that \((T(k), \cdot) \cong (k^\times, \cdot)\)
with \(\pm \sqrt{a}\) playing the (absorbing) roles of 0 and \(\infty\). In particular,
this proves that \((T(k), \cdot)\) is a group.

If \(a\) is not a square one can embed \(k\) in \(k(\sqrt{a})\) and still use this map
\(f\) to show that \(\cdot\) defines a group.

Whenever \(t \in T(k)\) we also define the rational \(a\)-sine and \(a\)-cosine as functions \(T(k) \to k\) with:
\[
rcos_a(t) = \frac{a + t^2}{a - t^2}, \quad rsin_a(t) = \frac{2t}{a - t^2},
\]
with of course \(rsin_a(\infty) = 0\) and \(rcos_a(\infty) = -1\). These provide a
rational parametrization of the curve \(x^2 - ay^2 = 1\):
\[
\{(x, y) \in k^2 \mid x^2 - ay^2 = 1\} = \{(rcos_a(t), rsin_a(t)) \mid t \in T(k)\}.
\]

The group \((T(k), \cdot)\) can thus be identified with the kernel of the norm
\(k(\sqrt{a})^\times \to k^\times : (u + \sqrt{av}) \mapsto u^2 - av^2\). In particular the group is (up
to isomorphism) only dependent on the class of \(a\) in \(k^\times/(k^\times)^2\).

If \(X\) is an endomorphism of a vector space \(V\) over \(k\) of dimension \(\geq 3\),
we will, for \(t \in k\), define the rational exponential
\[
rexp_a(tX) = rcos_a(t) + rsin_a(t)X.
\]

\(^2\)The reader may recognize the addition formula for \(\tan(x + y)\).
Proposition 4.2.1. If $X^2 = a$ then the map $t \mapsto \exp_a(tX)$ is a morphism $(T(k), \bullet) \to \text{GL}_V(k)$ of groups.

Proof. It suffices to observe that

$$\exp_a(sX) \circ \exp_a(tX) = \exp_a((s \bullet t)X).$$

This comes down to verifying the addition formulas

$$r\cos_a(s) r\cos_a(t) + a r\sin_a(s) r\sin_a(t) = r\cos_a(s \bullet t)$$
$$r\cos_a(s) r\sin_a(t) + r\cos_a(t) r\sin_a(s) = r\sin_a(s \bullet t),$$

which is straightforward.

Now if $X^2 = c^2a$ for some scalar $c$, it makes sense to define

$$\exp_a(\zeta X) = \exp_a(\zeta c \cdot \frac{X}{c}) = r\cos_a(\zeta c) + \frac{r\sin_a(\zeta c)}{c} X,$$

and when $X^2 = 0$, we define $\exp_a(\zeta X)$ by the familiar expression $1 + \zeta X$, which is also compatible with the heuristic $\lim_{c \to 0} r\sin_a(\zeta c)/c = \zeta$.

Similarly if $X = X_1 \oplus \ldots \oplus X_n$ on $V = V_1 \perp \ldots \perp V_n$, where $X_i^2 = c_i^2 a$, then we may simply define

$$\exp_a(tX) = \exp_a(tX_1) \oplus \ldots \oplus \exp_a(tX_n).$$

In the particular case that we will need, $X = X_1 \oplus X_2$ on $V = V_1 \perp V_2$, where $X_1^2 = a$ and $X_2 = 0$, so we have $\exp_a(tX) = \exp_a(tX_1) \oplus \text{id}_{V_2}$. Furthermore, we will also abbreviate $r\cos_a = \cos$ and $r\sin_a = \sin$ if we think it is clear from the context what $a$ is.

4.3 Generating orthogonal groups

Now as before, let $(V, q)$ be a quadratic space of dimension $\geq 3$ and consider the Lie algebra $\mathfrak{so}(V, q) \cong \wedge^2(V)$ with its adjoint action. If $(e_i)_{i=1}^n$ is an orthogonal basis for $V$, then the elements $(e_i \wedge e_j)_{1 \leq i < j \leq n}$ form an orthogonal basis for $\mathfrak{so}(V, q)$—it is useful to imagine this basis
as the set of edges of a complete graph on $n$ vertices. The action of $e_1 \wedge e_2$ on $\wedge^2 V$ has the following property:

\[
\begin{cases}
\text{ad}(e_1 \wedge e_2)^2 = -q(e_1)q(e_2) \text{id} & \text{on } (\wedge^2 \langle e_1, e_2 \rangle)^\perp \oplus \wedge^2 \langle e_1, e_2 \rangle \perp \\
\text{ad}(e_1 \wedge e_2) = 0 & \text{on } \wedge^2 \langle e_1, e_2 \rangle \perp \oplus \wedge^2 \langle e_1, e_2 \rangle.
\end{cases}
\]

(Seen as edges of a complete graph, we have partitioned the edges $E = \{u, v\}$ based on the parity of $E \cap \{e_1, e_2\}$.)

This means we can use our rational exponential to define $\text{rexp}_a \text{ad}(e_1 \wedge e_2)!$. We can do this for every pure multivector $x \wedge y$ and so we obtain a group which we denote

\[
\text{SO}_2(V, q) = \langle e^{\zeta(x \wedge y)} \mid \zeta \in T(k), x, y \in V, x \perp y, q(x)q(y) \neq 0 \rangle \\
\leq \text{GL}(\wedge^2 V),
\]

where we have used the abbreviation

\[
e^{\zeta(x \wedge y)} = \text{rexp}_{-q(x)q(y)}(\zeta \cdot \text{ad}(x \wedge y)).
\]

(In fact, if $u = x \wedge y$ where $x$ and $y$ are not orthogonal, one can write $y = \mu_1 x + y'$ with $y' \perp x$ and observe that $u = x \wedge y'$ and it does not change the space $\langle x, y \rangle$.)

The standard representation $\rho : \mathfrak{so}(V, q) \rightarrow \mathfrak{gl}(V)$, behaves very similarly, since

\[
\begin{cases}
\rho(e_1 \wedge e_2)^2 = -q(e_1)q(e_2) \text{id} & \text{on } \langle e_1, e_2 \rangle \\
\rho(e_1 \wedge e_2) = 0 & \text{on } \langle e_1, e_2 \rangle^\perp,
\end{cases}
\]

so we will also consider the group

\[
\text{SO}_1(V, q) = \langle e^{\zeta(x \wedge y)} \mid \zeta \in T(k), x, y \in V, x \perp y, q(x)q(y) \neq 0 \rangle \\
\leq \text{GL}(V),
\]

where the same abbreviation was used. This could lead to confusion, since the space on which $x \wedge y$ acts is certainly relevant in the description, but no real confusion can arise because of the following proposition.
Proposition 4.3.1. Whenever \( u \in \wedge^2 V \) is pure and anisotropic, the equation \( e^{\zeta u}(x \wedge y) = e^{\zeta u}x \wedge e^{\zeta u}y \) holds for all \( x, y \in V \). In other words, the following diagram commutes.

\[
\begin{array}{ccc}
SO_1 & \longrightarrow & SO_2 \\
\downarrow & & \downarrow \\
GL(V) & \longrightarrow & GL(\wedge^2 V).
\end{array}
\]

Proof. We can always write \( u \) as \( u_1 \wedge \ldots \wedge u_i \) with respect to some orthogonal basis \( u_1, \ldots, u_v \), and we will denote \( U = \langle u_1, \ldots, u_i \rangle \). First let \( x, y \in U^\perp \); then

\[
e^{\zeta u}x \wedge e^{\zeta u}y = x \wedge y = e^{\zeta u}(x \wedge y).
\]

Next, let \( x \in U^\perp \) and \( y \in U \); then we may assume that \( u = z \wedge y \) for some \( z \in U \), \( z \perp y \) and we have

\[
e^{\zeta u}x \wedge e^{\zeta u}y = x \wedge (\cos \zeta y + \sin \zeta (y \perp u))
= \cos \zeta (x \wedge y) + \sin \zeta q(y)(z \wedge x),
\]

while

\[
e^{\zeta u}(x \wedge y) = \cos \zeta (x \wedge y) + \sin \zeta [x \wedge y, u]
= \cos \zeta (x \wedge y) + \sin \zeta q(y)(z \wedge x).
\]

Finally, let \( x, y \in U \). Then \( u = r(x \wedge y) \) for some \( r \in k \), and we have \( x \perp u = rq(x)y \) and \( y \perp u = -rq(y)x \). Furthermore recall that we work with respect to \( a = -q(u) = -r^2q(x)q(y) \). This implies

\[
(e^{\zeta u}x) \wedge (e^{\zeta u}y) = (\cos \zeta x + \sin \zeta (x \perp u)) \wedge (\cos \zeta y + \sin \zeta (y \perp u))
= (\cos \zeta x + \sin \zeta rq(x)y) \wedge (\cos \zeta y - \sin \zeta rq(y)x)
= (\cos \zeta)^2 (x \wedge y) - q(u)(\sin \zeta)^2 y \wedge x
= ((\cos \zeta)^2 - a(\sin \zeta)^2)(x \wedge y)
= x \wedge y = e^{\zeta u}(x \wedge y).
\]

Of course, we want our generators to be contained in the orthogonal group!
Proposition 4.3.2. For \( u \in \wedge^2 V \) pure and anisotropic,
\[
(e^\zeta u_x) \cdot (e^\zeta u_y) = x \cdot y,
\]
therefore \( \text{SO}_1(V,q) \leq \text{SO}(V,q) \).

Proof. Let \( u = e_1 \wedge e_2 \). First, let \( x, y \in \langle e_1, e_2 \rangle^\perp \), then
\[
e^\zeta u_x = x \quad \text{and} \quad e^\zeta u_y = y,
\]
and the statement is true. Next, let \( x \in \langle e_1, e_2 \rangle^\perp \) and \( y \in \langle e_1, e_2 \rangle \). Then
\[
e^\zeta u_x = x \quad \text{and} \quad e^\zeta u_y = \cos \zeta y + \sin \zeta (y \cdot u),
\]
so that, using the fact that \( x \cdot (y \cdot u) = -(x \cdot u) \cdot y = 0 \), and recalling that \( x \cdot y = 0 \),
\[
e^{\zeta u} x \cdot e^{\zeta u} y = x \cdot \left( \cos \zeta \cdot y + \sin \zeta \cdot (y \cdot u) \right)
= (\cos \zeta)(x \cdot y)
= 0 = x \cdot y.
\]
Finally in the most general case, we have that
\[
e^{\zeta u} x \cdot e^{\zeta u} y = \left( \cos \zeta x + \sin \zeta (x \cdot u) \right) \cdot \left( \cos \zeta y + \sin \zeta (y \cdot u) \right)
= (\cos \zeta)^2 (x \cdot y) + (\sin \zeta)^2 (x \cdot u) \cdot (y \cdot u)
= (\cos \zeta)^2 x \cdot y - (\sin \zeta)^2 x \cdot (y \cdot u \cdot u)
= ((\cos \zeta)^2 - a(\sin \zeta)^2) (x \cdot y)
= x \cdot y \ \blacksquare
\]

Problem 4.3.3. We expect that these generators indeed generate the entire orthogonal group:
\[
\text{SO}_1(V,q) = \text{SO}(V,q).
\]

For a proof we suggest showing that every product of 2 reflections is in \( \text{SO}_1 \) and using the theorem of Cartan-Dieudonné to conclude that this generates the entire special orthogonal group.
In this chapter, we continue our discussion from Section 3.4 on superfabulous trivectors. We should warn the reader that this is a rather sketchy account of what we know about $G_2$-structures, with some computations omitted. Our reasons for this are mainly that the results are usually well known in one form or another, or otherwise can be proven by a direct computation; so we have prioritized other chapters in this thesis and restricted ourselves to the essentials here.

Before we begin, recall that we have defined:

1. A field $k$ and a $7$-dimensional vector space $V$ over $k$ and corresponding Grassman algebra $\wedge^\bullet V$. (Section 2.1)

2. A non-degenerate quadratic form $V \rightarrow k$ with a choice of orthogonal basis given by $(e_1, \ldots, e_7)$. (Section 2.2) Recall also that the associated bilinear form is denoted $\beta(x, y) = x \cdot y$ and that we have introduced the contraction products $x \frown y = a_{xy}$. (Notation 2.3.1)

3. An isomorphism $\wedge^2 V \cong so(V)$, and for each $u \in \wedge^2 V$ the corresponding derivation $D_u(x) = x \frown u$ such that $D_u x \cdot y + x \cdot D_u y = 0$. (Proposition 2.3.3)

4. A 2-fabulous trivector (Section 2.4) $F \in \wedge^3 V$ with multiplier $\lambda_2 = 1$ and corresponding maps $V \times V \rightarrow V : (u, v) \mapsto u \times v = (u \wedge v) \frown F$ and $V \times V \times V \rightarrow k : (u, v, w) \mapsto (u \wedge v \wedge w) \frown F = \langle u, v, w \rangle$ which is completely determined by $q$. (Theorem 3.4.8)
5.1 The Lie algebra $\mathfrak{g}_2$

The linear map $\pi$, defined by
\[ \pi : \wedge^2 V \to V : x \mapsto x \downarrow F, \]
restricts to the cross product on the pure bivectors. We will observe that the kernel $\ker \pi \leq \mathfrak{so}(V)$ is in fact a 14-dimensional Lie-subalgebra, by showing that it consists of all simultaneous derivations of the inner product and the cross product. This will be our definition of $\mathfrak{g}_2$.

**Proposition 5.1.1.** The following sets are equal:

(\(\mathfrak{g}_2\)) The kernel of the map $\pi : \mathfrak{so}(V) \to V : x \mapsto x \downarrow F$.

(\(\tilde{\mathfrak{g}}_2\)) The set of all $u \in \wedge^2 V$ such that $D_u x \times y + x \times D_u y = D_u (x \times y)$.

Furthermore, we define

(\(\check{\mathfrak{g}}_2\)) The set of all $\alpha \in \mathfrak{gl}(V)$ such that $\langle \alpha(x), y, z \rangle + \langle x, \alpha(y), z \rangle + \langle x, y, \alpha(z) \rangle = 0$.

(t) The set $t = \{(x \downarrow F) \downarrow F - 3x \mid x \in \wedge^2 V\}$.

Then all sets are Lie algebras; we have an inclusion of ideals $t \leq \mathfrak{g}_2 \leq \check{\mathfrak{g}}_2$ and the dimensions are given by:

<table>
<thead>
<tr>
<th>char $k$</th>
<th>dim $t$</th>
<th>dim $\mathfrak{g}_2$</th>
<th>dim $\check{\mathfrak{g}}_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k \neq 3$</td>
<td>14</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>$k = 3$</td>
<td>7</td>
<td>14</td>
<td>15</td>
</tr>
</tbody>
</table>

**Proof.** Clearly all sets are linear spaces. Let us first show that $\check{\mathfrak{g}}_2$ and $\tilde{\mathfrak{g}}_2$ are Lie-algebras by a direct computation. Take $u, v \in \check{\mathfrak{g}}_2$ and $x, y \in V$, then we have
\[
D_u D_v (x \times y) = D_u D_v x \times y + D_v x \times D_u y + D_u x \times D_v y + x \times D_u D_v y.
\]
If we subtract from this the same expression with the roles of $u$ and $v$ reversed, we find
\[
[D_u, D_v](x \times y) = [D_u, D_v] x \times y + x \times [D_u, D_v] y,
\]
5.1. The Lie algebra \( \mathfrak{g}_2 \)

thus \( \tilde{\mathfrak{g}}_2 \) is closed under the commutator bracket, as desired. Similarly, if we take \( \alpha, \beta \in \tilde{\mathfrak{g}}_2 \) then we have

\[
\langle \alpha \beta x, y, z \rangle = \langle x, \beta \alpha y, z \rangle + \langle x, \beta y, \alpha z \rangle
\]

Subtracting from this the same expression with the roles of \( \alpha \) and \( \beta \) reversed, we find

\[
\langle [\alpha, \beta] x, y, z \rangle = -\langle x, [\alpha, \beta] y, z \rangle - \langle x, y, [\alpha, \beta] z \rangle,
\]

which shows that \( \tilde{\mathfrak{g}}_2 \) is closed under the commutation bracket.

We will now show that \( \tilde{\mathfrak{g}}_2 = \mathfrak{g}_2 \). We have

\[
u \in \mathfrak{g}_2 \iff a_u F = 0 \iff \forall y \in V : u \perp a_y F \iff u \in \cap_{y \in V} a_y F
\]
on the one hand and

\[
u \in \tilde{\mathfrak{g}}_2 \iff ((x \cup u) \wedge y) \cup F + (x \wedge (y \cup u)) \cup F
\]

on the other hand. Therefore equality is just Remark 3.4.10. Clearly \( \pi \) is surjective, since every anisotropic vector is a cross product of two vectors and every vector is a sum of anisotropic vectors, and thus \( \dim \mathfrak{g}_2 = \dim \mathfrak{so}_7 - \dim V = \binom{7}{2} - 7 = 14 \).

The statements concerning \( \mathfrak{t} \) and \( \tilde{\mathfrak{g}}_2 \) we leave to the reader.\(^1\)

We will now study the structure of this Lie algebra.

\(^1\)We don’t know a coordinate-free proof although certainly for the fact that \( [\mathfrak{t}, \mathfrak{g}_2] \subseteq \mathfrak{t} \) we believe there should be one. Note that our proof of the equality \( \mathfrak{g}_2 = \tilde{\mathfrak{g}}_2 \) is also not really synthetic because Remark 3.4.10 relies on a normal form for \( F \).
Lemma 5.1.2. Consider a basis as in Theorem 3.4.8, denoted by $e_0, e_1, e_2, e_3, e'_1, e'_2, e'_3$. Define

$$h = h_0 = \langle e_1 \wedge e'_1, e_2 \wedge e'_2, e_3 \wedge e'_3 \rangle \cap (\ker \pi).$$

Then $h \leq \ker \pi$ is a 2-dimensional abelian Lie-algebra. If we define similarly $h_1, h'_1, \ldots$, then $[h_i, h_j] = h_k$ whenever $(i, j, k)$ is a line on the Fano plane. In particular $h_0 \oplus \bigoplus_{i=0}^3 (h_i \oplus h'_i) = g_2$.

Proof. Clearly, by interchanging the roles of the basis vectors, it is sufficient to show that $[h_0, h_i] = h_{-i}$ and this can be read off from the Fano plane. This implies that $h_0 \oplus \bigoplus_{i=0}^3 (h_i \oplus h'_i)$ is a Lie-subalgebra of $g_2 = \ker \pi$ of dimension $2 \cdot 7 = 14$; and therefore both are equal. \qed

Remark 5.1.3. This determines what Thompson [Tho76] calls a Dempwolff decomposition for $g_2$.

It is possible to refine this decomposition as follows. To assist the reader with the computations, we first recall that

$$e_0 = q_2q_3(e_1 \times e'_1) = q_1q_3(e_2 \times e'_2) = q_1q_2(e_3 \times e'_3)$$
$$e_1 = q_1(e_2 \times e'_3) = e'_1 \times e_0 = q_1(e'_2 \times e_3)$$
$$e'_1 = q_1(e'_3 \times e'_2) = \frac{1}{q_0}(e_0 \times e_1) = q_1'(e_2 \times e_3).$$

So if we define

$$h_1 = q_2q_3(e_1 \wedge e'_1)$$
$$h_2 = q_3q_1(e_2 \wedge e'_2)$$
$$h_3 = q_1q_2(e_3 \wedge e'_3)$$

then these are not elements of $g_2$, but the differences $h_i - h_j$ are and in fact $h = \langle h_1 - h_2, h_2 - h_3 \rangle$. Next, we define

$$\eta_1 = q_1e_2 \wedge e'_3 - 2e'_1 \wedge e_0 + q_1e'_2 \wedge e_3$$
$$\theta_1 = q_1e_2 \wedge e'_3 - q_1e'_2 \wedge e_3$$
$$\eta'_1 = q_1e'_3 \wedge e'_2 - \frac{2}{q_0}e_0 \wedge e_1 + q_1'e_2 \wedge e_3$$
$$\theta'_1 = q_1e'_3 \wedge e'_2 - q'_1e_2 \wedge e_3.$$
Note that $\eta_1, \theta_1 \in \mathfrak{h}_1$, and $\eta_1', \theta_1' \in \mathfrak{h}_1'$. It is then straightforward to compute the following expressions:

\[
q(h_1 - h_2) = q(h_2 - h_3) = q(h_3 - h_1) = 2q_0,
\]

\[
q(\eta_i) = 6q_i \quad q(\theta_i) = 2q_i.
\]

So we see that the spaces $\Theta_i = \langle \theta_i, \theta_i' \rangle$ and $H_i = \langle \eta_i, \eta_i' \rangle$ provide a decomposition of $g_2$ into eigenspaces for $\mathfrak{h}$. Let us record this:

**Proposition 5.1.4.** There is a decomposition

\[
g_2 = \mathfrak{h} \oplus \bigoplus_{i=1}^{3} (\Theta_i \oplus H_i)
\]

where $[\mathfrak{h}, \Theta_i] \leq \Theta_i$, $[\mathfrak{h}, H_i] \leq H_i$.

If the Lie algebra $g_2$ is non-split, one cannot do better. However, if the field $k$ contains the right scalar, which turns out to be $\sqrt{-q_0}$, it is possible to diagonalise the action of $\mathfrak{h}$ on the $\Theta_i$ and $H_i$ spaces, and obtain a decomposition in root spaces, see Section 5.4.

(There are many other field extensions of degree 2 that will split the Lie algebra, but then one must make different choices along the way to obtain the corresponding decomposition; in other words one must be careful to choose a *splitting* Cartan algebra.)

### 5.2 Small characteristic phenomena

Although investigating $G_2$-phenomena that occur in characteristic 3 was our primary motivation, we have not said much about it yet. Very much in the spirit of the rest of Part I, we will disappoint the reader and only give an informal account of what is special about characteristic 3.
First, we must recall that the classical representation theory of semi-simple complex Lie algebras provides us with the following decompositions of \( g_2 \) modules (for instance, see [FH91, §22.3]):

\[
\wedge^2 V \cong V \oplus \Gamma_{0,1} \quad \text{and} \quad \wedge^3 V \cong \mathbb{C} \oplus V \oplus \Gamma_{2,0}.
\]

Here the symbols \( \Gamma_{a,b} \) denote certain irreducible representations with highest weight \( a\omega_1 + b\omega_2 \), where \( \omega_1 \) and \( \omega_2 \) denote the fundamental weights of \( g_2 \). Of interest to our story is mainly the fact that \( \Gamma_{0,1} \) is called the \textit{adjoint representation}; it is just the Lie algebra \( g_2 \) itself. (In fact, \( V \) can also be written as \( \Gamma_{1,0} \) and is then called the \textit{standard} representation.)

To understand what happens to these decompositions over fields of small characteristic (but different from 2), we will need to reconstruct these decompositions by means of the (superfabulous) 3-vector \( F \) and the corresponding 4-vector \( T \) from Proposition 2.6.2. Recall from Proposition 2.5.4 that our approach with fabulous multivectors will give away a fundamental dichotomy in some characteristics, ultimately related to combinatorial properties of the underlying design (the Fano plane). The dichotomy comes from the maps

\[
\begin{align*}
    k & \to \wedge^3 V \to k : x \mapsto (x \downarrow F) \downarrow F = 7 \cdot x \\
    V & \to \wedge^3 V \to V : x \mapsto (x \downarrow T) \downarrow T = 4 \cdot x \\
    V & \to \wedge^2 V \to V : x \mapsto (x \downarrow F) \downarrow F = 3 \cdot x,
\end{align*}
\]

where the numbers \( \lambda_0 = 7 \), \( \lambda_0 - \lambda_1 = 4 \) and \( \lambda_1 = 3 \) find their combinatorial origin in the number of lines, lines not through a point, and lines through a point in the Fano plane. These maps govern the splitting behaviour of the representations \( \wedge^3 V \) and \( \wedge^2 V \). In particular, whenever \( \text{char}(k) \neq 3 \), we recover the splitting \( \wedge^2 V \cong V \oplus g_2 \); and whenever \( \text{char}(k) \neq 7 \), we obtain \( \wedge^3 V \cong k \oplus V \oplus U \) for some \( U \), where the isomorphism is one as \( g_2 \)-modules.

\textbf{Problem 5.2.1.} One result in the literature which distinguishes characteristic 7 states that the Witt algebra \( \text{Der}(k[x]/x^7) \) as well as all its Galois twists embed in \( g_2 \) only when the characteristic is 7; although we are not sure if this has anything to do with a failure of this representation to split. This result can be found in [CE16] (and
the references therein) along with many other relevant results for the cases of characteristic 2 and 3. A related issue is the inclusion of a irreducible simple subgroup of type $G_2$ in a simple group $E_6$ over an algebraically closed field, which fails in characteristic 7, because then this group $G_2$ ‘accidentally’ ends up in a subgroup $F_4 < E_6$; see [LS04].

In characteristic 3, something remarkable happens: although $\wedge^2 V$ still admits $V$ and $g_2$ as subrepresentations, they are now contained on one another via the inclusion

$$V \xrightarrow{x \mapsto x \cdot F} t \xrightarrow{\text{inc}} g_2,$$

as we saw in Proposition 5.1.1. It is ultimately this fact which gives rise to the Ree groups $^2G_2$ and the mixed groups $MG_2$.

For instance, Wilson [Wil10] constructs groups of type $^2G_2$ from first principles using these modules. (Although the main purpose of his work is to circumvent Lie theory, so he sees all spaces as modules for the group $H = AGL(1,8)$.) The connection between his work and Lie theory is summarized in the following diagram:

```
\begin{tikzpicture}
  \node (g2/t) at (0,0) {$g_2/t$};
  \node (g2) at (0,-1) {$g_2$};
  \node (V) at (-1,-2) {$V$};
  \node (m) at (2,-2) {$\wedge^2 V/t$};
  \node (lF) at (1,-3) {$\wedge^2 V$};
  \node (rF) at (3,-3) {$V$};

  \draw[->] (g2) -- (g2/t);
  \draw[->] (g2) -- (V);
  \draw[->] (g2) -- (m);
  \draw[->] (g2) -- (lF);
  \draw[->] (g2) -- (rF);
  \draw[->] (V) -- (lF);
  \draw[->] (lF) -- (rF);
  \draw[->] (m) -- (lF);
  \draw[->] (m) -- (rF);
\end{tikzpicture}
```

The diagram depicts the $g_2$-module $\wedge^2 V$ together with its submodules

$$0 \subset t \subset g_2 \subset \wedge^2 V,$$

and associated quotient modules. Our superfabulous trivector $F$ has kindly provided identifications $\wedge^2 V / g_2 \cong V \cong t$, as we saw in Proposition 5.1.1. Wilson calls this diagram the octopus algebra, because there are eight arrows, each of which he interprets as a left or
right semi-(co-)multiplication (!). (The dashed arrow \( m \) is the mouth of the octopus.)

With this information, we have enough to study \textit{mixed groups} of type \( M\mathbf{G}_2 \). (Wilson is not interested in these groups which become interesting only in presence of non-perfect fields; in fact he is only interested in groups of type \( ^2\!\!\!\mathbf{G}_2 \) over finite fields.) The deeper reason is that these groups ultimately only depend on the existence of an ideal in the Lie algebra \( \mathfrak{g}_2 \). Following Wilson, we can also provide an ad-hoc definition of the Ree groups \( ^2\!\!\!\mathbf{G}_2 \) by tweaking the map \( m \) a little bit with an automorphism of the field and defining the twisted group as the elements \( g \in \text{GL}(V) \) which commute with all arrows.

Nonetheless, there is something important missing on the diagram, which is an identification \( \mathfrak{g}_2/t \cong t \) or \( \wedge^2 V/t \cong \mathfrak{g}_2 \), and we don’t really see a way to construct it without choosing a basis:

\textbf{Problem 5.2.2.} Is there a canonical identification \( \mathfrak{g}_2/t \cong t \)? (Especially in the anisotropic case.)

This would be highly relevant, because it means that we have a composition

\[ \mathfrak{g}_2 \to \mathfrak{g}_2/t \to \mathfrak{g}_2, \]

and it is the ‘exponential’ of this map\(^2\) which is ultimately responsible for the important \( k\)-endomorphism of the algebraic group \( \mathbf{G}_2 \). This is related to the following question, to which I also could not find any relevant information in the literature.

\textbf{Problem 5.2.3.} Let \( G \) be a linear algebraic \( k\)-group and \( n \leq \mathfrak{g} = \text{Lie}(G) \) a Lie \( p\)-subalgebra which exponentiates to a closed subgroup scheme \( N \leq G \) of height \( \leq 1 \). Can one construct \( \text{Lie}(G/N) \) directly from the embedding of Lie \( p\)-algebras \( n \hookrightarrow \mathfrak{g} \)?

\(^2\)More precisely: the quotient with the ad-invariant Lie \( p\)-subalgebra [Bor91, §17] or equivalently the quotient with the corresponding closed normal subgroup scheme contained in the kernel of the relative Frobenius [CGP15, A.7.14].
5.3 Generators for anisotropic $G_2$

We can now apply the procedure that we introduced in Chapter 4 to write down generators for $G_2(q, k)$, where $q$ is the quadratic form under consideration. In fact, it is straightforward to exponentiate the Dempwolff decomposition from Lemma 5.1.2 in the exact same manner but this is not really what we want since that mixes up the long and short root spaces. What we want is to exponentiate in such a way that we can let the short roots take values in a different field than the long roots and hopefully the result is then something that deserves to be called an anisotropic mixed group of type $G_2$. Alas, this is again something that we will not show, since it is merely a sidequest in our hunt for the mixed octonion algebra.

Let us briefly explain how we can exponentiate the decomposition in Proposition 5.1.4. First, we write down the action of the $\eta_i$, $\eta'_i$, $\theta_i$ and $\theta'_i$ with respect to the basis $(e'_3, e'_2, e'_1, e_0, e_1, e_2, e_3)$.

We observe that—for some order on the basis—these matrices are of the form

$$M = \begin{pmatrix} X & 0 & 0 \\ 0 & Y & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

where $X^2 = a$ and $Y^2 = r^2a$ for some scalars $r$ and $a$. This implies that we can exponentiate them with the formula

$$\exp_a \zeta M = \begin{pmatrix} r\cos_a \zeta + X r\sin_a \zeta & 0 & 0 \\ 0 & r\cos_a(\zeta r) + \frac{1}{r} r\sin_a(\zeta r) & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

where as in Chapter 4, $t \mapsto (r\cos t, r\sin t)$ is some rational parametrisation of $(r\cos t)^2 - a(r\sin t)^2 = 1$.

To make this more concrete, we have worked out some of these matrices on a computer. (This seemed less error-prone than performing the computations by hand.) With respect to the aforementioned ordered basis we have—with $\sin = r\sin_{-q_1}$ and $\cos = r\cos_{-q_1}$:
\[ \text{rexp}_{-q_1}(\zeta \eta_1) = \begin{bmatrix} \cos \zeta & 0 & 0 & 0 & 0 & q_1 q_2 \sin \zeta & 0 \\ 0 & \cos \zeta & 0 & 0 & 0 & 0 & -q_1 q_3 \sin \zeta \\ 0 & 0 & \cos 2\zeta & q_0 \sin 2\zeta & 0 & 0 & 0 \\ 0 & 0 & -q_1' \sin 2\zeta & \cos 2\zeta & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ -1/q_2 \sin \zeta & 0 & 0 & 0 & \cos \zeta & 0 \\ 0 & 1/q_3 \sin \zeta & 0 & 0 & 0 & \cos \zeta & 0 \end{bmatrix} \]

\[ \text{rexp}_{-q_1}(\zeta \theta_1) = \begin{bmatrix} \cos \zeta & 0 & 0 & 0 & 0 & q_1 q_2 \sin \zeta & 0 \\ 0 & \cos \zeta & 0 & 0 & 0 & 0 & -q_1 q_3 \sin \zeta \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ -1/q_2 \sin \zeta & 0 & 0 & 0 & \cos \zeta & 0 \\ 0 & 1/q_3 \sin \zeta & 0 & 0 & 0 & \cos \zeta & 0 \end{bmatrix} \]

Similarly, with \( \sin = r \sin_{-q_1'} \) and \( \cos = r \cos_{-q_1'} \), we have

\[ \text{rexp}_{-q_1'}(\zeta \eta_1') = \begin{bmatrix} \cos \zeta & -1/q_3 \sin \zeta & 0 & 0 & 0 & 0 & 0 \\ 1/q_2 \sin \zeta & \cos \zeta & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \cos 2\zeta & q_1 \sin 2\zeta & 0 & 0 \\ 0 & 0 & 0 & -2 \sin 2\zeta & \cos 2\zeta & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \cos \zeta & -1/q_2 \sin \zeta \\ 0 & 0 & 0 & 0 & 1/q_3 \sin \zeta & 0 & \cos \zeta \end{bmatrix} \]

\[ \text{rexp}_{-q_1'}(\zeta \theta_1') = \begin{bmatrix} \cos \zeta & -1/q_3 \sin \zeta & 0 & 0 & 0 & 0 & 0 \\ 1/q_2 \sin \zeta & \cos \zeta & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cos \zeta & -1/q_2 \sin \zeta \\ 0 & 0 & 0 & 0 & 1/q_3 \sin \zeta & 0 & \cos \zeta \end{bmatrix} \]

**Problem 5.3.1.** Do these matrices indeed generate the full group of \( k \)-rational points of a (non-split) \( G_2 \)? If this is the case, consider a field extension \( \ell/k \) such that \( \ell^3 \subseteq k \), note that \( G_2 \) is still anisotropic when extended to \( \ell \). Let \( \zeta \) take values in \( \ell \) on the long roots and in \( k \) on the short roots, do we obtain something that deserves the name *mixed anisotropic* \( G_2 \)?
5.4 Split versions

Assume that $j = \sqrt{-q_0} \in k$. In this case the Lie algebra $\mathfrak{g}_2$ and corresponding group $G_2(q)$ are split. To see this, we define

$$\eta^\pm_i = \eta_i \pm j\eta'_i$$

and

$$\theta^\pm_i = \theta_i \pm j\theta'_i.$$

This determines a decomposition of $\mathfrak{g}_2$ into eigenspaces for the action of $\mathfrak{h}$. The elements $\theta_i^\pm$ correspond to long roots, while $\eta_i^\pm$ correspond to short roots. These roots are indeed nilpotent matrices and their exponentials are given by

$$\exp(\zeta r) = 1 + \zeta r + \frac{1}{2}(\zeta r)^2.$$

From there it is straightforward to explicitly compute these matrices. We will do so under the additional assumption that $q_1 = q_2 = q_3 = -1$ (and $\lambda = 1$, of course); since the underlying quadratic form is hyperbolic these assumptions do not harm the generality.
Chapter 5. Structures of type $G_2$

$$\exp(\zeta \eta^+_1) = \begin{bmatrix} 1 & \zeta & 0 & 0 & 0 & \zeta & 0 \\ -\zeta & 1 & 0 & 0 & 0 & 0 & -\zeta \\ 0 & 0 & 2\zeta^2 + 1 & -2\zeta & -2\zeta^2 & 0 & 0 \\ 0 & 0 & -2\zeta & 1 & 2\zeta & 0 & 0 \\ 0 & 0 & 2\zeta^2 & -2\zeta & -2\zeta^2 + 1 & 0 & 0 \\ \zeta & 0 & 0 & 0 & 0 & 1 & \zeta \\ 0 & -\zeta & 0 & 0 & 0 & -\zeta & 1 \end{bmatrix}$$

$$\exp(\zeta \theta^-_1) = \begin{bmatrix} 1 & -\zeta & 0 & 0 & 0 & \zeta & 0 \\ \zeta & 1 & 0 & 0 & 0 & 0 & -\zeta \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -\zeta & 0 \\ 0 & -\zeta & 0 & 0 & 0 & \zeta & 1 \end{bmatrix}$$

$$\exp(\zeta \eta^-_1) = \begin{bmatrix} 1 & -\zeta & 0 & 0 & 0 & \zeta & 0 \\ \zeta & 1 & 0 & 0 & 0 & 0 & -\zeta \\ 0 & 0 & 2\zeta^2 + 1 & -2\zeta & 2\zeta^2 & 0 & 0 \\ 0 & 0 & -2\zeta & 1 & -2\zeta & 0 & 0 \\ 0 & 0 & -2\zeta^2 & 2\zeta & -2\zeta^2 + 1 & 0 & 0 \\ \zeta & 0 & 0 & 0 & 0 & 1 & -\zeta \\ 0 & -\zeta & 0 & 0 & 0 & \zeta & 1 \end{bmatrix}$$

$$\exp(\zeta \theta^+_2) = \begin{bmatrix} 1 & 0 & -\zeta & 0 & -\zeta & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \zeta & 0 & 1 & 0 & 0 & 0 & \zeta \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ -\zeta & 0 & 0 & 0 & 1 & 0 & -\zeta \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & \zeta & 0 & \zeta & 0 & 1 \end{bmatrix}$$

$$\exp(\zeta \eta^+_2) = \begin{bmatrix} 1 & 0 & -\zeta & 0 & -\zeta & 0 & 0 \\ 0 & 2\zeta^2 + 1 & 0 & -2\zeta & 0 & -2\zeta^2 & 0 \\ \zeta & 0 & 1 & 0 & 0 & 0 & \zeta \\ 0 & -2\zeta & 0 & 1 & 0 & 2\zeta & 0 \\ -\zeta & 0 & 0 & 0 & 1 & 0 & -\zeta \\ 0 & 2\zeta^2 & 0 & -2\zeta & 0 & -2\zeta^2 + 1 & 0 \\ 0 & 0 & \zeta & 0 & \zeta & 0 & 1 \end{bmatrix}$$
5.4. Split versions

\[
\exp(\zeta \theta_\pm) = \begin{bmatrix}
1 & 0 & \zeta & 0 & -\zeta & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 \\
-\zeta & 0 & 1 & 0 & 0 & 0 & \zeta \\
0 & 0 & 0 & 1 & 0 & 0 & 0 \\
-\zeta & 0 & 0 & 0 & 1 & 0 & \zeta \\
0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & \zeta & 0 & -\zeta & 0 & 1 \\
\end{bmatrix}
\]

\[
\exp(\zeta \eta_\pm) = \begin{bmatrix}
1 & 0 & \zeta & 0 & -\zeta & 0 & 0 \\
0 & 2\zeta^2 + 1 & 0 & -2\zeta & 0 & 2\zeta^2 & 0 \\
-\zeta & 0 & 1 & 0 & 0 & 0 & \zeta \\
0 & -2\zeta & 0 & 1 & 0 & -2\zeta & 0 \\
-\zeta & 0 & 0 & 0 & 1 & 0 & \zeta \\
0 & -2\zeta^2 & 0 & 2\zeta & 0 & -2\zeta^2 + 1 & 0 \\
0 & 0 & \zeta & 0 & -\zeta & 0 & 1 \\
\end{bmatrix}
\]
\[
\exp(\eta_{G_2}) = \begin{bmatrix}
2\zeta^2 + 1 & 0 & 0 & -2\zeta & 0 & 0 & 2\zeta^2 \\
0 & 1 & -\zeta & 0 & \zeta & 0 & 0 \\
0 & \zeta & 1 & 0 & 0 & -\zeta & 0 \\
-2\zeta & 0 & 0 & 1 & 0 & 0 & -2\zeta \\
0 & \zeta & 0 & 0 & 1 & -\zeta & 0 \\
0 & 0 & -\zeta & 0 & \zeta & 1 & 0 \\
-2\zeta^2 & 0 & 0 & 2\zeta & 0 & 0 & -2\zeta^2 + 1
\end{bmatrix}
\]

5.5 The orbits of the mixed group

We are now finally ready to study the mixed group \(G_2(k, \ell)\), where \(k\) and \(\ell\) are fields of characteristic 3 such that \(k^3 \leq \ell \leq k\)!

To do so, we first identify the \(k\)-vector space \(V\) of imaginary octonions with \(\mathfrak{psl}_3(k) = \mathfrak{sl}_3(k) / \mathbb{Z}\). Any \(x \in V\) can be represented by a matrix \(A\) in \(\mathfrak{sl}_3(k)\) such that \(x = [A] \in \mathfrak{psl}_3(k)\), where \(\mathfrak{sl}_3(k) \rightarrow \mathfrak{psl}_3(k) : A \mapsto [A]\), the identification is given by

\[
\begin{bmatrix}
e_0 & e_1^+ & e_3^- \\
e_1^- & 0 & e_2^+ \\
e_3^+ & e_2^- & 0
\end{bmatrix},
\]

where \(e_i^\pm = e_i \pm e_i'\). For instance, \(e_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\) and \(e_1' = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\).

The advantage is that the action of the subgroup generated by the short root groups corresponds to the action of \(\text{GL}_3(k)\) (or more precisely \(\text{PGL}_3(k)\)) by conjugation, while the subgroup of long roots can be identified with the group \(\text{SL}_3(k)\), where a matrix \(M\) acts on \((a\ b\ c)^t\) by multiplication, on \((A\ B\ C)^t\) by multiplication with \((M^{-1})^t\) and trivially on \(x\).

So understanding the orbits under the short roots comes down to understanding matrices ‘up to conjugation’ for which we will use an

\[\text{We will omit a computational verification of this and other facts below. The computations were carried out on a computer in a SAGE-worksheet which minimizes the possibility of a computational error. (Although there are still plenty of opportunities for other errors to arise.)}\]
ad hoc version of what is known as the Frobenius\textsuperscript{4} rational form of a matrix.

On the space $\mathfrak{gl}_3(k)$ there exist linear, quadratic and cubic forms, which are invariant under conjugation by an invertible matrix: they are the coefficients of the characteristic polynomial

$$\det(\lambda + A) = \lambda^3 + \text{tr}(A) \lambda^2 + Q(A) \lambda + \det(A).$$

of the matrix. It is easy to verify that $Q(A) = q(x)^5$ and $\text{tr}(A) = 0$, whenever $[A] = x \in V$.

First, we will show that the orbits of $G_2(k^3, k)$ with $q(x) \neq 0$ correspond to the surfaces of constant quadratic invariant $q$. In other words, they are the same as the orbits of the group $\text{SO}(V, q)$ which contains $G_2$. To see this, we will show that every element $x \in V$ with $q(x) \neq 0$ can be mapped to the element of the form

$$\begin{bmatrix} 0 & q(x) & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = Q(x)e_1^+ + e_1^-.$$

So take an arbitrary $x \in V$ and represent it by $A \in \mathfrak{sl}_3(k)$. The characteristic polynomial of $A$ is given by

$$f(x) = x^3 + Q(A)x - \det(A) = 0.$$

Note that if $w$ is a root of this equation, then the other roots (in an algebraic closure) are given by $w \pm \delta$, where $\delta = \sqrt{-Q(A)}$ and by assumption $Q(A) \neq 0$ and thus $f$ is separable. We distinguish 3 cases:

- $f$ has its three roots $w, w \pm \delta$ in $k$ and then

$$[A] \simeq \begin{bmatrix} w + \delta & 0 & 0 \\ 0 & w - \delta & 0 \\ 0 & 0 & w \end{bmatrix} = \begin{bmatrix} \delta & 0 & 0 \\ 0 & -\delta & 0 \\ 0 & 0 & 0 \end{bmatrix} \simeq \begin{bmatrix} 0 & Q(A) & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

(Where $\simeq$ means ‘similar’, i.e. equal up to conjugation. As octonions it means: equal up to action by subgroup generated by the short root groups.)

\textsuperscript{4}Him again!

\textsuperscript{5}Maybe up to a sign.
• $f$ has a unique root $w \in k$ and $\delta \notin k$. Then consider $B = A - w$, note that $[B] = [A - w]$. Since

$$
\det(B) = \det(A - w) \\
= \det(A) - Q(A)w - w^3 \\
= -f(w) = 0,
$$

and $Q(B) = Q(A)$, the characteristic polynomial of $B$ is given by

$$f_B(x) = x^3 - Q(A)x,$$

which implies

$$[A] = [B] \simeq \begin{bmatrix} 0 & Q(A) & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

• $f$ is irreducible over $k$ then $A$ acts on the field $k[x]/f(x)$ as multiplication by $x$, which shows that

$$[A] \simeq \begin{bmatrix} 0 & 0 & \det(A) \\ 1 & 0 & -Q(A) \\ 0 & 1 & 0 \end{bmatrix}.$$

A computation shows that the element $\exp(\frac{\zeta^3}{2} \theta_1^+)$ with $\zeta = \det(A)/Q(A)$, which belongs to the long root groups—restricted to $k^3$ of course!—will map this to the matrix given by

$$[B] = \begin{bmatrix} 0 & 0 & \det(A) + \zeta^3 \\ 1 & 0 & -Q(A) \\ 0 & 1 & 0 \end{bmatrix}.$$

Its characteristic polynomial is of course given by

$$f_B(x) = x^3 + Q(A)x - \det(A) - \zeta^3 = 0.$$

Since $f_B(\zeta) = 0$ so we see that $x = [A]$ can be mapped to an element $[B]$ with reducible characteristic polynomial which reduces us to one of the other two cases.
Next, we consider the quadric $Q = \{ x \in V \mid q(x) = 0 \}$. Let $x = [A] \in Q$ then 

$$\det(\lambda + A) = \lambda^3 + \det(A).$$

This provides us with a well defined map 

$$\det : Q \rightarrow (k,+)/(\ell,+).$$

Explicitly if 

$$v = \begin{bmatrix} x & a & B \\ A & -x & c \\ b & C & 0 \end{bmatrix},$$

then 

$$\det(v) = abc + ABC + xbB - xcC \in k/\ell.$$ 

Now consider the subset 

$$\mathcal{H} = \{ v \in Q \mid \det(v) \in \ell \text{ and } \det(x_r(1) \cdot v) \in \ell, \text{ whenever } r \text{ is long} \}$$

where $x_r$ runs through the long roots. In coördinates, $\mathcal{H}$ is defined by the following 8 expressions: an ordinary equation $q = 0$ and 7 mixing equations:

\begin{align*}
q(v) &= x^2 + aA + bB + cC = 0 \\
A^2C - b^2c + xbA &\in \ell \\
A^2B - bc^2 - xcA &\in \ell \\
AB^2 - ac^2 + xcB &\in \ell \\
B^2C - a^2c - xaB &\in \ell \\
BC^2 - a^2b + xAC &\in \ell \\
AC^2 - ab^2 - xbC &\in \ell \\
\det(v) &= abc + xbB - xcC + ABC \in \ell
\end{align*}

(Note that $-x^3 = x(aA + bB + cC) \in \ell$, so the condition $\det(v) \in \ell$ is invariant under cyclic permutation of $(a,b,c)$ and $(A,B,C)$. )

Computations then verify that this set is indeed stable under the action of the group $G_2(k,\ell)$. Note that it is reasonably large, since it
contains, for instance all elements of the form

\[
\begin{bmatrix}
  x & \delta a & \frac{B}{\delta} \\
  \frac{\delta}{A} & -x & \delta c \\
  \delta b & \frac{C}{\delta} & 0
\end{bmatrix},
\]

where \(a, b, c, A, B, C \in \ell\) are arbitrary, \(\delta \in k\) is arbitrary, and \(x\) satisfies \(0 = x^2 + aA + bB + cC\) (and hence \(x \in \ell\)).

**Problem 5.5.1.** Does \(G_2(k, \ell)\) act transitively on \(Q \setminus \mathcal{H}\)?

This begs the question—what is this thing \(\mathcal{H}\)!? The answer is: it is the mixed hexagon!

### 5.6 The mixed hexagon

This mixed hexagon is an incidence geometric structure that pops up in the classification of the Moufang polygons [TW02]—see also Appendix A for an overview and some historical context. The mixed hexagon \(\mathcal{H}\) is for the mixed group \(G_2(k, \ell)\) what a projective space is for a special linear group; and what a quadric is for an orthogonal group associated to a quadratic form: It is in the first place a geometric object, on which the group acts transitively. It can be constructed by endowing the set of cosets of a particular subgroup (a maximal parabolic) with additional structure. This structure can be that of a projective variety or that of an incidence geometry, depending on what one is trying to achieve. But whereas projective spaces over fields and quadrics can be successfully studied with both the methods of algebraic geometry and incidence geometry, there is no algebro-geometric approach for the mixed hexagon.

This seems clear if we just look at the equations that we found: they do not describe an algebraic variety in any natural way; furthermore the associated group is a mixed group, which is not a group of rational points of an algebraic group. Nonetheless, these groups are clearly very closely related to groups of type \(G_2\) which are algebraic. This raises the question: *what is going on here?* and this is what will bring us to Part II.
5.6. The mixed hexagon

But before we get there, we will use the formulas from [VM98, Table 3.3, right column] to verify with some computer algebra that our equations actually describe the mixed hexagon. (These computations can also easily be done by hand, but this has a higher chance of introducing mistakes.) Let us start up a new Sage worksheet and enter the following command:

```python
tempdomain.<t,u,v,w,x,y,z,a,b,c,A,B,C,D,q> = \ PolynomialRing(GF(3))
R = FractionField(tempdomain)
R.inject_variables()
```

We can now freely use these letters as indeterminates in a field of characteristic 3. Next, we enter the following commands:

```python
def invariants(X):
    a = X[0,1]; b = X[2,0]; c = X[1,2]
    A = X[1,0]; B = X[0,2]; C = X[2,1]
    x = X[0,0]
    return (-b^2*c + x*b*A + A^2*C,
            -b*c^2 - x*c*A + A^2*B,
            -a*c^2 + x*c*B + A*B^2,
            -a^2*c - x*a*B + B^2*C,
            -a^2*b + x*a*C + B*C^2,
            -a*b^2 - x*b*C + A*C^2,
            a*b*c + x*b*B - x*c*C + A*B*C)
def from_hvm_co(t):
    return Matrix(R,3,3,[-t[3],t[0],-t[5],
                        -t[4],t[3],t[2],t[1],-t[6],0])
```

This defines two functions: the first function will compute the values of these 7 polynomials for a given $3 \times 3$ matrix, which is assumed to be traceless with $X_{2,2} = 0$; the second function will convert a point from the coordinates in [VM98] to our preferred representation as a traceless $3 \times 3$ matrix. Finally, we enter the following code:

```python
P1 = (1,0,0,0,0,0,0)
P2 = (a,0,0,0,0,0,1)
P3 = (b,0,0,0,0,1,-u)
P4 = (-u-a*A,1,0,-a,0,a^2,-A)
P5 = (v+b*B,u,1,b,0,B,b^2-B*u)
```

---

6Go to https://sagecell.sagemath.org/ for an easy to use web interface.
Note that the letters $a, b, c$ and $A, B$ do not have the same meaning they had before, but instead they correspond to the letters $k, k', k''$, ... in [VM98]. We have then computed the values of these polynomials for each of the points from the coordinatization provided by [VM98]. For instance, the invariants associated to the last point are given by the following expressions:

$$(v, -u, va^3 - u^2, ua^3c^3 + v^2a^3 + ub^3 + u^2v, \quad u^2c^3 + vb^3 - uv^2, -uc^3 - v^2, a^3c^3 + uv)$$

We can then easily inspect the output to observe that the variables $a, A, b, B$ and $c$ will always occur with a 3 in the exponent; this means that indeed the polynomial takes values in $\ell$.

On a sidenote, the same equations are also implicit in Tits’s description of the Ree groups $^2G_2$ for the Séminaire Bourbaki in [Tit61]: on page 75 he has variables $p_{3i}$, $p_{3i}'$, and $p_{3i}''$ which (he claims) are polynomials in other variables $p_{ij}'$ and $p_{00}' - p_{11}'$, but he does not work out what polynomials they are. It is not so easy to do so! We will leave it as an interesting exercise to the reader to work this out (we found this by staring intensively at the Fano plane) but the outcome is that the polynomials are precisely the ones that describe the mixed hexagon.

We like to think that these equations have not been described before, but the truth is that probably no-one really cared and it seems extremely likely that Tits knew of such a description.
Part II

Twisting and Mixing
“...he has no ambition and no energy. He will not
even go out of his way to verify his own solutions,
and would rather be considered wrong than take the
trouble to prove himself right. Again and again I
have taken a problem to him, and have received an
explanation which has afterwards proved to be the
correct one. And yet he was absolutely incapable of
working out the practical points...”

Sherlock Holmes (about his brother Mycroft)\textsuperscript{1}

So in October 2015 I was still not a single step closer to the mythical mixed octonion algebra. A comment, which I had hear Tom De Medts make a few times was how Richard Weiss often says that we need a version of algebraic geometry that works over two fields at once. I had noticed, some time ago, that if there are two fields $k, \ell$ such that $\ell^p \subseteq k \subseteq \ell$, this information could be expressed more elegantly by saying there is a pair of fields $k, \ell$ together with maps

\[
\begin{array}{c}
k \\
\downarrow \lambda
\end{array} \rightleftharpoons
\begin{array}{c}
\bullet \\
\downarrow \kappa
\end{array} \leftleftharpoons
\begin{array}{c}
\ell
\end{array}
\]

such that $\lambda \circ \kappa$ is the $p$-th power operation on $k$ and $\kappa \circ \lambda$ is the $p$-th power operation on $\ell$. This $p$-th power operation is called the Frobenius and it is drawn on this diagram as some sort of black hole around which the arrows revolve. This was a nice symmetric formulation but I didn’t really know what to do with it at that point.

\textsuperscript{1}Arthur C. Doyle, The Adventure of the Greek Interpreter
But perhaps if we want to define a mixed algebra, we can try something like this. Let us start from a pair $(V_k, V_{\ell})$ of vector spaces together with maps which compose to... Now we are in trouble. The problem with that approach is that there is no $p$-th power operation on a vector space so that wasn’t going anywhere. Unless we decide to skip a step and immediately define a mixed algebra as a pair $(A_k, A_{\ell})$ consisting of a $k$-algebra $A_k$, an $\ell$-algebra $A_{\ell}$, together with maps which compose to the $p$-th power operation. So we get something which looks like this

![Diagram](attachment:diagram.png)

It is important to note that since there is no black hole in the upper or lower part of the diagram, these parts of the diagram commute and there is no Frobenius involved.

The avid reader may have spotted a fatal flaw already: the algebras must be commutative and associative in order to possess a Frobenius endomorphism: so much for mixed octonion algebras! But I had a grand scheme: I knew that the essential part $(V, \times)$ of an octonion algebra (the orthogonal complement of the identity) is a Lie algebra in characteristic 3 and so I wanted to define mixed Hopf algebras; from there I would work out the notion of a mixed Lie algebra—which would specialize to the notion of a Lie algebra with $p$-operation in the case where $\kappa$ is an isomorphism—and the mixed octonion algebra I would obtain later by adding the unit back to it one way or another.

I thought it was a bit a wacky plan, even for my standards, but when I proposed it to Tom De Medts, to my surprise he said that he thought it was actually quite clever and he asked if there was a way in which I could see the mixed groups themselves arising from this sort of formalism. Although I speculated idly that perhaps one could formally take the dual of the category of such mixed Hopf algebras and associate abstract groups to some of them, all of that seemed wild imagination at that point.
After I had worked out a few of the basic properties of mixed algebras it seemed only natural to shift the focus from algebras together with fields to simply rings.\(^2\) Although I had no idea how to prove it, I gradually became convinced that mixed rings were precisely the objects needed to describe Tits’s mixed groups in full generality, and so in particular the groups of type $\text{MG}_2$ that all my troubles had started with!

To explain how I came to that realization, I need go back in time a few years, to my first class in algebra. The year is 2009 and the professor is Jan Van Geel, who had the habit of asking all the students what they were expecting from the course. A few weeks earlier, I had been trying to solve some olympiad-style problems with Gaussian integers $\mathbb{Z}[i]$, and I had been very puzzled with what a prime number in such a ring was supposed to be, and if it should then also be an ordinary prime number or not. So my answer was that I expected to learn methods which tell us when a ring is a domain, or a domain is a field. But Prof. Van Geel told me that I was wrong—he told all of us that we were wrong that day, except for the one guy who said he expected a lot from the course and that he expected a fun course. Algebra, said Prof. Van Geel, is about the study of equations and more precisely polynomial equations. An example of a polynomial equation would be

$$x^{97} + y^{97} + z^{97} = 0.$$  

So with that wisdom in mind, let us stare at the equations which describe the mixed hexagon in a 6-dimensional projective space.

$$x^2 + aA + bB + cC = 0$$

$$A^2C - b^2c + xbA \in \ell$$

$$A^2B - bc^2 - xcA \in \ell$$

$$AB^2 - ac^2 + xCB \in \ell$$

$$B^2C - a^2c - xAB \in \ell$$

---

\(^2\)It is in fact a very classic observation in algebraic geometry—perhaps originally due to Grothendieck, although I am not sure—that by treating a field itself as the simplest kind of variety many classical theorems about varieties over a field can be generalized to theorems about morphisms between varieties (and more generally, morphisms between schemes).
\[ BC^2 - a^2b + xAC \in \ell \]
\[ AC^2 - ab^2 - xbC \in \ell \]
\[ abc + xbB - xcC + ABC \in \ell \]

Clearly the first equation belongs to the realm of algebra. But for the other seven, it occurs to me that although these have a certain *equational flavor*, they do not end with \( = 0 \) and therefore are not polynomial equations; I call them mixing equations. Many readers will agree with this, but maybe a wise guy will note that if \( \ell = k^3 \), we can introduce an auxiliary variable \( u \) and rewrite the equation \( f(x) \in k^3 \) as \( f(x) = u^3 \). And in fact, one can always choose a basis for \( k \) as an \( \ell \)-vectorspace and then still expand the equations in that basis, at least when it is finite—this is how people traditionally deal with such things.\(^3\) But on a conceptual level, it seems poor practice to deal this way with things that are intrinsically, well, *mixed*.\(^4\)

The modern way of looking at such a polynomial is through the formalism of *rational points*. For instance, if we are looking for complex solutions of the equation

\[ f(z) = a_0 + a_1z + \ldots + a_nz^n \in \mathbb{C}[z], \]

then we construct the ring \( A = \mathbb{C}[z]/(f(z)) \), which has a structure morphism \( q_A : \mathbb{C} \to A \) and the *morphisms* \( x : A \to \mathbb{C} \) such that \( x \circ q_A \) is the identity on \( \mathbb{C} \) correspond to the solutions of the equation \( f \); these morphisms are called the *rational points* of \( A \) and denoted \( \text{hom}(A, \mathbb{C}) \).

Such a rational point \( x \) is depicted on the following diagram:

\[ A \xleftarrow{q_A} \mathbb{C}. \]

\(^3\)This is the *Weil restriction*; as pointed out by Hendrik Van Maldeghem during the private defense of this thesis, the former procedure also generalizes whenever \([\ell : k^3] < \infty\), providing a strange counterpart for the Weil restriction.

\(^4\)Something similar is true for equations that arise in the context of the groups of type \( ^2\text{G}_2 \). They look like polynomials, but the exponents are not elements of \( \mathbb{N} \) but rather of \( \mathbb{N}[\theta] \), where \( \theta \) is an endomorphism of the field. Over a finite field \( \mathbb{F}_{p^{2e+1}} \), this \( \theta \) is actually a natural number \( \theta = p^{e+1} \) and thus there is a natural way to see these equations as polynomials after all. But this too seems a poor crutch to study something that is intrinsically *twisted*.\)
So a logical next step was for me to find out where my formalism leads to if I were to deploy the same mechanism for *mixed rings* instead. It turned out that extra equations that this formalism of mixed rings allows you to deals with, are precisely these mixing equations that end with $\in \ell$.

But this is only half of the story of mixed rational points. Let me illustrate this by drawing a rational point of the mixed algebra $M = (A_k, A_\ell)$ over the mixed field $m = (k, \ell)$ in dashed arrows on the following diagram:

$$
\begin{array}{c}
\text{Hom}(M,m) \rightarrow \text{Hom}(A_k,k) \times \text{Hom}(A_\ell,\ell)
\end{array}
$$

As you can see, a rational point $(x, y) \in \text{hom}(M, m)$ is actually a pair $x \in \text{hom}(A_k, k)$ and $y \in \text{hom}(A_\ell, \ell)$ such that a certain correspondence between $x$ and $y$, which can be read off from the diagram, is satisfied. This implies that there is an embedding

$$
\text{hom}(M, m) \rightarrow \text{hom}(A_k, k) \times \text{hom}(A_\ell, \ell)
$$

which we can compose with the projection on either component to obtain a pair of maps

$$
\text{hom}(M, m) \rightarrow \text{hom}(A_k, k) : (x, y) \mapsto x \\
\text{hom}(M, m) \rightarrow \text{hom}(A_\ell, \ell) : (x, y) \mapsto y.
$$

The crucial observation here is that *both maps are injective*. (This relies on the fact that the Frobenius endomorphism of a field is always injective.) What this practically means is that the set $\text{hom}(M, m)$ can be described by ordinary equations over the field $k$, together with mixing equations ending with $\in \ell$; or by ordinary equations over the field $\ell$, together with mixing equations ending with $\in k$. This is precisely the sort of mixed personality that one would expect in a theory which aims to describe strange groups and at that point, I had no doubt anymore that they would fit into the picture.
One thing to keep in mind though is that even though it is \textit{in principle} possible to expand most algebra back into polynomial equations, it is not necessarily a clever idea to do so. In this context, I like to cite Galois\textsuperscript{5}:

If now you give me an equation that you have chosen at will, and about which you want to know if it is or is not solvable by radicals, I cannot do any more than indicate the means for answering your question, without wanting to charge myself or any other person with doing it. In a word, the calculations are impractical.

And so I developed the theory of mixed rings, which is really just a theory of equations that end with $\in \ell$ instead of $= 0$. At one point, I noted that it would not be enough to deal with rings alone. The reason is that certain objects, such as the mixed hexagon that I copy/pasted in all its glory just a few pages earlier—and more generally certain coset spaces of algebraic groups—are not \textit{affine} but \textit{projective} varieties. At that point I faced with the choice of either trying to understand what mixing means for graded rings or for schemes. It seemed to me the latter approach was the easier one\textsuperscript{6}, and so I made the transition from rings to schemes. (I didn’t really know what a scheme was when I started, but I learned along the way.)

What still nagged at me at that point, was that each time that I wanted to write down a mixed field $(k, \ell)$, I still had to choose which field to put in the first slot and which field to put in the second slot! So I had the idea to consider the ring $k \times \ell$ instead, which is canonically isomorphic to $\ell \times k$. The maps between $k$ and $\ell$ now give rise to a single map

$$
\varphi : k \times \ell \to k \times \ell : (u, v) \mapsto (\lambda(v), \kappa(u))
$$

with the property that $\varphi^2$ is the Frobenius on $k \times \ell$. A familiar sight, because this is precisely the type of thing that occurs in the study of the twisted groups! I have tried to draw this on a diagram as well, but this time I had to do something really awkward and draw the square

\footnotesize
\textsuperscript{5}The translation is from [Cox12]

\textsuperscript{6}I was completely horrified when I found out the Proj-construction is not functorial.
root of a black hole in the middle, since we need to loop around it twice to get to the Frobenius.

\[
\begin{array}{c}
\Delta \\
\varphi
\end{array}
\]

\[
k \times \ell
\]

\[
\sqrt{\bullet}
\]

It would lead me too far to explain precisely what the relation is between diagrams with a black hole in them and those with the square root of a black hole in them, but suffice it to say one can always untwist such a diagram back to a diagram that looks like this:

\[
\begin{array}{c}
\varphi \\
\Delta
\end{array}
\]

\[
k \times \ell
\]

\[
\bullet
\]

\[
k \times \ell
\]

Understanding which mixed diagrams arise in this fashion\(^7\) is a way to describe very precisely the interplay between structures of a twisted nature, such as \(2G_2\), and structures of a mixed nature such as \(MG_2\)—I dare say this is something that was not understood very well before my work.

With those discoveries in my pocket, I gave a 30 minute talk at the Mini Course in Lens on May 30th of 2016 on the subject of mixed groups as rational points of mixed algebraic groups. After the talk, Michel Brion came talk to me for a few minutes, and he asked if there was some relation with the pseudo-reductive groups that had been studied quite recently by Conrad, Gabber and Prasad. I had had the first edition of their book on my desk for a while at that point, but I

\(^7\)The answer is: those diagrams that are invariant under rotation over 180°! See Proposition 8.3.6.
had never looked at it much. The formalism seemed too heavy and my knowledge of algebraic geometry too poor. What I did know about their work was that many pseudo-reductive groups were obtained as Weil restrictions of reductive groups, and so I speculated that perhaps some of their weird pseudo-reductive groups were actually Weil restrictions of mixed groups. I’m not sure what Brion said after that, but he seemed not to disapprove of the idea which I found quite encouraging. So when I got back to my office later that week, I decided to open the big bad book on pseudo-reductive groups [CGP15] and have a better look at what was inside. I didn’t understand much of the book but it seemed blatantly obvious to me that one of their classes of pseudo-reductive groups should arise as a Weil restriction of a mixed group and I could prove this quite easily.

At least, that’s what I thought. In my alleged proof, I used a very standard mathematical formalism—an adjoint pair of functors—but in a context where it does not apply. What this formalism does is this: it allows you to start from an $S$-diagram, and draw a $T$-diagram and vice versa. Let me try to explain in some layman’s terms how this works. The objects that algebraic geometry is typically concerned with can be depicted as follows.

\[ \begin{array}{c}
X \\
\downarrow \\
S
\end{array} \quad \text{or} \quad \begin{array}{c}
Y \\
\downarrow \\
T
\end{array} \]

These are the simplest kinds of $S$-diagrams and $T$-diagrams. If in addition to such a diagram, we also have been given an arrow $\beta : T \to S$, then there exists a procedure $\beta^*$ which transforms an $S$-diagram into a $T$-diagram and there is a procedure $\beta_*$ in the opposite direction.

\[ \begin{array}{c}
X \\
\downarrow \sim \rightarrow \\
S
\end{array} \quad \beta^* \begin{array}{c}
X \\
\downarrow \\
T
\end{array} \quad \text{and} \quad \begin{array}{c}
Y \\
\downarrow \sim \rightarrow \\
T
\end{array} \quad \beta_* \begin{array}{c}
Y \\
\downarrow \\
S
\end{array} \]

These procedures go by many names, such as extension of scalars/base change and restriction of scalars/Weil restriction but I will now simply call this the $(\beta^* \dashv \beta_*)$-formalism. We can also apply this procedure
to vastly more complicated $S$-diagrams. For instance, consider the diagram

\[
\begin{array}{c}
\beta^*X \\
| \quad \beta^*w \quad \beta^*v |
\end{array}
\begin{array}{c}
\beta^*Y \\
| \quad f^2 \quad h |
\end{array}
\begin{array}{c}
\beta^*Z \\
| \quad g^2 \quad T |
\end{array}
\begin{array}{c}
\beta^*U \\
\beta^*V \\
\beta^*V
\end{array}
\]

This is an example of an $S$-diagram because every object on it has an arrow pointing down towards $S$. (We have only drawn the arrows $Z \to S$ and $\beta^*U \to S$ because the diagram would get too crowded with arrows.) It is also a commuting diagram, which in particular means that paths with the same source and target are the same; by convention the identity morphisms are implicitly present on the diagram as well, so this means in particular that all loops are equal to the identity. The $(\beta^* \dashv \beta_*)$-formalism then transforms this $S$-diagram into the following $T$-diagram:

My proof basically consisted of applying this formalism twenty zillion times\(^8\) until I obtained the right diagram. Unfortunately, I later noticed that I had also applied it in a context where this is not allowed! The reason is that my diagrams had a black hole $\bullet$ in them and therefore they are not commuting diagrams—which is required by the formalism. But the next week I found an easy crutch and considered the proof finished, up to adding a reference where my easy crutch was proven.

\(^8\)Well, maybe 5 times. See Proposition 8.7.1.
At that point, I strongly felt the limitations of my scheme-theoretic
description which couldn’t distinguish well between scheme-theoretic
properties of the Frobenius endomorphism (such as that it is injective
on fields) and categorical properties. This and some concrete problems
convinced me to execute a major overhaul of my work where I shifted
focus from the category of schemes to an arbitrary category with good
properties.

I had emailed Tom De Medts about this, but his reply didn’t sound
particularly enthusiastic. He asked me if I had actually won anything
with it, and he noted that many people that were interested in twisted
and mixed groups and varieties (incidence geometers, group theorists
and algebrists, including himself) would not be able to follow how the
framework was related to these objects. Although I regretted this a bit,
at the time I could not imagine that someone would prefer to ignore
possibly interesting work because the terminology was unfamiliar—
after all, if I had learned these ideas along the way, why couldn’t
they? (I have since turned around a bit and provided a crash course
in Appendix C.)

As I tried to wrap up my work and add detailed proofs, I found out
that my easy crutch—to apply the $(\beta^* \dashv \beta_*)$-formalism to $\bullet$-diagrams—
didn’t work as expected. It took me more than a month to come up
with a new crutch. Technically, this is certainly the hardest part of
my work and I’m afraid if some things will not stand up to scrutiny,
this will be it.

The last thing I did was proving the insight with which it had all
started: that all mixed groups are groups of rational points of an
algebraic mixed group. I had delayed this a bit because it required
me to delve deeper into algebraic group theory than I was comfortable
with but in the end it turned out that everything worked as expected.
In contrast to the Steinberg groups, the Suzuki–Ree groups cannot be easily viewed as algebraic groups over a suitable subfield; morally, one ‘wants’ to view $^2B_2(2^{2n+1})$ and $^2F_4(2^{2n+1})$ as being algebraic over the field of $2^{(2n+1)/2}$ elements (and similarly view $^2G_2(3^{2n+1})$ as algebraic over the field of $3^{(2n+1)/2}$ elements), but such fields of course do not exist.

Terence Tao

This chapter aims to present some of the core ideas behind our work, in a manner that is accessible to a wide mathematical audience, sometimes at the cost of precision. It is not meant as a historical overview of known facts, for that see Appendix A. We also point out that in Section 10.1 the technical content of our main theorems will be explained in more detail.

Section 7.1 sketches what our work accomplishes while shunning technical definitions or terminology. Section 7.2 explains some of the problems that we see with the existing theories about twisted and mixed groups; these problems are usually of the type ‘morally speaking $X$ is actually $Y$, but we cannot make this rigorous’. In Section 7.3 we will then explain how to come to proper definitions that accomplish the framework that we set out.

---

1https://terrytao.wordpress.com/2013/09/05/notes-on-simple-groups-of-lie-type
Chapter 7. Mathematical summary

7.1 Overview

One of the cornerstones of algebraic group theory is the structure theory of semi-simple groups over an algebraically closed field, which is largely due to Chevalley. The theory was quickly extended to a theory of (connected) reductive groups over arbitrary fields, and in fact over an arbitrary base scheme, by many others most notably Borel, Tits, and the authors of [SGA3]. However, during the second half of the 20th century, on a number of occasions, groups have been encountered which appear to be closely related to reductive groups, but in a strange, exotic manner.

The first time such an encounter happened was around 1960. In the process of classifying a class of finite simple groups, Suzuki discovered a new class, now known as the Suzuki groups. His discovery was a precursor for the discovery of a more general construction by Ree later that year which also produced other, similar classes of groups: the twisted Chevalley groups. Somewhat later—probably around 1970—Tits was studying reductive groups by means of his theory of buildings. As he was classifying certain buildings, he discovered that although most of these buildings came from reductive groups, there were a few that were related, but not so directly: these were the mixed buildings and groups. In 1997 then, Weiss completed the classification of another class of buildings and discovered groups that are arguably even stranger, but nonetheless still recognisable as distant cousins of reductive groups. Finally, around 2010 Conrad, Gabber and Prasad, as part of their structure theory and classification for a class of algebraic groups named pseudo-reductive groups, discovered that—as the name seems to suggest!—most pseudo-reductive groups are closely related to reductive groups. But again, there are some estranged and exotic family members that are more distantly related.

All of these occurrences share two features. The first of these is that it is always the combinatorics of root systems with roots of two different lengths that makes the construction work; in other words one of the Dynkin diagrams $B_n$, $C_n$, $F_4$ or $G_2$ plays an important role. The second feature is that the constructions require certain ingredients which are very specific to ‘positive characteristic mathematics’, in
particular they always require the Frobenius endomorphism of a field or algebraic group and sometimes also depend on the occurrence of inseparable field extensions.

These two features are of course closely related. For instance: the ratio of roots in the root system under consideration forces the characteristic of the field to be 2 or 3. Nonetheless, it has been our point of view that in order to deepen our understanding of these groups, we should untangle these two features as well as we can. Throughout most of our work, we will focus on the second aspect, i.e. the ‘positive characteristic mathematics’. It is easy enough to smuggle in the combinatorics of root systems via a backdoor, namely by assuming the existence of certain isogenies, thereby ultimately relying on combinatorial properties of root systems that are well understood. But in our approach, this is really more an afterthought.

Let us summarize the main idea behind our approach. Traditionally, an algebraic group is defined as a group object in the category of schemes. But if we are considering the category of schemes for a fixed positive characteristic only, there is a special gimmick under the guise of the absolute Frobenius. We use this gimmick to create new but closely related categories: the categories of twisted and mixed schemes. Then there are also group objects in these categories, and this is where these exotic abstract and algebraic groups find their origin.

More specifically, recall that an ordinary algebraic group over a field $k$ has a functor of points $K \mapsto G(K)$ which produces an abstract group $G(K)$ out of every $k$-algebra $K$. The same procedure, but in the setting of twisted and mixed group schemes, will produce a large number of abstract groups out of twisted or mixed rings and these are the groups that we are after.

The category of twisted schemes that we will construct lies a bit deeper than the category of schemes, and it can be thought of as the category of schemes over $\text{Spec} \mathbb{F}_{\sqrt{p}}$. (The field with $\sqrt{p}$ elements!) There is an embedding$^3$ $m : X \mapsto mX$ which takes an ordinary scheme$^3$

$^2$I.e. a faithful (but not full) functor.

$^3$Actually this functor will be denoted $\delta \circ m$ in Section 8.1 but here and in
and produces the corresponding twisted scheme. By an abuse of notation, we will then denote $mX$ simply by $X$; in particular the notation “$\text{Spec } F_p$” now also denotes a twisted scheme. If we apply the functor $m$ to a scheme $X$ together with its unique structural morphism $X \to \text{Spec } F_p$, we obtain a twisted scheme which is defined over $F_p$.

To make this a little less abstract, let us consider an ordinary algebraic group $G$ over $F_p$ together with a homogeneous space $\mathcal{H}$ and map $\pi : G \to \mathcal{H}$, all defined over $F_p$. This is usually depicted as a diagram of schemes as follows.

$$
\begin{tikzcd}
G \ar[r, \pi] \ar[d] & \mathcal{H} \ar[d] \\
\text{Spec } F_p
\end{tikzcd}
$$

But if we embed this in the category of twisted schemes, there are many other things happening that we were unaware of:

$$
\begin{tikzcd}
G \ar[r, \pi] \ar[rd] & \mathcal{H} \ar[d] \ar[ld] \\
\text{Spec } F_p & 2G \ar[r] & \text{Spec } F_{\sqrt{p}^3} \\
& \text{Spec } F_{\sqrt{p}}
\end{tikzcd}
$$

First of all, there are a number of alien objects which are defined over $F_{\sqrt{p}}$ but not over $F_p$. For instance, we have suggestively drawn an object denoted by $2G$ and an object $\text{Spec } F_{\sqrt{p}^3}$ on this diagram. But there is a second surprise: there are twisted schemes, which are defined over $F_p$, but do not come from an ordinary scheme! These imaginary or invisible objects are depicted by the $\Im$ on the diagram and they are the ones responsible for Tits’s mixed groups. So if the rest of this chapter we have given up on precision in favor of the clarity of the exposition.
we dub the twisted schemes over $F_p$ *mixed schemes*, then we have observed that there are mixed schemes which are not ordinary.

There are some very important subtleties to look out for. For instance, if $X \to \text{Spec} F_p$ is a mixed scheme then this structural morphism is unique as a morphism of mixed schemes, but not as a morphism of twisted schemes. For instance, within the mixed schemes, there is just one morphism from $\text{Spec} F_p$ to itself; but as a twisted scheme, there are two! Another way of saying this is that its automorphism group in the category of twisted schemes has order 2; in other words we have that “$\text{Gal}(F_p/F_{\sqrt{p}}) \cong C_2$”.

Our main theorems are three applications which show how the strange groups that we introduced at the beginning of this section arise naturally in this set-up. Let us quickly summarize this. Theorem 10.3.1 says that the twisted groups of Suzuki and Ree arise as group objects in the category of twisted schemes. Theorem 10.4.1 says that the mixed groups of Tits arise as group objects in the category of mixed schemes. Moreover, the twisted and mixed groups are closely related with one another: extending scalars along the extension $F_p/F_{\sqrt{p}}$ turns a twisted group into a mixed group, and there is a descent criterium Proposition 8.3.6 which can be used to figure out if a given twisted group arises this way out of a mixed group. The exotic groups of Conrad, Gabber and Prasad are also closely related to these invisible mixed reductive groups: our Theorem 10.5.1 states that they arise as Weil restrictions of mixed reductive groups. Finally, the exotic groups of Weiss are one class of groups for which we do not explicitly show that they can be described with the tools under our belt. We believe that there is enough evidence to be convinced that this is indeed the case in a precise manner, but it may require some technical virtuosity to show this rigorously.

Finally, a note on the techniques used. Since the essential bit of information that we wish to manipulate consists of a category $\mathcal{C}$ together with a gadget $F$, this is exactly what we study throughout most of the work. We believe that this approach offers a resting point halfway in between the difficulties of scheme theory and the

\footnote{Technically correct: an endomorphism of the identity functor.}
greater level of abstraction that would be required to explore the
depths below $F_p$ and $F_{\sqrt{p}}$. If we worked with schemes from the get-
go, we would be tempted to prove superficial theorems only about
the schemes that we like—reduced, of finite type over a field,—without achieving a deep appreciation for those facts which originate
from the categorical nature of the Frobenius endomorphism, thereby
contrasting those facts which ultimately say more about ourselves
and our preference for certain schemes over others than about the
Frobenius endomorphism. But if we would work too abstractly with
monoids of endomorphisms of identity functors, or in the framework
of $m$-categories for which we have sketched the hypothetical countours
in Section B.1, it would become much harder to figure out the right
statements and proofs and connect these with known objects such as
Suzuki-Ree groups. We hope that the reader, even if they do not
agree with our choices in this matter, will at least appreciate that we
did not strafe further away from their own point of equilibrium.

7.2 Three classes of exotic groups: problems

Let us examine the exotic groups that we spoke of in some more detail,
and focus on the conceptual problems with their standard definitions.

In 1960, Suzuki found an infinite class of finite simple groups, explicit-
itly described by matrices as subgroups of $GL(4, 2^{2e+1})$. Not much
later, Ree showed how these Suzuki groups can be obtained from a
Chevalley group of type $B_2$ as follows: if the characteristic of the
underlying field $k$ is 2, the Chevalley group has an exceptional graph
automorphism $g$ with the strange property that it squares to the
Frobenius endomorphism. So if we assume that the field $k$ admits
an automorphism $\varphi$ such that $x^{2^2} = x^2$, then $\varphi^{-1} \circ g$ is an auto-
morphism of order 2 of the group $B_2(k)$ and its fixed points form
a subgroup $^2B_2(k, \varphi)$ which is a Suzuki group. Extending this pro-
cedure to Chevalley groups of type $G_2$ in characteristic 3—where
the condition on $\varphi$ becomes $x^{\varphi^2} = x^3$—and $F_4$ in characteristic 2,
Ree found the small Ree groups $^2G_2(k, \varphi)$ and the large Ree groups
$^2F_4(k, \varphi)$. Somewhat later, Tits showed how to define these groups
over non-perfect fields, when $\varphi$ is a non-invertible endomorphism of $k$;
essentially by rewriting \( x = \varphi^{-1}(g(x)) \) as \( \varphi(x) = g(x) \). These groups are now known as twisted (Chevalley) groups or as the Suzuki-Ree groups; an endomorphism of a field which squares to the Frobenius endomorphism is often called a Tits endomorphism.

There are a number of moral grounds based on which one could object to this description. Perhaps most strikingly, in the celebrated classification of finite simple groups, all the infinite non-abelian classes are closely related to a (semi-simple) algebraic group, and this is a useful tool to initiate the study of these groups. The exceptions to this rule are the alternating groups \( \text{Alt}_n \)—algebraic groups over the ‘field of one element’, whatever that is!—and the Suzuki-Ree groups. The analogy with the Steinberg groups such as the unitary groups, which are superficially abstract groups ‘over \( \mathbb{C} \)’ but realized as algebraic groups over \( \mathbb{R} \), suggests that the Suzuki-Ree groups are algebraic groups over a field with \( \sqrt{p} \) elements—see the blog post by T. Tao cited under the title of this chapter. At the same time, it is well understood that for many purposes algebraic groups over \( \mathbb{F}_p \) should really be treated as a pair \( (G, F_G) \) consisting of a group \( G \), together with its Frobenius \( F_G \)—but since the Frobenius is canonical, it can be omitted from the description. Similarly, the Suzuki-Ree groups can be seen as a pair \( (G, \Phi_G) \) where \( G \) is an algebraic group and \( \Phi_G \) a square root of \( F_G \) and it turns out that one can often recycle techniques used for \( (G, F_G) \) on \( (G, \Phi_G) \) without adaptation.

The second exotic class of groups which merits our attention has not been studied nearly as well: the groups of mixed type or mixed groups for short. Besides a parenthetical remark in Steinbergs lecture notes [Ste68, p. 153], their first appearance in the literature seems to be in Tits’s lecture notes [Tit74, (10.3.2)] on buildings. Tits introduces buildings as a tool to study algebraic groups, and achieves a complete classification of the important class of spherical buildings of rank \( \geq 3 \). The classification does mostly what it was intended to do: with some exceptions all of these buildings originate from algebraic groups or classical groups.\(^5\) The exceptions are the buildings of mixed type. These are buildings of type \( \Gamma_n = B_n/C_n, F_4 \) or \( G_2 \) whose definition

\(^5\)For instance \( \text{PSL}_n(D) \) is classical but non-algebraic if \( D \) is a division ring of infinite dimension over its center.
requires a pair of fields $k, \ell$ of respective characteristic $p = 2, 2$ or $3$ such that $\ell^p \subseteq k \subseteq \ell$. In a nutshell, Tits’s construction comes down to observing that one can restrict some of the generators of the group $X_n(\ell)$ to the subfield $k$ in a way which preserves the commutation relations which define the group. In this way one gets a subgroup which acts on a subbuilding. Tits’s notation for these abstract groups is then $X_n(k, \ell)$, where of course $X_n(\ell, \ell) = X_n(\ell)$. These groups act on a building, and there is a corresponding BN-pair which can be used to study them.

Several problems arise with these mixed groups. First, they are merely abstract groups and there is no corresponding algebraic group. Second, since they are defined very explicitly by means of a set of generators, the construction only works well for split groups. Third, it was one of Tits’s favorite observations that the fields $k$ and $\ell$ involved play essentially the same role! In his lectures, he mentioned that for him, a mixed group was defined over the infinite sequence of fields

$$\ldots \subseteq \ell^2 \subseteq k^p \subseteq \ell^p \subseteq k \subseteq \ell \subseteq k^{1/p} \subseteq \ell^{1/p} \subseteq \ldots$$

Moreover, Tits notes that [Tit74, p. 204] $B_n(k, \ell) \cong C_n(\ell^2, k)$, while $G_2(k, \ell) \cong G_2(\ell^3, k)$ and $F_4(k, \ell) \cong F_4(\ell^2, k)$. (And of course $\ell^p \cong \ell$.) But this phenomenon is not apparent from the description, where one field is the larger field and one the smaller. Finally, it is intuitively clear that there is some connection with the twisted groups, but it seemed difficult to make this precise.

Finally, we introduce the class of pseudo-reductive groups. Recall that a linear algebraic group $G$ over a field $k$ is reductive if the unipotent radical of $G_\overline{k}$, the base change of $G$ to the algebraic closure $\overline{k}$, is trivial. The corresponding notion over $k$ is weaker and called pseudo-reductivity. Where the structure theory of reductive groups has been an important part of algebraic group theory for more than half a century, a structure theory for pseudo-reductive groups is a rather recent addition due to Conrad–Gabber–Prasad [CGP15], building on some older work of Tits. Although we found the structure theory in its entirety not particularly accessible, the gist of it can be phrased rather elegantly by saying that that most pseudo-reductive groups arise from a certain standard construction. The standard construction takes as input a reductive group $G'$ over a field $k'$, and a purely
inseparable field extension $k'/k$. It then applies a process known as Weil restriction to $G'$ to obtain a group $G = R_{k'/k}G'$ over $k$ which is pseudo-reductive. In fact, this is only the easy part of the standard construction: there is another important step which consists of replacing the Cartan subgroups by a different commutative group which satisfies certain conditions, but this will not play a role in our work and we refer to [CGP15, §1.4] for details. Exceptional groups which do not arise from the standard construction exist only in characteristics 2 and 3. A first class of exceptions consists of the exotic pseudo-reductive groups, which are constructed and studied in Chapter 7 of [CGP15]. The authors first construct what they call very special isogenies between certain algebraic groups and then perform a rather elaborate construction which starts from such an isogeny $\pi : G \to \overline{G}$ together with a purely inseparable field extension $k'/k$ to produce a pseudo-reductive group $\mathcal{G}$. In their own words, the construction roughly comes down "thickening the short root groups from $k$ to $k'$ and part of the torus from $k^\times$ to $k'^\times". Clearly as an abstract group, this should correspond to the mixed group $X_n(k,k')$ and in fact, we have no doubt that this is how Tits found them!

We cannot stop the reader from concluding that these pseudo-reductive groups are the groups associated to the mixed buildings, but we do not share this belief for a couple of reasons. First of all, it is not very elegant that most of Tits’s spherical buildings correspond to reductive groups, but the mixed buildings suddenly require a few pseudo-reductive groups to finish the picture. Another issue is that if both dimensions $[k : k']$ and $[k' : k^p]$ are infinite, these groups cannot be constructed (except perhaps as pro-algebraic groups). Worse even, the symmetry between $k$ and $k'$ that Tits was so fond of is completely destroyed in the process. On the other hand, we do get a pseudo-reductive $k'$-group $\mathcal{G}'$ with the same group of rational points: $\mathcal{G}'(k) \cong \mathcal{G}'(k')$, but in general the dimensions of these groups will be different and depend on $[k : k']$ and $[k' : k]$. Finally, the theorem which states that most pseudo-reductive groups arise from Weil restrictions of reductive groups, already suggests that perhaps there should exist some kind of imaginary of invisible reductive groups $G$, defined over some invisible field $m$ which extends both $k$ and $k'$ of which these exotic pseudo-reductive groups should be Weil restrictions: $\mathcal{G} = R_{m/k}G$
Chapter 7. Mathematical summary

and $\mathcal{G'} = R_{m/k} G$, so that $\mathcal{G}(k) \cong G(m) \cong \mathcal{G'}(k')$.

So with a little dramatization and within the approximations mentioned, we come to the following picture. Our work consists of enlarging the class of algebraic groups, and in particular the reductive groups, to obtain something that provides a better match with the other three classes.

### 7.3 Three classes of exotic groups: solutions

As the notation $^2X_n(k, \theta)$ suggests, the groups of Suzuki and Ree can be defined over an arbitrary pair $(k, \theta)$ consisting of a field $k$ of the appropriate characteristic together with an endomorphism $\theta : k \to k$ which squares to the Frobenius—this is often called a *Tits endomorphism*. It is a very fruitful idea in algebraic geometry that a field is the simplest kind of variety (or scheme). This suggests that the Suzuki-Ree groups should be some sort of algebraic groups in a context where the fields with Tits-endomorphisms are the simplest objects. So we define a twisted ring (resp. scheme) as a pair $\tilde{X} = (X, \theta)$ where
$X$ is a ring (resp. scheme) endowed with a *twister* $\theta : X \to X$, which is an endomorphism of $X$ which squares to the absolute Frobenius. With the obvious notion of morphism, we obtain a category of *twisted* rings (resp. schemes).

Before we continue to the mixed rings, we must clear up an important point. The reader probably thinks that there are far fewer twisted rings than ordinary rings, since once can always strip a twisted ring of its twister to obtain an ordinary ring. In fact, the opposite is true! After all, every ordinary ring $R$ gives rise to a twisted ring

$$m(R) = \left( R \times R, (x, y) \mapsto (y^p, x) \right).$$

This assignment can be extended to morphisms and by an abuse of notation, we may also denote $m(R)$ simply by $R$. In particular every ring $R$ together with its unique structural morphism $q_R : \mathbb{F}_p \to R$ gives rise to a morphism

$$m(q_R) : m(\mathbb{F}_p) \to m(R) : (x, y) \mapsto (q_R(x), q_R(y))$$

of twisted rings. We should warn the reader though that as twisted rings, there are additional morphisms between $m(\mathbb{F}_p)$ and $m(R)$, which the reader can try to determine.

Clearly, the simplest twisted ring is given by $(\mathbb{F}_p, \text{id})$; we define this twisted ring to be $\mathbb{F}_{\sqrt{p}}$. Note that every twisted ring $\tilde{X}$ admits a unique morphism $\mathbb{F}_{\sqrt{p}} \to \tilde{X}$, so the role of $\mathbb{F}_{\sqrt{p}}$ in the theory of twisted schemes is comparable to that of $\mathbb{F}_p$ for ordinary schemes in characteristic $p$. Let us justify the “$\sqrt{p}$” in the notation $\mathbb{F}_{\sqrt{p}}$ in two different ways. A first heuristic is found by observing that the set underlying $m(k)$, for a field $k$, is given by $k \times k$. This suggests defining the cardinality of a finite mixed ring as the square root of the cardinality of the underlying set and so $\mathbb{F}_{\sqrt{p}}$ should be some analogon of $\mathbb{F}_p$ with $\sqrt{p}$ elements. Another, more compelling argument is that it extends the validity of the identity $\text{Aut}(\mathbb{F}_{p^2}/\mathbb{F}_p) \cong \mathbb{C}_2$ to all half-integers.

---

6 With a cheap analogy: when an infinite bus of new schemes arrived at Hilbert’s hotel for schemes, he kindly requested every field $k$ to move to $k \times k$, and suddenly there were plenty of empty rooms for the newcomers.
The upshot of all this is that this provides a context where the groups of Suzuki and Ree are algebraic groups over the field with \( \sqrt{p} \) elements. (Theorem 10.3.1) The reader may wonder what happens if one considers the base change of, say a Ree group \( ^2G_2 \otimes_{F_3} F_3 \) and speculate that the outcome is the usual Chevalley group \( G_2 \). But actually the answer is a bit surprising: we get a mixed group instead!

Before we can explain this, let us delve a bit deeper into the theory of mixed groups, which Tits sometimes saw as being defined over an infinite chain of fields. A different way of saying this is that these groups are defined over 4-tuples \( (k, \ell, \kappa, \lambda) \), where \( k \) and \( \ell \) are fields, and \( \kappa : k \to \ell \) and \( \lambda : \ell \to k \) are maps such that the compositions \( \kappa \circ \lambda \) and \( \lambda \circ \kappa \) are equal to the Frobenius on \( \ell \) and \( k \). If we call such an object a mixed field, and we continue our philosophy that these fields should be the simplest objects in a context where the mixed groups are algebraic groups, we are automatically led to the definition of a mixed ring (resp. scheme) by replacing \( k \) and \( \ell \) by rings (resp. schemes). With the obvious notion of morphism, we obtain a category of mixed rings (resp. schemes).

We can now observe that every mixed ring \( M = (R, S, \varphi, \psi) \) immediately defines a twisted ring

\[
(R \times S, (u, v) \mapsto (\psi(v), \varphi(u))).
\]

This twisted ring admits a unique map from \( F_p \) (which is still an abuse of notation for \( m(F_p) \)) which encodes the fact that \( R \) is the first element of the tuple, and \( S \) the second. So every mixed ring becomes an algebra over the twisted ring \( F_p \) and in fact, the converse is also true, and so we obtain a one-to-one correspondence between twisted \( F_p \)-algebras and mixed rings. For instance, we saw that every ordinary ring \( R \) with its structural morphism \( q_R : F_p \to R \) gives rise to a twisted \( F_p \)-algebra \( m(R) \), and this algebra is an incarnation of the mixed ring

\[
(R, R, x \mapsto x^p, x \mapsto x).
\]

So we now have a context where Tits’s mixed groups are indeed algebraic groups and there exist mixed groups which are responsible for these mixed buildings. (Theorem 10.4.1) For now, we will denote
them by $\text{MB}_n$, $\text{MC}_n$, $\text{MF}_4$ in characteristic 2 and $\text{MG}_2$ in characteristic 3. (Of course the ordinary groups $\text{B}_n$, $\text{C}_n$, $\text{F}_4$, $\text{G}_2$ and all the others are still there in arbitrary characteristic.)

We can now understand the base change $\mathbb{R} \rightsquigarrow \mathbb{R} \otimes_{\mathbb{F}_p} \mathbb{F}_p$ as follows: it sends twisted rings to mixed rings, in the following manner:

$$(R, \theta) \rightsquigarrow (R, R, \theta, \theta).$$

In general, it is very unlikely that this mixed ring will be of the form $m(S)$ for an ordinary ring $S$. Conversely, if we have an ordinary mixed ring $R$ (so $R$ is an abuse of notation for $m(R)$) then it is very unlikely that there will be a twisted ring $S$ such that $R = S \otimes_{\mathbb{F}_p} \mathbb{F}_p$: the reader can try to verify that this will only happen if there is an automorphism $\alpha : R \to R$ such that $\text{Fr}_R \alpha^2 = \text{id}_R$. In other words, not only must the Frobenius have a square root, it must also be invertible! (This is why it is easier to define Suzuki- and Ree-groups over perfect fields.) So the reader will now not be too surprised to learn that

$$\text{MG}_2 = {^2}\text{G}_2 \otimes_{\mathbb{F}_\sqrt{3}} \mathbb{F}_3.$$ 

This suggests the following approach to twisted rings (or schemes). Instead of studying them directly, study the mixed rings and try to endow them with extra information which permits to define a twisted ring. (Proposition 8.3.6.)

Finally, we must note that sometimes mixed groups will invade our ordinary, unmixed world, instead of minding their own business. What goes on is that every mixed field $m = (k, \ell)$ is closely related to an ordinary field $k = (k, k)$ as follows

$$\begin{array}{ccc}
k & \xrightarrow{\kappa} & \ell \\
\downarrow{\text{id}_k} & \text{Fr}_k & \uparrow{\kappa} \\
k & \xrightarrow{\lambda} & k.
\end{array}$$

There exists a construction $\mathbb{R}_{m/k}$ called Weil restriction which takes a $(k, \ell)$-algebra $A = (A_k, A_\ell)$ and turns it into a mixed $k$-algebra $\mathcal{A}$:

$$\mathcal{A} = \mathbb{R}_{m/k} A.$$
This construction is easiest to describe at the level of the functor of points: an algebra $A$ can be thought of as a functor $M \mapsto (\text{Spec } A)(M)$ which sends every $m$-algebra $M$ to a set $(\text{Spec } A)(M) = \text{hom}_{m\text{-alg}}(A, M)$. The algebra $A$ is uniquely determined by the following functor on (mixed) $K$-algebras:

$$K \mapsto (\text{Spec } A)(K) = (\text{Spec } A)(K \otimes_k m).$$

Although $(\text{Spec } A)(K)$ is defined for every mixed $K$-algebra, we can restrict the domain to the ordinary $K$-algebra’s. This gives us an ordinary algebra which must carry features from both halves of $A$ since, for instance, there is an embedding $(\text{Spec } A)(k) = (\text{Spec } A)(m) \hookrightarrow (\text{Spec } A_\ell)(\ell)$.

This is exactly where the exotic pseudo-reductive groups come from: they are (one half of) a Weil restriction of one of Tits’s mixed groups. (Theorem 10.5.1.) The conceptual advantage is twofold: on the one hand, Tits’s spherical buildings still correspond to (mixed) reductive groups. On the other hand, the standard construction of Conrad–Gabber–Prasad now also produces the exotic examples.

### 7.4 The language of category theory

We will rely heavily on the language and elementary properties of categories; in particular we will frequently use the notions of slice categories, (co)limits, adjoint functors and the Yoneda embedding. The uninitiated reader may find it useful to keep a reference at hand or to take a look at Appendix C where we recall the most important definitions.

The standard reference in the field is [Mac71], but we also suggest the more recent [Lei14] which is freely available online on the arXiv (albeit with humongous margins). The latter reference covers less ground but is very accessible and contains everything that is required for our purposes. Finally, exposés I and II of [SGA3] and the nLab on ncatlab.org are highly recommended for some more in-depth coverage of certain topics that we will indicate throughout the text.
7.4. The language of category theory

Our notations are mainly inspired by [SGA3]: we denote a slice category of a category \( \mathcal{C} \) over an object \( X \) by \( \mathcal{C}/X \); we denote the structural morphism of an object \( Y \in \text{Ob}(\mathcal{C}/X) \) typically by \( q_Y \) (this makes \( Y \) an abuse of notation for \( q_Y \)); we denote the categories of sets, rings, schemes by \((\text{set}), (\text{ring}), (\text{sch})\); we use the arrow \( \sim \) to define a functor on objects if we leave the definition on arrows to the reader; we denote the category of presheaves on \( \mathcal{C} \) by \( \mathbb{C} = \text{hom}(\mathcal{C}^{\text{op}}, (\text{set})) \); we denote the internal hom in a category by the boldface \( \text{hom} \); we denote the Yoneda embedding by \( \mathcal{C} \rightarrow \mathcal{C} : X \leadsto h_X \) with \( h_X(Y) = \text{hom}(Y,X) \); we denote the functor of base change along an arrow \( f \) in a category which admits fibered products by \( f^* \); we denote an adjunction of functors \( L : \mathcal{C} \rightarrow \mathcal{D} \) and \( R : \mathcal{D} \rightarrow \mathcal{C} \) by \( L \dashv R \) and we denote the unit and counit transformations of an adjunction typically by \( \eta \) and \( \varepsilon \).
Twisting and mixing objects

"Pour moi, l’endomorphisme de Frobenius n’était pas un « alpha et oméga » pour le formalisme co-homologique, mais un endomorphisme parmi bien d’autres…"

Alexander Grothendieck

In this section $\mathcal{C}$ denotes an arbitrary category endowed with an endomorphism $F : \text{id}_\mathcal{C} \to \text{id}_\mathcal{C}$ of the identity functor, this means that for every object $X$ there is an endomorphism $F_X \in \text{End}_\mathcal{C}(X)$ such that for every arrow $f : X \to Y$ in $\mathcal{C}$, we have $F_Y \circ f = f \circ F_X$.

For the applications that we have in mind, we will always want to take for $\mathcal{C}$ the category of schemes in characteristic $p > 0$ and for $F$ the absolute Frobenius. Certainly one could think of other interesting situations, for instance by taking $\mathcal{C}$ arbitrary and $F$ trivial, or $\mathcal{C}$ an abelian category and $F$ the zero endomorphism, but then we don’t know any applications.

8.1 The twisted and mixed categories

Definition 8.1.1. The twisted category $\tilde{\mathcal{C}}$ is defined as follows. The objects are the pairs $\tilde{X} = (X, \Phi_X)$ where $X \in \text{Ob}(\mathcal{C})$ and $\Phi_X \in \text{End}_\mathcal{C}(X)$ satisfies $\Phi_X \circ \Phi_X = F_X$. The morphisms $f : \tilde{X} \to \tilde{Y}$ are those morphisms $f : X \to Y$ for which $\Phi_Y \circ f = f \circ \Phi_X$.

For a twisted object $\tilde{X} = (X, \Phi_X)$, we call $X$ the underlying ordinary object and $\Phi_X$ the twister. Note that $\tilde{\mathcal{C}}$ is itself a category with an

---

1Récoltes et Semailles, p. 881
endomorphism $\Phi$ of the identity functor, and that there is a forgetful functor $f : t\mathcal{C} \to \mathcal{C} : \tilde{X} \mapsto X$ which we call the untwisting functor.

**Definition 8.1.2.** The mixed category $m\mathcal{C}$ is defined as follows. The objects are the quadruples

$$\tilde{X} = (X_1, X_2, \Phi_{X_1}, \Phi_{X_2})$$

where $X_1, X_2 \in \text{Ob}(\mathcal{C})$ and $\Phi_{X_i} \in \text{hom}_\mathcal{C}(X_i, X_{2-i})$ satisfy $\Phi_{X_{2-i}} \circ \Phi_{X_i} = F_{X_i}$. The morphisms $f : \tilde{X} \to \tilde{Y}$ are those pairs $(f_1, f_2)$ of morphisms $f_i : X_i \to Y_i$ for which $\Phi_{Y_i} \circ f_i = f_{2-i} \circ \Phi_{X_i}$.

We will depict a morphism of mixed objects diagrammatically as

$$
\begin{array}{ccc}
\tilde{X} & \xrightarrow{\Phi_{X_1}} & X_1 \\
\downarrow f & & \downarrow \Phi_{X_2} \\
\tilde{Y} & \xrightarrow{\Phi_{Y_1}} & Y_1 \\
\end{array}
\quad \text{or} \quad
\begin{array}{ccc}
X_1 & \xleftarrow{\Phi_{X_2}} & \tilde{X} \\
\downarrow f_1 & & \downarrow \Phi_{X_1} \\
Y_1 & \xleftarrow{\Phi_{Y_2}} & \tilde{Y} \\
\end{array}
\quad f_2
$$

This is not a commutative diagram since the pair of arrows $\bullet \xmapsto{\Phi_{X_1}} \circ \Phi_{Y_1} \neq \Phi_{X_2}$ does not compose to the identity but rather to $F_{\bullet}$ and $F_\circ$; one should think of it as an abbreviation for the bigger (commutative) diagram

$$
\begin{array}{ccc}
X_1 & \xrightarrow{\Phi_{X_1}} & X_2 \\
\downarrow f_1 & & \downarrow \Phi_{X_2} \\
Y_1 & \xrightarrow{\Phi_{Y_1}} & Y_1 \\
\end{array}
\quad \text{or} \quad
\begin{array}{ccc}
\tilde{X} & \xleftarrow{\Phi_{X_2}} & \tilde{Y} \\
\downarrow f_2 & & \downarrow \Phi_{X_1} \\
\tilde{Y} & \xleftarrow{\Phi_{Y_2}} & \tilde{X} \\
\end{array}
\quad f_1
$$

The maps $\Phi_{X_1}$ and $\Phi_{X_2}$ are called the mixing maps or mixers. If they are clear from the context, we will also denote $\tilde{X}$ simply by $(X_1, X_2)$.

To make constructions, it is important that some of the good properties of $\mathcal{C}$ are carried over to $m\mathcal{C}$ and $t\mathcal{C}$. For instance, if $\mathcal{C}$ admits fibered products $X \times_S Y$ or coproducts $X \sqcup Y$, we would like the same thing to be true for $t\mathcal{C}$. The following lemma reassures us that this is always the case.
Lemma 8.1.3. If \( \mathcal{C} \) admits (co)limits for diagrams of shape \( \mathcal{J} \), then so do \( t\mathcal{C} \) and \( m\mathcal{C} \).

Proof. Consider a diagram \( \mathbf{D} : \mathcal{J} \to t\mathcal{C} \), set \( \mathbf{D}' = \mathbf{f} \circ \mathbf{D} \) and let \( X = \lim \mathbf{D}' \) together with the morphisms \( \chi_U : X \to \mathbf{D}'(U) \) for every \( U \in \text{Ob}(\mathcal{J}) \). Then the object \( X \) together with the morphisms \( \mathbf{f}(\Phi_{\mathbf{D}(U)}) \circ \chi_U \) forms a cone, i.e. a collection of morphisms from \( X \) to all the object in the diagram such that all paths together with arrows in the diagram commute. The universal property of a limit will associate to this cone a unique morphism \( \Phi_X : X \to X \). Moreover, it is clear that the morphism \( \Phi_X \circ \Phi_X \) is the unique morphism which makes the cone coming from all \( \Phi_{\mathbf{D}(U)}^2 \) commute, but the morphism \( F_X \) has this property as well so \( \Phi_X \circ \Phi_X = F_X \). It is then immediately verified that \((X, \Phi_X)\) is a limit for \( \mathbf{D} \). An analogous argument holds for colimits. The proof for \( m\mathcal{C} \) is similar but using each of the functors \((X_1, X_2, \Phi_{X_1}, \Phi_{X_2}) \rightsquigarrow X_i\) to construct an appropriate object. (We omit the proof because we will only use this in a context where it follows directly from the case of \( t\mathcal{C} \), thanks to Proposition 8.3.2.)

In particular, if \( \mathcal{C} \) has a terminal object \( 1 \), then \( 1_{t\mathcal{C}} = (1, \text{id}_1) \) is a terminal object for \( t\mathcal{C} \) and \( 1_{m\mathcal{C}} = (1, 1, \text{id}_1, \text{id}_1) \) is a terminal object for \( m\mathcal{C} \).

Throughout the rest of this chapter, we will always assume that the conditions \((\dagger)\) hold:

\[
(\dagger) \left\{ \begin{array}{ll}
\mathcal{C} \text{ has a terminal object } 1_{\mathcal{C}} \\
\mathcal{C} \text{ admits fibered products } X \times_S Y \\
\mathcal{C} \text{ admits binary coproducts } X \sqcup Y
\end{array} \right.
\]
8.2 Many functors

Definition 8.2.1. We now define a number of functors between the categories $C$, $tC$ and $mC$. The following overview will be helpful:

(i) We already introduced the untwisting functor, which forgets the twister:

$$f : tC \rightarrow C : (X, \Phi_X) \rightsquigarrow X.$$  

(ii) We have two functors $\delta_! : mC \rightarrow tC$ and $\delta_* : mC \rightarrow tC$ called the left and right misting functors (mixed to twisted). They are defined by

$$\delta_! : (X_1, X_2, \Phi_1, \Phi_2) \rightsquigarrow (X_1 \uplus X_2, \Phi_1 \sqcup \Phi_2),$$
$$\delta_* : (X_1, X_2, \Phi_1, \Phi_2) \rightsquigarrow (X_1 \times X_2, \Phi_1 \times \Phi_2).$$

(iii) The twixing functor (twisted to mixed) $\delta^* : tC \rightarrow mC$ and the twisting functor $\tau^* : mC \rightarrow mC$ are given by

$$\delta^* : (X, \Phi) \rightsquigarrow (X, X, \Phi, \Phi),$$
$$\tau^* : (X_1, X_2, \Phi_1, \Phi_2) \rightsquigarrow (X_2, X_1, \Phi_2, \Phi_1).$$

(iv) We define the component functors by

$$c_i : mC \rightarrow C : (X_1, X_2, \Phi_{X_1}, \Phi_{X_2}) \rightsquigarrow X_i.$$  

(v) Finally, we define the mixing and anti-mixing functors by

$$m : C \rightarrow mC : X \rightsquigarrow (X, X, F_X, \text{id}_X)$$
$$\overline{m} : C \rightarrow mC : X \rightsquigarrow (X, X, \text{id}_X, F_X).$$

Proposition 8.2.2. We have the following properties:

(i) There are adjunctions $c_1 \dashv m \dashv c_2 \dashv \overline{m} \dashv c_1$.

(ii) Each of the functors $c_1$, $c_2$, $m$ and $\overline{m}$ preserves all limits and colimits (hence in particular products, coproducts and terminal objects).
(iii) The functors \( \mathfrak{m} \) and \( \mathfrak{m} \) are full and faithful.

(iv) There is an adjunction \( \delta_1 \dashv \delta^* \dashv \delta_2 \).

(v) \( \tau^* \) is an equivalence and \( \tau^* \circ \tau^* \simeq \text{id}_{\mathcal{C}} \).

(vi) \( \tau^* \circ \mathfrak{m} = \mathfrak{m} \); \( c_i = c_{2-i} \circ \tau^* \); \( \delta^* = \tau^* \circ \delta^* \).

(vii) \( f = c_2 \circ \delta^* = c_1 \circ \delta^* \).

\[ \text{Proof.} \quad (i) \] We will only verify that \( c_1 \dashv \mathfrak{m} \); the other pairs are completely analogous. Consider objects \( \tilde{X} = (X_1, X_2, \Phi_{X_1}, \Phi_{X_2}) \in \text{Ob}(\mathfrak{mC}) \), \( Y \in \text{Ob}(\mathcal{C}) \) and the map

\[ \text{hom}_{\mathfrak{mC}}(\tilde{X}, \mathfrak{m}(Y)) \to \text{hom}_{\mathcal{C}}(c_1(X), Y) : (\alpha, \beta) \mapsto \alpha. \]

Clearly the map \( \alpha \mapsto (\alpha, \alpha \circ \Phi_{X_2}) \) is an inverse and these bijections are natural in \( \tilde{X} \) and \( Y \).

(ii) Since each of the functors \( c_1, \mathfrak{m}, c_2 \) and \( \mathfrak{m} \) is now both a left and right adjoint, they must preserve all limits, colimits, epimorphisms and monomorphisms.

(iii) Clearly the counit \( Y \to (c_1 \circ \mathfrak{m})(Y) = Y \) is the identity, which implies that \( \mathfrak{m} \) is full and faithful by [Mac71, (IV.3.1)]. A similar argument holds for \( \mathfrak{m} \).

The remaining statements are obvious from the definitions.

\[ \square \]

**Definition 8.2.3.** Since \( \mathfrak{m} \) (resp. \( \mathfrak{m} \)) is fully faithful, its essential image is equivalent to \( \mathcal{C} \), so we call the mixed objects isomorphic to \( \mathfrak{m}(X) \) (resp. \( \mathfrak{m}(X) \)) for some \( X \in \text{Ob}(\mathcal{C}) \) visible (resp. anti-visible). The mixed objects that are not visible are called invisible. In what follows we will occasionally identify an object \( X \in \text{Ob}(\mathcal{C}) \) with the corresponding visible object \( \mathfrak{m}(X) \in \text{Ob}(\mathfrak{mC}) \). In other words, we consider \( \mathcal{C} \) as a full subcategory of \( \mathfrak{mC} \) through \( \mathfrak{m} \).

**Remark 8.2.4.** The verification of the following observations is straightforward and left to the reader.

(i) A mixed object \((X_1, X_2, \Phi_1, \Phi_2)\) is visible (resp. anti-visible) if and only if the map \( \Phi_2 \) (resp. \( \Phi_1 \)) is an isomorphism.

(ii) If \( X \in \text{Ob}(\mathfrak{mC}) \) is visible then we have the following bijection, natural in \( X \) and \( \tilde{Y} \):

\[ \text{hom}(X, \tilde{Y}) \to \text{hom}(X, c_2\tilde{Y}) : (f_1, f_2) \mapsto f_2. \]
Note that the hom-set \( \text{hom}(X, c_2 \tilde{Y}) \) can be interpreted in either \( \mathcal{C} \) or \( m\mathcal{C} \), but in the latter case the map is given by \((f_1, f_2) \mapsto (f_2, f_2) \) instead.

(iii) If \( \tilde{X} = (X_1, X_2) \) and \( \tilde{Y} = (Y_1, Y_2) \) are mixed objects, then there is a fibered product in (\text{set})

\[
\text{hom}(\tilde{X}, \tilde{Y}) = \text{hom}(X_1, Y_1) \times_{\text{hom}(X_1, Y_2) \times \text{hom}(X_2, Y_1)} \text{hom}(X_2, Y_2),
\]

where the maps are \( u \mapsto (\Phi_{Y_1} \circ u, u \circ \Phi_{X_2}) \) and \( v \mapsto (v \circ \Phi_{X_1}, \Phi_{Y_2} \circ v) \).

(iv) The projection of this fibered product to its first component corresponds to the map \( c_1 : \text{hom}(\tilde{X}, \tilde{Y}) \to \text{hom}(X_1, Y_1) : (f_1, f_2) \mapsto f_1 \).

If \( F_{X_2} \) is epic or \( F_{Y_2} \) is monic, the reader may verify that this map is injective. We will postpone a much more precise version of this statement until Proposition 9.2.3.

We conclude this section with the following proposition, which claims good behaviour of all these constructions under functors that are sufficiently nice.

**Proposition 8.2.5.** The formation of \( t\mathcal{C} \) and \( m\mathcal{C} \) commutes with op: \((t\mathcal{C})^{\text{op}} = t\mathcal{C}^{\text{op}} \) and \((m\mathcal{C})^{\text{op}} = m\mathcal{C}^{\text{op}} \). Furthermore, if \( G : \mathcal{D} \to \mathcal{C} \) is a functor between categories with endomorphisms of the identity functor, both denoted by \( F \) and \( H \), such that \( G(F_X) = H_G(X) \) for all \( X \in \text{Ob}(\mathcal{D}) \), then there are functors \( tG : t\mathcal{D} \to t\mathcal{C} \) and \( mG : m\mathcal{D} \to m\mathcal{C} \) which commute with all the functors defined in Definition 8.2.1, i.e. with \( f, c_1, c_2, m, \delta, \delta^*, \delta^*_s \) and \( \tau^* \). For instance, the diagram

\[
\begin{array}{ccc}
m\mathcal{D} & \xrightarrow{\delta} & t\mathcal{D} & \xrightarrow{f} & \mathcal{D} \\
\downarrow mG & & \downarrow tG & & \downarrow G \\
m\mathcal{C} & \xrightarrow{\delta} & t\mathcal{C} & \xrightarrow{f} & \mathcal{C}
\end{array}
\]

commutes.

**Proof.** This is immediately clear from the definitions. \( \square \)
Note however, that taking $\text{op}$ will exchange the roles of the functors $\delta_!$ and $\delta^*$, and also the functors $\text{m}$ and $\text{M}$.

More generally, this will be the case for any contravariant functor, such as $\text{Spec} : (\text{ring}) \to (\text{sch})$.

8.3 Twisting versus mixing

The categories $\text{mC}$ and $\text{tC}$ are closely related: we will now observe that under a mild assumption $\text{mC}$ is a slice category of $\text{tC}$. Let us first briefly recall what this means:

**Definition 8.3.1.** If $\mathcal{C}$ is an arbitrary category, and $S \in \text{Ob}(\mathcal{C})$ an object, then the *slice category* $\mathcal{C}/S$ is the category which has as objects the arrows $q_X : X \to S$ in $\mathcal{C}$, and as morphisms between $q_X : X \to S$ and $q_Y : Y \to S$ those arrows $f : X \to Y$ in $\mathcal{C}$ such that the corresponding diagram commutes.

\[
\begin{array}{ccc}
X & \xrightarrow{f} & Y \\
\downarrow{q_X} & & \downarrow{q_Y} \\
S
\end{array}
\]

A frequent abuse of notation is to denote $q_X$ simply by $X$; but sometimes this identification can be dangerous—for instance, it is ambiguous for the arrow $F_S : S \to S$. With this notation in place, what we will show is that (under suitable conditions) the categories $\text{mC}$ and $\text{tC}/E$ are equivalent, where $E$ is a particular twisted object that we will first introduce.

Recall from the conditions $(\dagger)$ stated at the end of Section 8.1 that we assume that $\mathcal{C}$ has a terminal object $1 = 1_{\mathcal{C}}$. Then $\text{mC}$ has a terminal object $1_{\text{mC}} = \text{m}(1)$, which is given by $(1, 1)$ with the only possible choice $(\text{id}_1, \text{id}_1)$ for the mixers. If we denote $E = \delta!(1_{\text{mC}}) = \delta!\text{m}1$ then this is *not* a terminal object in $\text{tC}$. (This also shows that $\delta!$ cannot have a left adjoint.) In fact $f(E) = 2 = 1 \sqcup 1$ and the twister $\tau$ swaps both components.
Chapter 8. Twisting and mixing objects

Every object $X \in \text{Ob}(m\mathcal{C})$, together with its unique morphism $q_X : X \to 1_{m\mathcal{C}}$, now gives rise to an arrow $\delta q_X : \delta_i(X) \to E$ in $m\mathcal{C}$. So we obtain a functor:

$$Q : m\mathcal{C} \to t\mathcal{C}/E : X \rightsquigarrow \left( \begin{array}{c} \delta_iX \\ \downarrow \\ E \end{array} \right)$$

We will show that under a mild condition, this functor is an equivalence. Our assumption on $\mathcal{C}$ is that in addition to the conditions ($\dagger$) the natural functor

$$\mathcal{C} \times \mathcal{C} \to \mathcal{C}/2 : (X,Y) \to (X \sqcup Y \to 2)$$

is an equivalence. In this case all the natural functors

$$\mathcal{C}/A \times \mathcal{C}/B \to \mathcal{C}/A \sqcup B$$

are equivalences [CLW93, Proposition 4.1] and the category is said to be extensive. We will henceforth denote this assumption by ($\ddagger$).

**Proposition 8.3.2.** If the condition ($\ddagger$) holds, then $Q : m\mathcal{C} \to t\mathcal{C}/E$ is an equivalence of categories. Consequently, for every $\tilde{X} \in \text{Ob}(m\mathcal{C})$, there is an equivalence $m\mathcal{C}/\tilde{X} \simeq t\mathcal{C}/Q(\tilde{X})$.

Essentially we need to construct an inverse to the functor $Q$. One way of writing this down makes use of the arrow category $\text{Ar}(\mathcal{C})$ of an arbitrary category $\mathcal{C}$, sometimes also denoted by $\tilde{\mathcal{C}}$. This category has as objects the arrows in $\mathcal{C}$ and as morphisms the commuting squares.

**Proof.** We first claim that under the condition ($\ddagger$) there is a cartesian square in the category of categories given by

$$\begin{array}{ccc}
m\mathcal{C} & \xrightarrow{Q} & t\mathcal{C}/E \\
\downarrow & & \downarrow \\
\text{Ar}(\mathcal{C}) \times \text{Ar}(\mathcal{C}) & \longrightarrow & \text{Ar}(\mathcal{C}/2),
\end{array}$$
where the right vertical arrow is given by sending the object $\eta_Z : (Z, \Phi_Z) \to (2, \tau)$ to

$$
\begin{array}{ccc}
Z & \xrightarrow{\Phi_Z} & Z \\
\tau \circ \eta_Z & \searrow & \eta_Z \\
\downarrow & & \downarrow \\
2 & & 2
\end{array}
$$

Every object in the fibered product under consideration is therefore a tuple

$$(f : X \to X', g : Y \to Y'; \eta_Z : Z \to 2, \Phi_Z : Z \to Z)$$

such that $\Phi_Z^2 = F_Z$ and the following diagrams are equal

$$
\begin{array}{ccc}
X \sqcup X' & \xrightarrow{f \sqcup g} & Y \sqcup Y' \\
\downarrow \tau \circ (q_X \sqcup q_X') & & \tau \circ (q_Y \sqcup q_Y') \\
2 & \xrightarrow{\tau} & 2
\end{array} \equiv
\begin{array}{ccc}
Z & \xrightarrow{\Phi_Z} & Z \\
\eta_Z & \downarrow & \eta_Z \\
2 & \xrightarrow{\tau} & 2
\end{array}
$$

By the condition (‡), we can conclude that $Y' \cong X$, $X' \cong Y$ and the compositions $f \circ g$ and $g \circ f$ are equal to $F_X$ and $F_Y$. Thus such a tuple can be identified with an object of $m\mathcal{C}$, unique up to isomorphism. An arrow between an object as described above, and a different object

$$(\tilde{f} : \tilde{X} \to \tilde{X}', \tilde{g} : \tilde{Y} \to \tilde{Y}'; \eta_{\tilde{Z}} : \tilde{Z} \to 2, \Phi_{\tilde{Z}} : \tilde{Z} \to \tilde{Z}),$$

where we may also assume that $\tilde{X}' = \tilde{Y}$, $\tilde{Y}' = \tilde{X}$, $\tilde{Z} = \tilde{X} \sqcup \tilde{Y}$ and $\Phi_{\tilde{Z}} = \tilde{f} \sqcup \tilde{g}$, boils down to a collection of arrows $X \to \tilde{X}$, $X' \to \tilde{X}'$, $\ldots$, such that the corresponding squares commute. Therefore these are exactly the arrows in the category $m\mathcal{C}$. Conversely every arrow in $m\mathcal{C}$ gives rise to a collection of such arrows between the corresponding objects. This shows that the fibered product is isomorphic to $m\mathcal{C}$, in other words the given diagram is indeed a cartesian square.

Furthermore, under the condition (‡), the bottom arrow in this cartesian diagram is an isomorphism, and therefore the arrow $Q$ is an isomorphism as well.

For the final statement: if we take an object $\tilde{X}$ in $m\mathcal{C}$ and the corresponding object $Q(\tilde{X})$ in $t\mathcal{C}$ then the equivalence $Q$ will induce an
equivalence on the slice categories, and this implies the last statement.

\[ \square \]

**Remark 8.3.3.** Proposition 8.3.2 justifies the notations for some of the functors we introduced in Section 8.2:

- The twixing functor \( \delta^* \) and the twisting functor \( \tau^* \) can be interpreted as base changes in \( t\mathcal{C} \) along the arrows
  \[
  \delta : E \to 1 \quad \text{and} \quad \tau : E \to E,
  \]
  i.e. \( \delta^*(X) = X \times E \) and \( \tau^*(X) = X \times_{q \times \tau} E \), hence the notation. Note however that \( \tau^* \circ \tau^* \) is the identity functor on \( m\mathcal{C} \), whereas it is merely isomorphic to it when considered on \( (t\mathcal{C})/E \).
- The functor \( \delta_! \) can be interpreted as the functor \( t\mathcal{C}/E \to t\mathcal{C} \) which *forgets* the structural morphism.

**Remark 8.3.4.**

(i) It seems unlikely that one could construct \( m\mathcal{C} \) directly out of \( t\mathcal{C} \) under a much weaker condition than (†): a priori \( t\mathcal{C} \) could contain far less information than \( \mathcal{C} \), or even degenerate into the empty category, whereas \( m\mathcal{C} \) will always contain \( \mathcal{C} \) as a full subcategory. So the condition must in some sense guarantee that \( t\mathcal{C} \) is sufficiently large.

(ii) The practical importance of this proposition is that in Chapter 9, we will have:

\[
\text{“mixed ring” = “twisted algebra over } \mathbb{F}_p \text{”,}
\]
\[
\text{“mixed scheme” = “twisted scheme over Spec } \mathbb{F}_p \text{”}
\]

We may now study the twixing functor \( \delta^* : t\mathcal{C} \to m\mathcal{C} \) in greater detail and characterize its essential image. Considered as a functor of base change \( \delta^* : (t\mathcal{C})/1 \to (t\mathcal{C})/E \), this is an instance of a descent problem—although our characterization will not depend on this. The underlying philosophy is that the category \( m\mathcal{C} \) is much better behaved than \( t\mathcal{C} \) thanks to the (anti)-mixing and component functors, so we prefer to perform constructions in \( m\mathcal{C} \) and understand how they descend to \( t\mathcal{C} \).
Definition 8.3.5. A (twisted) descent datum on $\tilde{X} \in \text{Ob}(m\mathcal{C})$ is a morphism $f: \tilde{X} \to \tau^*\tilde{X}$ such that $\tau^*f \circ f = \text{id}_{\tilde{X}}$. We form a category $m\mathcal{C}[\text{tdd}]$ where objects are pairs $(\tilde{X}, f)$ consisting of an object $\tilde{X}$ of $m\mathcal{C}$ together with a descent datum $f$ on $\tilde{X}$; an $m\mathcal{C}[\text{tdd}]$-arrow $u: (\tilde{X}, f) \to (\tilde{Y}, g)$ is an $m\mathcal{C}$-arrow $u: \tilde{X} \to \tilde{Y}$ such that $g \circ u = \tau^*u \circ f$:

Since applying $\tau^*$ to $\tau^*f \circ f = \text{id}_{\tilde{X}}$ yields $f \circ \tau^*f = \text{id}_{\tau^*\tilde{X}}$, we see that a descent datum $f$ is always an isomorphism, with $\tau^*f$ as its inverse.

Proposition 8.3.6. The functor $\delta^*$ factors as

$$t\mathcal{C} \xrightarrow{\alpha} m\mathcal{C}[\text{tdd}] \xrightarrow{\text{forget}} m\mathcal{C}$$

where $\alpha$ is an equivalence and $\text{forget} : (\tilde{X}, f) \rightsquigarrow \tilde{X}$ forgets the descent datum.

Proof. Since $\delta^* = \tau^*\delta^*$, the identity is a map $\text{id}: \delta^*\tilde{X} \to \tau^*\delta^*\tilde{X}$, for every $\tilde{X} \in \text{Ob}(t\mathcal{C})$ which is trivially a descent datum on $\delta^*\tilde{X}$. So we define

$$\alpha : t\mathcal{C} \to m\mathcal{C} : \tilde{X} \rightsquigarrow (\delta^*\tilde{X}, \text{id}),$$

and it is clear that composing $\alpha$ with the functor forgetting the descent datum is indeed $\delta^*$. From this it is also clear that $\alpha$ is faithful.

We now show that $\alpha$ is essentially surjective. Consider an arbitrary object $(\tilde{X}, f) \in \text{Ob}(m\mathcal{C}[\text{tdd}])$, and write $\tilde{X} = (X_1, X_2, \Phi_1, \Phi_2)$ and $f = (f_1, f_2)$. Then the diagram

$$\begin{array}{ccc}
X_1 & \xleftarrow{\Phi_1} & X_2 \\
\downarrow{f_1} & & \downarrow{f_2} \\
X_2 & \xleftarrow{\Phi_2} & X_1
\end{array}$$
commutes, and moreover $f_1 = f_2^{-1}$. It follows that also the diagram

$$
\begin{array}{ccc}
X_1 & \xrightarrow{\Phi_1} & X_2 \\
\downarrow f_1 & & \downarrow \text{id} \\
X_2 & \xleftarrow{\Phi_1 \circ f_2} & X_2
\end{array}
$$

commutes. Hence there is an isomorphism $(f_1, \text{id}) : \tilde{X} \to \delta^*(X_2, \Phi_1 \circ f_2)$ which respects the descent data $(f_1, f_2)$ and $(\text{id}, \text{id})$, so it determines an isomorphism $(\tilde{X}, f) \cong \alpha(X_2, \Phi_1 \circ f_2)$ in $m\mathcal{C}[\text{tdd}]$. Thus $\alpha$ is essentially surjective.

Finally, we show that $\alpha$ is full. Let $u : (\tilde{X}, f) \to (\tilde{Y}, g)$ be a morphism in $m\mathcal{C}[\text{tdd}]$. Since $\alpha$ is essentially surjective, we may as well assume that $\tilde{X} = \delta^*(X, \Phi)$ and $\tilde{Y} = \delta^*(Y, \Psi)$ with in both cases the identity as descent datum. The morphism $u = (u_1, u_2)$ respects the descent data, which boils down to $u_2 = u_1$ and then $u = \alpha(u_1)$ so $\alpha$ is full. 

The following corollary is particularly useful to remember.

**Corollary 8.3.7.** A mixed object $\tilde{X}$ descends to a twisted object if and only if it has an endomorphism $f : \tilde{X} \to \tilde{X}$ such that $\tau^* f \circ f = \text{id}_{\tilde{X}}$. In particular, it is necessary that $c_1(\tilde{X}) \cong c_2(\tilde{X})$.

As an example of how this is useful, let us mention that later in this work we will construct what we call mixed algebraic groups $\tilde{X}$ of type $(B_n, C_n)$. The components $c_1(\tilde{X})$ and $c_2(\tilde{X})$ are isomorphic only when $n = 2$, so the group can only admit twisted descent in this case. It turns out that when $n = 2$, the group indeed admits twisted descent and this produces the Suzuki groups $^2B_2$.

### 8.4 Categories of presheaves

The category of presheaves on $\mathcal{C}$ is denoted by $\widehat{\mathcal{C}} = \text{hom}(\mathcal{C}^{\text{op}}, (\text{set}))$. We can canonically endow it with an endomorphism of the identity functor $\widehat{F} : \text{id}_{\widehat{\mathcal{C}}} \to \text{id}_{\widehat{\mathcal{C}}}$ by defining for an arbitrary presheaf $G \in \text{Ob}(\widehat{\mathcal{C}})$ the endomorphism $\widehat{F}_G : G \to G$ as the natural transformation

$$
\begin{array}{ccc}
G & \xrightarrow{\text{id}} & G \\
\downarrow \text{id} & & \downarrow \text{id} \\
G & \xrightarrow{\text{id}} & G
\end{array}
$$
with components \((\tilde{F}_G)_X = G(F_X)\). The pair \(F\) and \(\tilde{F}\) behaves well with respect to the Yoneda embedding \(Y_\phi : \mathcal{C} \to \hat{\mathcal{C}} : X \mapsto h_X\) in the sense of Proposition 8.2.5, so we have the following corollary of that proposition:

**Corollary 8.4.1.** Formation of \(t\mathcal{C}\) and \(m\mathcal{C}\) commutes with the Yoneda embedding.

**Remark 8.4.2.** To the functor \(m : \mathcal{C} \to m\mathcal{C}\) corresponds a functor \(m^*\), given by

\[
m^* : m\mathcal{C} \to \mathcal{C} : F \mapsto F \circ m.
\]

We call \(m^*F\) the *mixtor restriction* (mixed to ordinary) of the presheaf \(F\). If \(F\) is represented by \(\tilde{X}\), then \(m^*F\) is represented by \(c_2 \tilde{X}\).

This implies the following statement, which is of fundamental importance for understanding mixed objects: *if \(\tilde{X}\) is a mixed object, but we are only willing to probe it—in the sense of computing \(h_{\tilde{X}}(-)\)—on visible objects, then we can only observe its second component.*

The following proposition answers a natural question, although it will play no role in what follows.

**Proposition 8.4.3.** If \(\mathcal{C}\) is small, there is an equivalence \(t\hat{\mathcal{C}} \simeq t\mathcal{C}\).

**Proof.** By Proposition 8.2.5 there is a functor \(tY_\phi : t\mathcal{C} \to t\hat{\mathcal{C}} : (X, \Phi) \mapsto (h_X, h_\Phi)\). We may construct a functor \(G : t\hat{\mathcal{C}} \to t\mathcal{C}\) such that its composition with \(tY_\phi\) is the Yoneda embedding for \(t\mathcal{C}\):

\[
Y_{t\phi} : t\mathcal{C} \xrightarrow{tY_\phi} t\hat{\mathcal{C}} \xrightarrow{G} t\mathcal{C}.
\]

Indeed, we can define \(G : \tilde{X} = (X, \Phi) \mapsto G(\tilde{X})\) by

\[
G(X, \Phi) : (Y, \Psi) \mapsto eq\left( X(Y) \xrightarrow{\Phi_Y} X(Y) \right),
\]

where eq denotes the equalizer in the category of sets, i.e.,

\[
G(X, \Phi)(Y, \Psi) = \{ u \in X(Y) \mid \Phi_Y(u) = X(\Psi)(u) \}.
\]

In particular, if \(G(X, \Phi) = (h_X, h_\Phi)\) for some \((X, \Phi) \in \text{Ob}(t\mathcal{C})\), we get

\[
G(h_X, h_\Phi)(Y, \Psi) = \{ u \in h_X(Y) \mid (h_\Phi)_Y(u) = h_X(\Psi)(u) \}
\]
\[ \{ u \in h_X(Y) \mid \Phi \circ u = u \circ \Psi \} = h_{(X,\Phi)}(Y,\Psi), \]

from which it follows that \( \mathbf{G} \circ tY \mathcal{C} \mathcal{C} = Y t \mathcal{C} \). On the other hand, we know that \( \mathcal{C} \) is a cocomplete category—i.e. all small colimits exist, and therefore \( t\mathcal{C} \) is a cocomplete category by Lemma 8.1.3. Thus by the universal property of \( t\mathcal{C} \) as the free cocompletion of \( \mathcal{C} \), the functor \( tY \mathcal{C} \) extends uniquely to a cocontinuous functor \( \mathbf{F} : t\mathcal{C} \to t\mathcal{C} \) ([MM94, p. 43 Cor. 4]). The pair \( \mathbf{F}, \mathbf{G} \) provides the equivalence of categories.

\[ \square \]

### 8.5 Mixed objects over a visible base object

In this section, we discuss a first way in which mixed objects can appear in the ordinary world. The simplest occasion is when one encounters an absolute factorization, i.e. a factorization of \( F_X : X \to X \) through another object \( Y \) such that the other composition is \( F_Y \).

However, absolute factorizations rarely occur in practice, because often there is a base object \( S \) and one prefers to work relatively with respect to \( S \). In such situations the entire reasoning takes place in the category \( \mathcal{C}/S \) where the only admitted arrows between two \( S \)-objects are the \( S \)-linear ones; but if \( X \) is an \( S \)-object, then \( F_X \) is not expected to be \( S \)-linear unless \( F_S = \text{id}_S \). The role of the absolute factorizations is then played by relative factorizations which we proceed to introduce. The upshot of it all will be that mixed objects over visible base objects \( m(S) = (S,S,F_S,\text{id}_S) \) are still easily described in terms of linear arrows.

First, recall that for a morphism \( f : T \to S \) in \( \mathcal{C} \) and objects \( X \in \text{Ob}(\mathcal{C}/S) \) and \( Y \in \text{Ob}(\mathcal{C}/T) \), we have a natural identification between the following sets:

\[ \text{hom}_f(Y, X) = \{ g \in \text{hom}_\mathcal{C}(Y, X) \mid q_Y \circ g = f \circ q_X \} \]
\[ \simeq \{ h \in \text{hom}_\mathcal{C}(Y, X \times T) \mid q_Y = p_2 \circ h \} \]
\[ \simeq \text{hom}_{\mathcal{C}/T}(Y,f^* X), \]
where \( f^*X = X \times_{q_X,f} T \) is called the \textit{pullback} of \( X \) from \( \mathcal{C}/S \) to \( \mathcal{C}/T \), with the projection to \( T \) as its structural morphism.

Let us now consider an object \( X \in \text{Ob}(\mathcal{C}) \) together with a structural morphism \( q_X : X \to S \). Then the arrow \( F_X : X \to X \) is an element of the set \( \text{hom}_{\mathcal{C}/S}(X,X) \). Following the above isomorphism of hom-sets, we obtain a relative version of this arrow in \( \text{hom}_{\mathcal{C}/S}(X,F_S^*X) \), for which we introduce the notation

\[
F_{X/S} : X \to F_S^*X.
\]

Let us also introduce the notation \( \Delta X = F_S^*X \), which should stress the fact that we want to consider this object as an object over \( S \) via the projection \( p_1 : F_S^* = S \times_{F_S,q_X} X \to S \). Since this projection \( p_1 \) must be thought of as the structural morphism of the object \( \Delta X \) a more appropriate notation for it is \( q_{\Delta X} : \Delta X \to S \).

The arrow \( F_{X/S} : X \to \Delta X \) is by definition the unique arrow for which the following diagram commutes.

\[
\begin{array}{ccc}
X & \xrightarrow{F_X} & X \\
\downarrow{q_X} & & \downarrow{q_X} \\
\Delta X & \xrightarrow{P_{X/S}} & \Delta X \\
\downarrow{q_{\Delta X}} & & \downarrow{q_{\Delta X}} \\
S & \xrightarrow{F_S} & S \\
\end{array}
\]

In other words, \( F_{X/S} \) is completely determined by the following properties:

- The composition with \( q_{\Delta X} \) is the structural morphism \( q_X \) on \( X \), in other words \( F_{X/S} \) is \( S \)-linear.
- The composition with the canonical projection \( P_{X/S} : \Delta X \to X \) is equal to \( F_X \).

We will sometimes write this as \( F_{X/S} = q_X \times F_X \). Before we continue, let us formalize the good properties of \( F_{S/S} \):

\[\footnote{Our main excuse for introducing these notations is that we want \( F_S^* \) to be mentally processed in one step and not via the chain of thoughts \( S \to F_S \to F_S^* \). It also allows to suppress the base \( S \) in the notation, which is often convenient.}\]
Proposition 8.5.1. The arrows $F_{X/S}, X \in \text{Ob}(\mathcal{C}/S)$, form the components of a natural transformation

![Diagram]

\[ \begin{array}{ccc}
\mathcal{C}/S & \xrightarrow{id} & \mathcal{C}/S \\
\downarrow{\Delta} & & \downarrow{\Delta} \\
\mathcal{C}/S & & \mathcal{C}/S.
\end{array} \]

Proof. We must show that for an arbitrary morphism $f : X \to Y$ the following diagram commutes

\[
\begin{array}{ccc}
X & \xrightarrow{f} & Y \\
\downarrow{F_{X/S}} & & \downarrow{F_{Y/S}} \\
\Delta X & \xrightarrow{\Delta f} & \Delta Y.
\end{array}
\]

Equivalently, we must show equality of both paths, after composing them with the canonical projections $p_i : Y \times_{q_Y,F_S} S$, which we denote by $P_{Y/S} : \Delta Y \to Y$ and $q_{\Delta Y} : \Delta Y \to S$. But we have the diagrams

\[
\begin{array}{ccc}
X & \xrightarrow{f} & Y \\
\downarrow{F_{X/S}} & \downarrow{F_{Y/S}} & \downarrow{F_{X/S}} \\
\Delta X & \xrightarrow{\Delta Y} & \Delta Y \\
\downarrow{P_{X/S}} & \downarrow{P_{Y/S}} & \downarrow{q_{\Delta Y}} \\
X & \xrightarrow{f} & Y \\
\downarrow{S} & \downarrow{id_S} & \downarrow{S} \\
S & \xrightarrow{id_S} & S,
\end{array}
\]

which commute because $F$ is a natural transformation and $f$ is $S$-linear, and also the defining diagram for $\Delta f$

\[
\begin{array}{ccc}
\Delta X & \xrightarrow{P_{X/S}} & X \\
\downarrow{q_{\Delta X}} & \downarrow{\Delta f} & \downarrow{f} \\
\Delta Y & \xrightarrow{P_{Y/S}} & Y \\
\downarrow{q_{\Delta Y}} & \downarrow{q_Y} & \downarrow{q_Y} \\
S & \xrightarrow{id_S} & S \\
S & \xrightarrow{F_S} & S.
\end{array}
\]

From these diagrams one can immediately read off the required equalities. \qed
Let us define formally:

**Definition 8.5.2.** A *relative S-factorization* in \( \mathcal{C} \) is a diagram of \( S \)-morphisms

\[
X \overset{\pi}{\longrightarrow} \overline{X} \overset{\pi'}{\longrightarrow} \triangle X,
\]

such that \( \pi' \circ \pi = F_{X/S} \) and \( \triangle \pi \circ \pi' = F_{\overline{X}/S} \), where \( \triangle \pi : \triangle X \to \triangle \overline{X} \) is the base change of \( \pi \) along \( F_S \).

The following observation is helpful to construct examples:

**Lemma 8.5.3.** An \( S \)-morphism \( X \overset{\pi}{\longrightarrow} \overline{X} \overset{\pi'}{\longrightarrow} \triangle X \) with \( \pi' \circ \pi = F_{X/S} \) and \( \pi \) epic, is a relative factorization.

**Proof.** Since \( \pi \) is epic, it suffices to show that

\[
\triangle \pi \circ \pi' \circ \pi = F_{\overline{X}/S} \circ \pi.
\]

But then the left hand side is equal to \( \triangle \pi \circ F_{X/S} \) and we are left with an equality that is true by naturality of \( F_S \) (Proposition 8.5.1). \( \square \)

**Proposition 8.5.4.** Mixed objects over a visible base correspond to relative factorizations.

**Proof.** More specifically, for \( S \in \text{Ob}(\mathcal{C}) \), we will show that \( S \)-factorizations correspond to \( m(S) \)-objects in \( m' \mathcal{C} \).

Let us start from an \( S \)-factorization \( \pi : X \to \overline{X} \). If we glue a pullback square for \( \triangle X \) to the diagram defining the relative factorization, we obtain:

\[
\begin{array}{c c c c c c}
X & \overset{\pi}{\longrightarrow} & \overline{X} & \overset{\pi'}{\longrightarrow} & \triangle X & \overset{p_1}{\longrightarrow} & X \\
\downarrow q_X & & \downarrow q_{\overline{X}} & & \downarrow p_2 & & \downarrow q_X \\
S & \overset{id_S}{\longrightarrow} & S & \overset{id_S}{\longrightarrow} & S & \overset{F_S}{\longrightarrow} & S
\end{array}
\]

Now \( (\overline{X}, X, p_1 \circ \pi', \pi) \) will be our \( m(S) \)-object; all we must do is compute the compositions:

\[
p_1 \circ \pi' \circ \pi = p_1 \circ F_{X/S} = F_X,
\]

\[
\pi \circ p_1 \circ \pi' = p_1' \circ \triangle \pi \circ \pi' = p_1' \circ F_{\overline{X}/S} = F_{\overline{X}},
\]

where \( p_1' : \triangle \overline{X} \to \overline{X} \) is the canonical projection.
Conversely, it is clear that starting from a \( m(S) \)-object \((X, \varphi, \pi)\), one can linearize \( \varphi \), i.e. factor it uniquely through \( \Delta X \) into \( \varphi = p_1 \circ \pi' \) and verify immediately that \( \pi' \circ \pi \) and \( \Delta \pi \circ \pi' \) satisfy the defining properties of \( F_{X/S} \) and \( F_{\overline{X}/S} \), namely that they are \( S \)-linear and that composition with the projections gives \( F_X \) and \( F_{\overline{X}} \).

By combining a relative factorization with a base change \( \tilde{S} \to m(S) \) to an invisible object, we obtain a large number of mixed objects over invisible base objects \( \tilde{S} \):

**Corollary 8.5.5.** The following data in \( C \) determines an \( \tilde{S} \)-object \( \tilde{X} \): an object \( S \), a relative \( S \)-factorization \( X \xrightarrow{\pi} \overline{X} \xrightarrow{\pi'} \Delta X \), and morphisms \( S \xrightarrow{\alpha} S' \xrightarrow{\beta} S \) composing to \( F_S \) resp. \( F_{S'} \).

**Proof.** Of course \( \tilde{S} = (S, S', \alpha, \beta) \) is a mixed object and there is a morphism \( \rho : \tilde{S} \to m(c_1(\tilde{S})) = m(S) \) explicitly determined by \( \rho : (\text{id}_S, \beta) \)—this is the unit of the adjunction \( c_1 \dashv m \). On the other hand, by Proposition 8.5.4, the relative factorization corresponds to a morphism \( \tilde{Y} \to m(S) \). So the pullback \( \rho^*(\tilde{Y}) = \tilde{Y} \times_{m(S)} \tilde{S} \to \tilde{S} \) determines an \( \tilde{S} \)-object with components \((\overline{X}, \beta^*X)\). \( \square \)

**Remark 8.5.6.**

(i) One can turn the collection of relative factorizations into a category and interpret Proposition 8.5.4 as an equivalence of categories.

(ii) If \( \tilde{S} = m(S) \) is visible, every mixed \( \tilde{S} \)-object arises from Proposition 8.5.4, but it is not the case that for arbitrary \( \tilde{S} = (S, S') \) every \( \tilde{S} \)-object can be constructed as in Corollary 8.5.5! This is clear because in general one does not expect the functor of base change \( \tilde{X} \to \tilde{X} \times_{m(S)} \tilde{S} \) to be essentially surjective. In some sense the objects that do not arise this way are even less accessible.

(iii) Starting from an \( \tilde{S} \)-object \( \tilde{X} \), one may base change through the morphism \( m(c_2(S)) = (S', S', \text{id}_{S'}, F_{S'}) \to \tilde{S} \)—the counit of \( m \dashv c_2 \)—to obtain a \( m(S') \)-object and thus a relative \( S' \)-factorization. So it is true that every mixed \( \tilde{S} \)-object is a form
of a relative factorization. In particular, starting from an $S$-factorization one obtains an $S'$-factorization by base changing twice, which comes down to base changing along $\beta : S' \to S$.

(iv) We suspect that the reason why the mixed quadrangles of type $F_4$ are so peculiar is because they are such inaccessible objects which do not arise from a base change. See Section A.4 for some additional discussion.

## 8.6 The concept of a fairy

This section, which is independent of the previous sections, consists of preparations for the next section, and in particular Proposition 8.7.1, which is itself the crucial ingredient for our Theorem 10.5.1. The main proposition in this section is Proposition 8.6.3 but we need to make a few definitions before we can even state it.

Let us first recall some basic facts about an arbitrary functor $u : \mathcal{C} \to \mathcal{D}$. The Yoneda extension of $u$ is a functor

$$u^! : \mathcal{C} \to \mathcal{D} : h_X \mapsto h_{u(X)},$$

defined here on representable objects and extended to arbitrary presheaves by taking limits, as explained in [MM94, Cor. 4 p. 44]. Furthermore, composition with $u$ defines a functor in the opposite direction

$$u^* : \mathcal{D} \to \mathcal{C} : F \mapsto F \circ u$$

which is a right adjoint: $u^! \dashv u^*$. Since $u^!$ extends $u$, one expects that $u^*$ extends a right adjoint to $u$, whenever it exists.

To make this last claim more precise, we must consider a representable presheaf $h_X$. Then representability of $h_X \circ u$ implies that there exists an isomorphism $h_X \circ u \simeq h_U$ of functors, for a certain object $U \in \text{Ob}(\mathcal{C})$. In other words, we obtain a collection of bijections

$$\text{hom}_\mathcal{D}(uV, X) = h_X(uV) \simeq h_U(V) = \text{hom}_\mathcal{C}(V, U),$$

natural in $V$ and in $X$ whenever $h_X \circ u$ is representable. Therefore the assignment $v : \mathcal{D} \to \mathcal{C} : X \mapsto U$ is a partially defined functor, which is right adjoint to $u$. 
Constructing $v$ explicitly in this way requires the global choice of a representing object each time $h_X \circ u$ is representable, this is a technical difficulty that we will pass over quickly by stating that all such functors are naturally isomorphic.

Now we consider a fixed arrow $f : T \to S$ in our category $\mathcal{C}$ and consider specifically functors between the corresponding slice categories $\mathcal{C}/T$ and $\mathcal{C}/S$—the arrows in these categories are said to be $T$-linear and $S$-linear.

There is an adjoint pair of functors

$$f_! \dashv f^* : \mathcal{C}/T \leftrightarrow \mathcal{C}/S.$$ 

The functors $f_!$ and $f^*$ are defined on objects and their structure morphisms by

$$f_! : \mathcal{C}/T \to \mathcal{C}/S : (X, q_X) \mapsto (X, f \circ q_X)$$

$$f^* : \mathcal{C}/S \to \mathcal{C}/T : (X, q_X) \mapsto (X \times_S T, p_2),$$

and the fact that these form an adjoint pair follows immediately from the universal property of a pullback as fibered product. The question whether $f^*$ admits a right adjoint $f^*$ is of particular interest—this is closely related to the existence of an internal Hom-functor. The right adjoint can be found through the formalism introduced above, with $f^*$ playing the role of $u$. In general, $f^*$ will only be partially defined. If for some $X \in \text{Ob}(\mathcal{C}/T)$ the corresponding object $f_* X \in \text{Ob}(\mathcal{C}/S)$ is defined, there are bijections

$$\text{hom}_{\mathcal{C}/T}(f^* Y, X) \to \text{hom}_{\mathcal{C}/S}(Y, f_* X) : f \mapsto f^b,$$

natural in $Y$ and $X$, whenever $f_*$ is defined at $X$. The inverse map of $b$ will be denoted by $\#$. All this information is implied by writing

$$f_! \dashv f^* \dashv f_* : \mathcal{C}/T \leftrightarrow \mathcal{C}/S.$$ 

We generalize this to our context of a category $\mathcal{C}$ with endomorphism $F$ of the identity functor. In such a situation, we are often confronted
with a diagram as depicted here on the left and want to obtain a new diagram, as depicted on the right.

\[
\begin{align*}
\begin{array}{ccc}
\text{\(f^* A\)} & \xrightarrow{u} & \text{\(B\)} \\
\downarrow & & \downarrow \\
\text{\(T\)} & \xrightarrow{F_T} & \text{\(T\)}
\end{array}
\quad \Longleftrightarrow \quad
\begin{array}{ccc}
\text{\(A\)} & \xrightarrow{u^\flat} & \text{\(f_* B\)} \\
\downarrow & & \downarrow \\
\text{\(S\)} & \xrightarrow{F_S} & \text{\(S\)}
\end{array}
\end{align*}
\]

The difficulty is that it is not \textit{a priori} possible to consider \(u^\flat\), because \(u\) is not \(T\)-linear. In other words, \(u\) is simply not in the domain of a \(\flat\)-map. So we seek to extend the calculus of the adjunction \(f^* \dashv f_*\) to \(F_T\)-linear maps, and more generally \(F_T^n\)-linear maps, for any natural number \(n\).

Our first step is to formulate this problem. This requires us to define the \textit{fairies} \(\mathcal{C}_{/S}^{(F)}\) as follows. (The word fairy is short for \(F\)-ary category; one could also call them more verbosely \textit{semi-linear slice categories}.)

\textbf{Definition 8.6.1.} The objects of the fairy \(\mathcal{C}_{/S}^{(F)}\) are the arrows \(q_X : X \to S\). A morphism between \(X \to S\) and \(Y \to S\) is a pair \((f, n)\), where \(f : X \to Y\) and \(n\) is a natural number, such that \(F^n_S \circ q_X = q_Y \circ f\).

\textbf{Remark 8.6.2.} The following remarks are all important for working with fairies. We leave the easy proofs to the reader, insofar required.

(i) Let \(\mathcal{N}\) be a category with a single object \(\bullet\) such that \(\text{End}_{\mathcal{N}}(\bullet) = (\mathbb{N}, +)\). Then we have a diagram of categories

\[
\mathcal{C}_{/S} \xrightarrow{\text{inc}} \mathcal{C}_{/S}^{(F)} \xrightarrow{\text{pr}} \mathcal{N},
\]

where the first functor sends an object to itself and an arrow \(u\) to \((u, 0)\) and the latter functor sends every object to \(\bullet\) and an arrow \((u, n)\) to \(n\). We will always see \(\mathcal{C}_{/S}\) as the \textit{wide subcategory} (i.e. a subcategory containing all the objects) of \textit{linear morphisms} in the fairy \(\mathcal{C}_{/S}^{(F)}\) through the functor \(\text{inc}\). We also denote the inclusion simply by \(\text{inc} : X \mapsto \overline{X}\) and \(u \mapsto \overline{u}\). We will use the notation \(\rightarrow\) to warn the reader that an arrow is possibly non-linear when drawing fairy diagrams.
(ii) There is a functor $G : \mathcal{N} \to \mathcal{C}$ defined by $(\bullet \xrightarrow{n} \bullet) \mapsto (S \xrightarrow{F^n} S)$ and with the help of this functor one may define $\mathcal{C}^{(F)}_{/S}$ as either the fibered product $\vec{\mathcal{C}} \times_{p,\mathcal{C},G} \mathcal{N}$ of categories, where $\vec{\mathcal{C}}$ is the category of arrows in $\mathcal{C}$ and $p : \vec{\mathcal{C}} \to \mathcal{C} : (X \to Y) \mapsto Y$ the codomain fibration, or as the comma category $\text{id}_\mathcal{C} \downarrow G$. These constructions provide a natural variation on the theme of a slice category.

(iii) We will say a functor $u : \mathcal{C}/S \to \mathcal{D}/T$ between slice categories extends semi-linearly or simply extends if there is a functor $u : \mathcal{C}^{(F)}_{/S} \to \mathcal{D}^{(F)}_{/T}$ such that $u \circ \text{inc} = \text{inc} \circ u$.

(iv) The Yoneda extension of the inclusion $\text{inc}$ is a functor $\text{inc}$ which fits into the following diagram with the Yoneda embeddings:

\[
\begin{array}{ccc}
\mathcal{C}/S & \xrightarrow{\text{inc}} & \mathcal{C}^{(F)}_{/S} \\
\downarrow y & & \downarrow y \\
\mathcal{C}^{(F)}_{/S} & \xrightarrow{\text{inc}} & \mathcal{C}^{(F)}_{/S}
\end{array}
\]

If we also denote $\text{inc}! (\mathcal{F}) = \overline{\mathcal{F}}$, commutativity of the diagram can be written as $h_X = h_{\overline{X}}$.

(v) The category $\mathcal{C}^{(F)}_{/S}$ is itself a category with an endomorphism of the identity functor, also denoted by $F$ or by $F_{/S}$ if confusion is possible, with component at $X$ given by $F_X = (F_X, 1)$.

(vi) The universal property of the pullback $Y \leadsto Y \times_{q_Y,F^n_S} S = (F^n_S)^*Y$ implies that every arrow $(f, n + m) : X \to Y$ in $\mathcal{C}^{(F)}_{/S}$ factors into a morphism denoted $\langle f \mid m \rangle$ followed by a projection $(p_1, n)$:

\[
X \xrightarrow{(f|m)} Y \times_{q_Y,F^n_S} S \xrightarrow{(p_1,n)} Y.
\]

Thanks to natural isomorphism $(F^n_S)^* \simeq (F^n_S)^n$ we may and will in practice identify $\langle \cdots \langle (f \mid 1) \mid 1 \cdots | 1 \rangle \rangle$ with $\langle f \mid n \rangle$ and also $\langle f \mid 0 \rangle$ with $f$. This is mainly of importance when $m = 0$ and we have factored $(f, n)$ into a linear morphism $\langle f \mid n \rangle$ followed by a projection.
(vii) For instance $F_X$ factors via $\langle F_X | 1 \rangle$ which we identify with $F_{X/S}$ (see Section 8.5) through the inclusion $\text{inc}$. Therefore, recalling our notation $\Delta X = F^*_S X = X \times_{q_X,F_S} S$, we obtain a canonical factorization

$$X \xrightarrow{F_{X/S}} \Delta X \xrightarrow{P_{X/S}} X.$$ 

We may organise this information as follows: there is a functor $\Delta$, sometimes denoted $\Delta_S$ for clarity, together with natural transformations $F/S$ and $P/S$ between $\Delta$ and the identity functor as follows:

$$\begin{array}{c}
\text{id} \\
\id
\end{array} \xrightarrow{\eta} \xrightarrow{\varepsilon} \xrightarrow{\varepsilon} \xrightarrow{\eta} \xrightarrow{\id}$$

Our goal for the rest of the paragraph is to prove the following proposition.

**Proposition 8.6.3.** The adjunction $f_! \dashv f^* \dashv f_*$ extends (see Remark 8.6.2.(iii)) to the fairies

$$f_! \dashv f^* \dashv f_* : \mathcal{C}^{(F)}_{/T} \xrightarrow{\eta} \mathcal{C}^{(F)}_{/S}.$$ 

Proving this proposition requires us to extend these functors to all arrows of the fairies $\mathcal{C}^{(F)}_{/S}$ and proving that the resulting functors are still adjoint pairs. This holds no serious difficulty for $f_!$ and $f^*$, as we will see immediately in Proposition 8.6.4. For $f_*$ however, we know no direct way to extend the domain of definition to the non-linear morphisms and this causes most of the technical difficulties in this section. So we will use an indirect approach where we introduce the notion of a bewitched functor and study its right adjoints; the proof then follows by observing that $f^*$ is indeed bewitched.

**Proof of Proposition 8.6.3.**

The first part will be shown in Proposition 8.6.4; for the second part apply Proposition 8.6.8 to Remark 8.6.6.(iv).
Proposition 8.6.4. The adjunction $f_! \dashv f^* : \mathcal{C}/T \to \mathcal{C}/S$ extends semi-linearly.

Proof. Of course we must define $f_!(X) = f_!(X)$ on objects. (And in fact, $f_!(u,0) = (f_!u,0)$.) An arrow $(g,n) : X \to Y$ induces a commutative diagram in $\mathcal{C}$

$$
\begin{array}{ccc}
X & \xrightarrow{g} & Y \\
\downarrow^{q_X} & & \downarrow^{q_Y} \\
T & \xrightarrow{F^n_T} & T \\
\downarrow^{f} & & \downarrow^{f} \\
S & \xrightarrow{F^n_S} & S,
\end{array}
$$

and therefore an arrow $f_!(g,n) : f_!X \to f_!Y$. We leave to the reader the straightforward verification that this defines a functor.

Defining $f^*$ on objects—and linear arrows—is trivial, so let $(g,n) : U \to V$ be an arrow in $\mathcal{C}^F_{/S}$. The following diagram commutes, since every square commutes:

\begin{tikzcd}
& U \times_S T \\
T & S & U \\
& T & S \\
& F^n_T & F^n_S \\
T & S & V
\end{tikzcd}

It we erase the interior and replace it with a pullback square for $V \times_S T$, we get that the following diagram (without the dashed arrow)
8.6. The concept of a fairy

\[ U \times_S T \xrightarrow{p_1} U \]
\[ V \times_S T \xrightarrow{p_1} V \]
\[ T \xrightarrow{F^\alpha_T} T \xrightarrow{q_T} S \]

This implies the dashed arrow is uniquely defined by the pullback \( V \times_S T = f^*V \) and this is \( f^*g : f^*U \to f^*V \). We leave to the reader the straightforward verification that this defines a functor. Also the fact that the extended functors define an adjoint pair \( f \dashv f^* \) on the fairies is easy to verify and left to the reader. (Note that the unit and counit transformations are already determined by \( f_1 \) and \( f^* \).) \qed

Now comes the hard part: extending \( f_* \) on its domain. It is easy enough to define a functor \( f_* \) formally as the partially defined right adjoint of \( f^* \) but what is not obvious is that the functor \( f^* \) extends \( f_* \), i.e. that both functors agree on objects and linear arrows.

**Definition 8.6.5.** A functor \( \alpha : \mathcal{C}^{(F)}_{/S} \to \mathcal{D}^{(G)}_{/T} \) between fairies is bewitched if \( G_\alpha(X) = \alpha(F_X) \) for every \( X \in \text{Ob}(\mathcal{C}_{/S}) \) and it preserves the decomposition from Remark 8.6.2.(vi).

**Remark 8.6.6.**

(i) To avoid any confusion, we provide some details concerning the Definition 8.6.5: it says that there is a natural isomorphism \( \Delta_T \circ \alpha \simeq \alpha \circ \Delta_S \) (with linear components) such that for every arrow \( (u,n) : X \to Y \) in \( \mathcal{C}^{(F)}_{/S} \) with its decomposition into \( \langle u \mid n \rangle : X \to \Delta^nY \) and \( P^n_{Y/S} : \Delta^nY \to Y \) we have \( \alpha \langle u \mid n \rangle = \langle \alpha(u) \mid n \rangle \) and \( \alpha(P^n_{Y/S}) = P^n_{\alpha(Y)/T} \), up to the identification \( \alpha \Delta^nY \cong \Delta^n\alpha Y \), as in the following diagram:

\[
\begin{array}{ccc}
\alpha(X) & \xrightarrow{\alpha} & \alpha(\Delta^nY) \\
\alpha(\langle u \mid n \rangle) & \downarrow & \alpha(P^n_{Y/S}) \\
\Delta^nT(\alpha Y) & \xrightarrow{\alpha} & \alpha(Y) \\
\end{array}
\]
(ii) Let us take in particular $n = 0$. Since the lower path in the above diagram is the (unique) decomposition in a linear arrow and a projection, and since the vertical identification is linear, we have that $\alpha\langle u | n \rangle = \alpha(u)$ is linear. Therefore, a bewitched functor sends linear arrows to linear arrows and thus restricts to a functor $\alpha^0 : \mathcal{E}/S \to \mathcal{D}/T$.

\[
\begin{array}{c}
\mathcal{E}/S \\
\alpha^0 \downarrow \quad \quad \downarrow \text{inc}
\end{array}
\begin{array}{c}
\mathcal{D}/T
\end{array}
\begin{array}{c}
\mathcal{E}/S^{(F)} \\
\alpha \downarrow \quad \quad \downarrow \text{inc}
\end{array}
\begin{array}{c}
\mathcal{D}/T^{(F)}
\end{array}
\]

In Proposition 8.6.8 we will denote a bewitched functor by $\overline{\alpha}$ and its restriction by $\alpha$ so that this diagram reads on objects: $\overline{\alpha}(X) = \alpha(X)$.

(iii) Another immediate consequence of the definition is that a bewitched functor preserves the entire diagram that we drew in Remark 8.6.2.(vii), since this diagram encodes the decomposition of $F = P/S \circ F/S$.

(iv) The functor $f^*$ as defined in Proposition 8.6.4 is bewitched, this is easy to verify with the natural isomorphism $F_T^* \circ f^* \simeq (f \circ F_T)^* = (F_S \circ f)^* \simeq f^* \circ F_S^*$.

**Proposition 8.6.7.** Let $\alpha : \mathcal{E}/S^{(F)} \to \mathcal{D}/T^{(G)}$ be a bewitched functor together with its (perhaps partially defined) left and right adjoints $\gamma \dashv \alpha \dashv \beta$. If $X \in \text{Ob}(\mathcal{D}/T^{(G)})$, then $\beta G_X = F_{\beta(X)}$ and $\gamma G_X = F_{\gamma(X)}$.

**Proof.** Let us show this for $\beta$, the proof for $\gamma$ is similar. Consider an arbitrary $X \in \text{Ob}(\mathcal{D}/T^{(G)})$. Since $\alpha$ is bewitched, we have that $\alpha F_{\beta X} = G_{\alpha \beta X}$. Therefore there is a diagram

\[
\begin{array}{c}
\alpha \beta X \\
\alpha F_{\beta X} \downarrow \quad \quad \downarrow G_X
\end{array}
\begin{array}{c}
\alpha \beta X \\
\alpha \beta X \eta_X \rightarrow X
\end{array}
\begin{array}{c}
X
\end{array}
\]
where the horizontal arrows are units of the adjunction $\alpha \dashv \beta$. We can use the adjunction to push $\alpha$ to the right, taking into account that $(\eta_X)^\flat = \text{id}_{\beta X}$, we get that $\beta G_X = F\beta X$.

\textbf{Proposition 8.6.8.} Consider a bewitched functor $\overline{\alpha} : C^{(F)}_{/S} \to C^{(F)}_{/T}$ with restriction $\alpha$. Consider the partial right adjoints $\alpha \dashv \beta$ and $\overline{\alpha} \dashv \overline{\beta}$. Then $\overline{\beta}$ extends $\beta$.

\textit{Proof.} Consider the left diagram below, which is a commuting diagram in the category of categories. Each of the occurring functors $\gamma : \mathcal{U} \to \mathcal{V}$ lifts in two ways to functors $\gamma_! : \mathcal{U} \to \mathcal{V}$ and $\gamma^* : \mathcal{V} \to \mathcal{U}$ between the corresponding categories of presheaves, and these functors form adjoint pairs $\gamma_! \dashv \gamma^*$.

This implies that we may form the diagram on the right, which is \textit{not} a commuting diagram; but the square is inhibited by a natural transformation $\text{inc}_! \circ \alpha^* \Rightarrow \overline{\alpha}^* \circ \text{inc}_!$ which is called the \textit{Beck–Chevalley transformation}.

\[
\begin{array}{ccc}
C^{(F)}_{/S} & \xrightarrow{\alpha} & C^{(F)}_{/T} \\
\text{inc} \downarrow & & \downarrow \text{inc} \\
C^{(F)}_{/S} & \xrightarrow{\overline{\alpha}} & C^{(F)}_{/T}
\end{array}
\quad
\begin{array}{ccc}
\overline{C}^{(F)}_{/S} & \xleftarrow{\alpha^*} & \overline{C}^{(F)}_{/T} \\
\text{inc}_! \downarrow & & \downarrow \text{inc}_! \\
\overline{C}^{(F)}_{/S} & \xleftarrow{\overline{\alpha}^*} & \overline{C}^{(F)}_{/T}
\end{array}
\]

This natural transformation arises as follows:

\[
\text{inc}_! \circ \alpha^* \Rightarrow \text{inc}_! \circ \alpha^* \circ \text{inc}_!^* \circ \text{inc}_! \\
\Rightarrow \text{inc}_! \circ \text{inc}_!^* \circ \overline{\alpha}^* \circ \text{inc}_! \\
\Rightarrow \overline{\alpha}^* \circ \text{inc}_!
\]

Here in the successive steps we have applied the unit and counit natural transformations of the adjunctions associated to each of the inclusion functors, as well as the fact that the left square commutes. We will now show that this transformation is actually natural isomorphism; one then says that the \textit{Beck–Chevalley condition holds}.

Let us first explain how this implies the statement of the proposition. If $\beta$ is defined at $X$ then we have $h_X \circ \alpha \simeq h_{\beta(X)}$; applying the
inclusion we get
\[ \overline{\text{h}_X} \circ \alpha \simeq \text{h}_{\beta(X)} \simeq \text{h}_{\overline{\beta(X)}}. \]

The Beck–Chevalley isomorphism says that
\[ \overline{\text{h}_X} \circ \alpha = \text{inc}_!(\alpha^*(\overline{\text{h}_X})) \simeq \overline{\alpha}^*(\text{inc}_!(\text{h}_X)) = \overline{\text{h}_X} \circ \overline{\alpha} \simeq \text{h}_X \circ \overline{\alpha}. \]

Therefore \[ \overline{\text{h}_X} \circ \overline{\alpha} \simeq \text{h}_{\overline{\beta(X)}} \] and thus \( \overline{\beta} \) is defined at \( X \) and in fact represented by \( \overline{\beta(X)} \).

We will now apply the criterium [Gui14, 1.18]—proven in [Gui80]—to verify that that the diagram on the left, depicted again here for the reader’s convenience, is an exact square. This implies the Beck-Chevalley condition as in [Gui14, 1.15]. There is also an exposition of this material on the nlab [nLab].

\[
\begin{array}{ccc}
\mathcal{C}/S & \xrightarrow{\alpha} & \mathcal{C}/T \\
\text{inc} & & \text{inc} \\
\mathcal{C}^{(F)}/S & \xrightarrow{\overline{\alpha}} & \mathcal{C}^{(F)}/T
\end{array}
\]

To apply the criterium, we need to consider arbitrary objects \( Y \) in \( \mathcal{C}/T \) and \( Z \) in \( \mathcal{C}^{(F)}/S \) and a morphism \( (u, n) : Y \to \overline{\alpha}(Z) = \overline{\alpha}(Z) \) in \( \mathcal{C}^{(F)}/T \). First we need to show that there exists a triple \( (X, y, (z, m)) \) where \( X \) is an object in \( \mathcal{C}/S \) and
\[
y : Y \to \alpha(X) \\
(z, m) : X \to Z
\]
are morphisms such that the composition in the fairy \( \mathcal{C}^{(F)}/T \)
\[
Y \xrightarrow{y} \overline{\alpha}(X) \xrightarrow{\overline{\alpha}(z, m)} \overline{\alpha}(Z)
\]
is equal to \( (u, n) \). To see this, let us choose \( X = \overline{\triangle^n Z} \), and recall that we have an identification \( \overline{\alpha}(X) \cong \triangle^n \overline{\alpha}Z \). Thus there is a decomposition of \( (u, n) \) into a linear morphism \( \langle u \mid n \rangle : Y \to \alpha(Z) \) and a projection \( \triangle^n \overline{\alpha}Z \to \overline{\alpha}Z \) which, since the functor is bewitched, is also
of the projection $P_{\mathbb{Z}/T}^n : \triangle^n \mathbb{Z} \to \mathbb{Z}$. Thus the choice $y = \langle u \mid n \rangle$ and $(z, n) = (P_{\mathbb{Z}/T}^n, n)$ provides the sought decomposition.

Next, we need to assume that $(X', y', (z', m))$ is another solution to the same problem and show that $X$ and $X'$ are connected in the category $\mathcal{C}/S$ through a zigzag in such a way that the induced diagram commutes. (This is the so called lantern diagram.) More precisely, we will show that there is a morphism $v : X' \to X$ in $\mathcal{C}/S$ such that the following diagrams commute:

\[
\begin{array}{c}
Y \\
\downarrow y \ & \ & \downarrow y' \\
\alpha(X') \ & \ & \alpha(X) \\
\downarrow \alpha(v) \\
\overline{X}' \ & \ & \overline{X} \\
\downarrow (z', m) \ & \ & \downarrow (z, n) \\
\mathbb{Z} \\
\end{array}
\]

Obviously $m = n$ and it is clear that linearizing the morphism $(z', n) : \overline{X}' \to \mathbb{Z}$ via $\overline{X} = \triangle^n \mathbb{Z}$, we obtain a linear morphism $v : X' \to X$ which fits into the lower diagram. Applying $\alpha$ to this diagram we obtain the following commuting diagram, where the dashed arrow is defined by composition.

\[
\begin{array}{c}
Y \\
\downarrow y' \ & \ & \downarrow \alpha(v) \\
\overline{X}' \ & \ & \overline{Z} \\
\end{array}
\]

\[
\begin{array}{c}
\alpha(X) \\
\downarrow \alpha(v) \\
\overline{X}' \\
\end{array}
\]

\[
\begin{array}{c}
\alpha(Z) \\
\downarrow \alpha(z, n) \\
\overline{Z} \\
\end{array}
\]

\[
\begin{array}{c}
\alpha(X) \\
\downarrow \alpha(v) \\
\overline{X}' \\
\end{array}
\]

\[
\begin{array}{c}
\alpha(Z) \\
\downarrow \alpha(z', n) \\
\overline{Z} \\
\end{array}
\]

Up to the canonical identification $\alpha(X) \cong \triangle^n(\alpha(Z))$, the dashed arrow is a linear arrow with the property that composing it with the projection $\triangle^n(\alpha(Z)) \to \alpha(Z)$ yields $u$. But this property defines $y = \langle u \mid n \rangle$. \qed
8.7 Restriction to visible objects

In Section 8.6 we studied the functor of base change $f^*: \mathcal{C}_S \rightarrow \mathcal{C}_T$ and its partially defined right adjoint $f_*: \mathcal{C}_T \rightarrow \mathcal{C}_S$ and extended this formalism to semi-linear morphisms as provided by the fairies. We will now apply these ideas to study a right adjoint to base change in the mixed category $m\mathcal{C}$. Before we do so, let us motivate why we are interested in this situation. The real motivation is of course that we would like to prove Theorem 10.5.1 eventually, which states that certain exotic pseudo-reductive groups are Weil restrictions of mixed reductive groups. This theorem will ultimately rely on the work in the present section, which formulates the precise way in which the mixed and ordinary versions of right adjoints of base changes interact—in the context of algebraic geometry, this will tell us how to take a Weil restriction of a mixed scheme.

Nonetheless, we provide some intrinsic motivation, based on the observation that this functor $f_*$ provides a way for invisible mixed objects to invade the visible world, giving rise to exotic phenomena.

Recall once again that we consider the category $\mathcal{C}$ as a (full) subcategory of $m\mathcal{C}$ through the functor $m$; recall that these objects are called visible. Also recall the adjunctions

$$c_1 \dashv m \dashv c_2 : \quad m\mathcal{C} \leftrightarrow \mathcal{C}$$

It is thanks to these adjunctions that the mixed object $\tilde{X}$ can be understood as some kind of mixture of its components $c_i(\tilde{X})$.

Let us now focus on the relative situation, with respect to a base object $\tilde{S}$. If we assume that $\tilde{S} = m(S)$ is itself visible then by good behaviour of adjunctions with slice categories, as is explained nicely on the nLab [nLaa, (3.1)], we get new adjunctions for free:

$$\left( c_1 \right)_{/S} \dashv m_{/S} \dashv \left( c_2 \right)_{/S} : \quad m\mathcal{C}_{/S} \leftrightarrow \mathcal{C}_{/S}$$

This too holds for us the interpretation that a mixed $S$-object is a mixture of two ordinary $S$-objects. This is not a surprise, since they are related to relative factorizations as we saw in Section 8.5.
If $\tilde{S} = (S, S', \alpha, \beta)$ is invisible, the situation becomes more interesting. The morally correct way of seeing an object $\tilde{X} = (X, X', \varphi, \psi)$ as a mixture of two ordinary objects, would be to see it as a mixture of an $S$-object $X$ and $S'$-object $X'$ through the adjunctions between $m'\mathcal{C}$ and $\mathcal{C}$. But often one insists on working over a fixed base object $S$ and there, more adjunction hocus-pocus as in [nLaa, (3.1)] can only give us:

$$
\begin{array}{c}
\begin{array}{c}
(c_1)_{/\tilde{S}} \dashv m_{/\tilde{S}} \dashv ?? : m'\mathcal{C}_{/\tilde{S}} \rightleftarrows \mathcal{C}/S
\end{array}
\end{array}
$$

The question marks mean that defining a right adjoint to $m_{/\tilde{S}}$ is not in general possible. But as we explained in the introduction to Section 8.6, such a functor can still be partially defined at some objects. In these cases, we obtain objects in the category $\mathcal{C}/S$ which somehow look exotic. (Our Theorem 10.5.1 says that this is how the exotic pseudo-reductive groups arise.)

To better understand the right adjoint to $m_{/\tilde{S}}$, let us recall how $m_{/\tilde{S}}$ is defined in this case. Let us denote by $f : \tilde{S} \to S = m(c_1(\tilde{S}))$ the unit of the adjunction $c_1 \dashv m$:

$$
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array{c}
**Proposition 8.7.1.** Consider a mixed object \( \tilde{S} \) together with its morphism \( f : \tilde{S} \to S = m_{c_1}(\tilde{S}) \). Let \( \tilde{X} = (X, X', \varphi, \psi) \) be an \( \tilde{S} \)-object and assume \( \beta_* \beta^* X \) and \( \beta_* X' \) exist. If we define

\[
f_*(\tilde{X}) = (X, X \times_{\beta_* \beta^* X} \beta_* X', \pi, p_1),
\]

then for all \( S \)-objects \( \tilde{T} \):

\[
\text{hom}_S(\tilde{T}, f_*(\tilde{X})) \simeq \text{hom}_{\tilde{S}}(f^* \tilde{T}, \tilde{X})
\]

The map \( \pi \) and the maps defining \( Y = X \times_{\beta_* \beta^* X} \beta_* X' \) are specified in the proof.

**Proof.** In the following reasoning, we will be working with the adjunction of fairies

\[
\beta^* \dashv \beta_* : \mathcal{E}^{(F)}_S \leftrightarrow \mathcal{E}^{(F)}_{S'}.
\]

Although we will still use the notation \( \Rightarrow \) to warn the reader for non-linear arrows, we will denote the arrow \( (u, n) \) simply by \( u \). The number \( n \) is always 0 or 1 so it can be read off from the diagrams.

**Step 1: construction of \( \pi \).** Let us start from the base change \( \tilde{X} \times_{\tilde{S}} S' \). It is given by the \( S' \)-object

\[
(\beta^* X, X', \varphi \circ p_1, \psi'),
\]

where \( p_1 \circ \psi' = \psi \) and \( p_1 : \beta^* X \to X \) denotes the canonical projection. So if start from the diagram

\[
F_{\beta^* X} : \beta^* X \xrightarrow{\varphi \circ p_1} X' \xrightarrow{\psi'} \beta^* X
\]

and apply the adjunction \( \beta^* \dashv \beta_* \), we get

\[
(F_{\beta^* X})^b : X \xrightarrow{(\varphi \circ p_1)^b} \beta_* X' \xrightarrow{\beta_* \psi'} \beta_* \beta^* X. \quad (Y_1)
\]

On the other hand, starting from \( F_{\beta^* X} : \beta^* X \xrightarrow{\text{id}_{\beta^* X}} \beta^* X \xrightarrow{\beta_* \beta^* X} \beta^* X \) instead, we obtain

\[
(F_{\beta^* X})^b : X \xrightarrow{(\text{id}_{\beta^* X})^b} \beta_* \beta^* X \xrightarrow{\beta_* \beta^* X} \beta_* \beta^* X.
\]
The first arrow is just a cumbersome notation for $\eta_X$, the component at $X$ of the unit $\eta$ of the adjunction $\beta^* \dashv \beta_*$. Expressing that $\eta$ is a natural transformation $\text{id}_\mathcal{C} \to \beta_* \beta^*$, we have a diagram

$$
\begin{array}{ccc}
X & \xrightarrow{\eta_X} & \beta_* \beta^* X \\
\downarrow F_X & & \downarrow \beta_* \beta^* F_X \\
X & \xrightarrow{\eta_X} & \beta_* \beta^* X,
\end{array}
$$

where we note that $\beta_* \beta^* F_X = \beta_* F \beta^*_X$ by Proposition 8.6.7 to obtain

$$(F_{\beta^* X})^\flat : \ X \xrightarrow{F_X} X \xrightarrow{\eta_X} \beta_* \beta^* X.$$  \hfill (Y_2)

Combining $(Y_1)$ and $(Y_2)$ we get the following diagram, where the dotted arrow is implied by the pullback square, i.e. we define $\pi = F_x \times (\varphi \circ p_1)^\flat$.

\begin{tikzpicture}
  \node (X) at (0,0) {$X$};
  \node (Y) at (4,0) {$Y$};
  \node (Xp) at (8,0) {$\beta_* X'$};
  \node (Yp) at (4,4) {$Y'$};
  \node (Xpp) at (0,4) {$\beta_* X$};
  \node (Ypp) at (8,4) {$\beta_* \beta^* X$};

  \draw[->] (X) to node[above]{$(\varphi \circ p_1)^\flat$} (Y);
  \draw[->] (X) to node[above]{$\pi$} (Xp);
  \draw[->] (Y) to node[above]{$p_2$} (Xp);
  \draw[->] (Y) to node[below]{$p_1$} (X);
  \draw[->] (Xp) to node[above]{$\beta_* \psi'$} (Yp);
  \draw[->] (Xp) to node[below]{$\beta_* \beta^* Y$} (Ypp);
  \draw[->] (X) to node[below]{$\eta_X$} (Xpp);
  \draw[->] (Y) to node[below]{$\eta_Y$} (Ypp);
\end{tikzpicture}

**Step 2: verifying** $p_1 \circ \pi = F_X$ and $\pi \circ p_1 = F_Y$. The first of these identities is clear from the construction of $\pi$. To show the second one\(^3\), we first consider the following diagram:

\begin{tikzpicture}
  \node (Y) at (0,0) {$Y$};
  \node (Yp) at (4,0) {$\beta_* X'$};
  \node (Ypp) at (4,4) {$\beta_* \beta^* Y$};
  \node (Xpp) at (0,4) {$\beta_* \beta^* X$};
  \node (X) at (0,4) {$X$};

  \draw[->] (Y) to node[above]{$p_2$} (Yp);
  \draw[->] (Y) to node[below]{$p_1$} (Ypp);
  \draw[->] (Yp) to node[below]{$\beta_* \psi'$} (Ypp);
  \draw[->] (Ypp) to node[below]{$\beta_* \beta^* p_1$} (Xpp);
  \draw[->] (Xpp) to node[below]{$\eta_X$} (X);
  \draw[->] (Ypp) to node[below]{$\eta_Y$} (Y);
\end{tikzpicture}

\(^3\)This is trivial if $\pi$ is epic since then $\pi \circ p_1 \circ \pi = F_Y \circ \pi = \pi \circ F_X$.\vspace{2cm}
The bottom triangle commutes again because $\eta$ is a natural transformation; the big square commutes by definition of $Y$. Thus the upper triangle commutes. Using this triangle, but noting that the diagonal is also $(\beta^* p_1)^\flat$, we obtain the diagram below on the left; using the adjunction we can push $\beta_*$ from the diagonal to $Y$ and get the diagram on the right:

$$
\begin{array}{ccc}
Y & \xrightarrow{p_2} & \beta_* X' \\
\downarrow{(\beta^* p_1)^\flat} & & \downarrow{\beta_* \psi} \\
\beta_* \beta^* X & \Rightarrow & \beta^* X
\end{array}
= 
\begin{array}{ccc}
\beta^* Y & \xrightarrow{p_2^\sharp} & X' \\
\downarrow{\beta^* p_1} & & \downarrow{\psi} \\
\beta^* X
\end{array}
$$

Let us add to this diagram two more triangles which clearly commute in order to obtain the diagram on the left below; omitting the dotted arrows and pushing $\beta^*$ to the right with the adjunction again, we get the diagram on the right.

$$
\begin{array}{ccc}
\beta^* Y & \xrightarrow{p_2^\sharp} & X' \\
\downarrow{\beta^* p_1} & & \downarrow{\psi} \\
\beta^* X & \Rightarrow & \beta_* X' \\
\downarrow{\beta^* \pi} & & \downarrow{\phi p_1} \\
\beta^* Y & \xrightarrow{p_2^\sharp} & X' \\
\end{array}
= 
\begin{array}{ccc}
Y & \xrightarrow{p_2} & \beta_* X' \\
\downarrow{p_1} & & \downarrow{\phi \beta_* F_{X'}} \\
X & \Rightarrow & \beta_* X' \\
\downarrow{\phi \pi} & & \downarrow{\phi p_2} \\
Y & \xrightarrow{p_2} & \beta_* X'. \\
\end{array}
$$

On the other hand, we clearly have a square

$$
\begin{array}{ccc}
Y & \xrightarrow{p_1} & X & \xrightarrow{\phi} & \bar{Y} \\
\downarrow{p_1} & & \downarrow{\text{id}_X} & & \downarrow{p_1} \\
X & \xrightarrow{\text{id}_X} & X & \xrightarrow{F_{X}} & X. \\
\end{array}
$$

Combining $(\square_1)$ and $(\square_2)$ with the pullback defining $Y$, we get a
commuting diagram

\[
\begin{array}{c}
\begin{array}{ccc}
Y & \xrightarrow{p_2} & \beta_*X' \\
\downarrow & & \downarrow \\
Y & \xrightarrow{p_2} & \beta_*X' \\
\downarrow & & \downarrow \\
X & \xrightarrow{FX} & X \\
\end{array}
\end{array}
\end{align*}
\]

\[\pi \circ p_1 \quad \Downarrow \quad \beta_*F_{X'} \]

\[\begin{array}{c}
\begin{array}{ccc}
P_1 & \xrightarrow{p_1} & \beta_*X' \\
\downarrow & & \downarrow \\
P_1 & \xrightarrow{p_1} & \beta_*X' \\
\downarrow & & \downarrow \\
X & \xrightarrow{\eta_X} & \beta_*\beta*X. \\
\end{array}
\end{array}
\]

On the other hand, it is clear that if we replace the dotted arrow with \(F_Y\), the diagram commutes as well. But this arrow is uniquely determined since the square in the bottom right is a pullback. So \(\pi \circ p_1 = F_Y\).

**Step 3: the maps between homsets.** Now we consider an arbitrary mixed \(S\)-object \(\tilde{T}\), say \(\tilde{T} = (T, T^\circ, \zeta, \xi)\). We compute \(f^*\tilde{T} = T \times_S \tilde{S} = (T, \beta^*T^\circ, \zeta', \xi \circ p_1)\), where \(p_1 : \beta^*T^\circ \to T^\circ\) is the canonical projection and \(\zeta = p_1 \circ \zeta'\). Now consider the following diagrams:

\[
\begin{array}{cc}
\begin{array}{c}
T \xleftarrow{\zeta'} \xrightarrow{\xi \circ p_1} \beta^*T^\circ \\
\downarrow u \quad \downarrow v \\
X \xleftarrow{\phi} \xrightarrow{\psi} X' \\
\downarrow q_X \quad \downarrow q_{X'} \\
S \xleftarrow{\beta} \xrightarrow{\alpha} S' \\
\end{array}
\end{array}
\begin{array}{c}
T \xrightarrow{\zeta} T^\circ \\
\downarrow x \quad \downarrow y \\
X \xleftarrow{\pi} \xrightarrow{p_1} Y' \\
\downarrow \quad \downarrow \\
S \xleftarrow{\text{id}_S} \xrightarrow{FS} S' \\
\end{array}
\end{array}
\]

What we must show is this: to every pair of morphisms \((u, v)\) which makes the left diagram commute, there corresponds uniquely a pair \((x, y)\) which makes the right diagram commute. We stress again that, as in Section 8.1, this is an abbreviation for a bigger commutative diagram since \(\bullet \circ \equiv \circ \) does not compose to the identity but rather to \(F_*\) and \(F_c\); furthermore all vertical compositions give the appropriate structural morphism. In fact, the correspondence is given by the formulas

\[
\begin{cases}
x = u \\
y = (u \circ \xi) \times v^b
\end{cases}
\quad \text{and} \quad
\begin{cases}
u = (p_2 \circ y)^z \\
\end{cases}
\]

\[\begin{cases}
x = u \\
y = (u \circ \xi) \times v^b
\end{cases}
\quad \text{and} \quad
\begin{cases}
u = (p_2 \circ y)^z
\end{cases}
\]

\[\begin{cases}
x = u \\
y = (u \circ \xi) \times v^b
\end{cases}
\quad \text{and} \quad
\begin{cases}
u = (p_2 \circ y)^z
\end{cases}
\]
Let us first verify that \((u \circ \xi) \times v^b\) actually defines a morphism \(y : T^o \to Y\), i.e. that \(\eta_X \circ u \circ \xi = \beta_* \psi' \circ v^b\). Using that \(v^b = \beta_* v \circ \eta_T\), where \(\eta_T : T \to \beta_* \beta^* T\) is the unit morphism, we must show that \(\eta_X \circ (u \circ \xi) = \beta_* (\psi' \circ v) \circ \eta_T\). By naturality of the adjunction, it suffices to verify that \(\psi' \circ v = \beta^*(u \circ \xi)\). To see this, we observe that the equality \(\psi \circ v = u \circ \xi \circ p_1\) implies that there is a diagram

\[
\begin{array}{ccc}
\beta^* T^o & \xrightarrow{v} & X' & \xrightarrow{\psi'} & \beta^* X \\
\downarrow p_1 & & \downarrow p_1 & & \\
T^o & \xrightarrow{\xi} & T & \xrightarrow{v} & X
\end{array}
\]

Since the arrows \(v\) and \(\psi'\) are both \(S'\)-linear, we have that \(\psi' \circ v\) is an \(S'\)-linear arrow which makes the square above commute. The unique arrow with that property is \(\beta^*(u \circ \xi)\), which completes our verification.

**Step 4: verifying bijectivity.** We can quickly verify that the two maps defined in step 3 are inverses of one another. In one direction, it is immediately clear that

\[
\left( p_2 \circ ((u \circ \xi) \times v^b) \right)^\sharp = (v^b)^\sharp = v.
\]

In the other direction, we must verify that

\[
(x \circ \xi) \times (p_2 \circ y) = y
\]

but this is a direct consequence of \(p_1 \circ y = x \circ \xi\).

**Step 5: naturality.** We now verify that the constructed isomorphism is natural in both arguments. Let us only illustrate this in one case and leave the other verifications to the reader. Let \(\widetilde{W}\) be another \(\tilde{S}\)-object together with a morphism \((a, b) : \tilde{X} \to \widetilde{W}\). Then we have a corresponding square:

\[
\begin{array}{ccc}
\text{hom}_S(\tilde{T}, f_* \tilde{X}) & \longrightarrow & \text{hom}_S(f^* \widetilde{T}, \tilde{X}) \\
\downarrow & & \downarrow \\
\text{hom}_S(\tilde{T}, f_* \widetilde{W}) & \longrightarrow & \text{hom}_S(f^* \widetilde{T}, \widetilde{W})
\end{array}
\]
The vertical arrows are given by composition with \((a, s)\) and \((a, b)\) respectively, where we wish to construct

\[
s : X \times_{\beta_{s} \beta^{*} X} \beta_{s} X' \to W \times_{\beta_{s} \beta^{*} W} \beta_{s} W'
\]

by pairing \(a : X \to W\) and \(\beta_{s} b : \beta_{s} X' \to \beta_{s} W'\); such an arrow \(s\) will then in particular satisfy the equation \(p_{2} \circ s = \beta_{s} b \circ p_{2}\) and this immediately implies naturality condition for the square

\[
(p_{2} \circ s \circ y)^{y} = b \circ (p_{2} \circ y)^{y}.
\]

So to conclude the proof, we must only show that this pairing \(s\) indeed exists, i.e. that the diagram depicted below (without the dotted arrows) commutes, where \(\psi'_{W}\) denotes the analogon of the arrow \(\psi'\) for \(X\).

\[
\begin{array}{ccc}
X \times_{\beta_{s} \beta^{*} X} \beta_{s} X' & \xrightarrow{p_{2}} & \beta_{s} X' \\
\downarrow p_{1} & & \downarrow \beta_{s} b \\
X & \xrightarrow{\eta_{X}} & \beta_{s} \beta^{*} X \\
\downarrow a & & \downarrow \beta_{s} \psi' \\
W & \xrightarrow{\eta_{W}} & \beta_{s} \beta^{*} W
\end{array}
\]

Taking into account the dashed arrows, it is now sufficient that each of the three squares commutes, and only for the right square is this not trivial. To verify this, it again suffices to look at the upper square in the following diagram. (Ignore the lower square for now.)

\[
\begin{array}{ccc}
X' & \xrightarrow{b} & W' \\
\downarrow \psi' & & \downarrow \psi'_{W} \\
\beta^{*} X & \xrightarrow{\beta^{*} a} & \beta^{*} W \\
\downarrow p_{2} & & \downarrow p_{2} \\
X & \xrightarrow{a} & W
\end{array}
\]

Since we must now verify that the two paths from \(X' \to \beta^{*} W\) are equal, where we recall that \(\beta^{*} W\) is itself a fibered product \(W \times_{S} S'\), it suffices to verify that the compositions with both projections are
equal. For one projection \( \beta^* W \to S' \) this is trivial, since the entire upper square is an \( S' \)-linear diagram. For the other projection we can see this by gluing the lower square to the diagram: if we observe that the vertical arrows compose precisely to \( \psi \) and \( \psi_W \) then the big rectangle commutes precisely because \((a,b) : \tilde{X} \to \tilde{W} \) is a morphism of mixed objects, which concludes our verification that \( a \) and \( \beta_* b \) can be paired to the said arrow \( s \).

**Remark 8.7.2.** A few notes.

(i) Here is a heuristic: the object that one would like to write there is \( f_* \tilde{X} = (X, \beta_* X') \). But there is no obvious map \( \beta_* X' \to X \), so we form a product with \( X \) and let the projection play the role of that map.

(ii) In truth, we guessed this description of \( Y \) from [CGP15, §7.2] which in turn relies on Tits’s notes [Tit92, §4]. Tits probably started from his description of mixed abstract groups [Tit74, (10.3.2)] and somewhere along the way used some version of our Proposition 9.2.3 to come up with the corresponding exotic pseudo-reductive groups. In Remark 9.2.4.5 we will explain this more in detail.

**Proposition 8.7.3.** Let \( f : S \to T \) be a morphism of visible objects and let \( \tilde{X} = (X_1, X_2, \Phi_1, \Phi_2) \) be a mixed \( S \)-object. Then a right adjoint \( f_* \) to the base change \( f^* \) in \( C \) also defines a right adjoint to base change in \( mC \).

**Proof.** It is easily verified that \( f_*(\tilde{X}) = (f_* X_1, f_* X_2, f_* \Phi_1, f_* \Phi_2) \) has the required property that \( \hom_T(\tilde{Y}, f_* \tilde{X}) = \hom_S(f^* \tilde{Y}, \tilde{X}) \) for every mixed \( T \)-object \( \tilde{Y} \).  \( \square \)
To state our result, we need a definition.

Robin Hartshorne\textsuperscript{1}

We will now take a step towards the applications that we have in mind and apply the results of the previous section to the categories of schemes and rings of a fixed characteristic $p > 0$, choosing the absolute Frobenius as endomorphism of the identity functor.

It should be noted that one could also choose the identity endomorphism of the identity functor. We will only mention that this comes down to the study of schemes with an involution and would eventually lead to a slightly different description of the Steinberg groups $2\text{A}_n$, $2\text{D}_n$, $2\text{E}_6$. The main difference here is a shift of viewpoint: usually one regards, say, PSU in the real/complex case as an algebraic group over $\mathbb{R}$, whereas in our approach it would become a twisted group over the twisted field $(\mathbb{C}, \tau)$, where $\tau$ denotes complex conjugation. Since our main goal is to study and describe groups such as $2\text{B}_2$, $2\text{G}_2$, $2\text{F}_4$ we will not pursue this route any further. One could also consider schemes over $\mathbb{F}_q$, with $q = p^e$, together with the $e$'th power of the Frobenius for $F$.

Some examples will be grouped together in Section 9.4; the reader is encouraged to skip ahead.

\textsuperscript{1} Algebraic geometry. Graduate Texts in Mathematics, No. 52, 1977, p. 77.
9.1 Twisted and mixed schemes

Applying the results of the previous section to the category \((sch)_p\) of schemes of characteristic \(p\) with their absolute Frobenius \(F\) provides us with a number of categories and functors. We use the following notations for the occurring categories and their terminal objects, where we recall that \(1_{tC} = (1_C, \text{id})\) and \(1_{mC} = (2_C, \tau)\):

<table>
<thead>
<tr>
<th>general:</th>
<th>(\mathcal{C})</th>
<th>(F)</th>
<th>(t\mathcal{C})</th>
<th>(m\mathcal{C})</th>
<th>(1_C)</th>
<th>(1_{tC})</th>
<th>(1_{mC})</th>
</tr>
</thead>
<tbody>
<tr>
<td>schemes:</td>
<td>((sch)_p)</td>
<td>(Fr)</td>
<td>((tsch)_p)</td>
<td>((msch)_p)</td>
<td>(\text{Spec} F_p)</td>
<td>(F)</td>
<td>(E)</td>
</tr>
<tr>
<td>rings:</td>
<td>((ring)_p)</td>
<td>(fr)</td>
<td>((tring)_p)</td>
<td>((mring)_p)</td>
<td>(F_p)</td>
<td>(f)</td>
<td>(e)</td>
</tr>
</tbody>
</table>

We call \((tsch)_p\) resp. \((msch)_p\) the category of twisted resp. mixed schemes (of characteristic \(p\)). From now on, we will often omit the subscript \(p\). Recall the notion of the underlying ordinary object, in this case the underlying ordinary scheme of a twisted scheme \((X, \Phi_X)\), which is just the \(F_p\)-scheme \(X\).

**Proposition 9.1.1.** \((tsch)\) and \((msch)\) have fibered products \(X \times S Y\) and terminal objects.

**Proof.** This is an immediate consequence of Lemma 8.1.3. \(\square\)

Although the category \((ring)\) is not extensive, its opposite category \((ring)^{\text{op}} = (affsch)\) is. Actually, let us show this to convince the reader it is not a deep fact. If we have an algebra morphism \(f : F_p \times F_p \to A\), then \(e_1 = f(1,0)\) and \(e_2 = f(0,1)\) are orthogonal idempotents, so we obtain a decomposition \(A \cong A_1 \times A_2\) and the map \(f\) is actually built from \(f_1 : F_p \to A_1\) and \(f_2 : F_p \to A_2\), and the product functor

\[
(ring) \times (ring) \to (ring) / (F_p \times F_p) = (F_p \times F_p)\text{-alg}:
\]

\[
(A_1, A_2) \mapsto A_1 \times A_2
\]

is an equivalence.

If \(\tilde{S}\) is a twisted scheme, then by an \(\tilde{S}\)-scheme, we mean an arrow \(\tilde{X} \to \tilde{S}\), in other words an object of the slice category \((tsch)/\tilde{S}\). Similarly
if \( \tilde{R} \) is a twisted ring, then by an \( \tilde{R} \)-algebra we mean an arrow \( \tilde{R} \to \tilde{S} \), in other words an object of the coslice category \(((\text{ring})^{\text{op}})/(\tilde{R})^{\text{op}}\).

**Proposition 9.1.2.** We have equivalences \((\text{tsch}) \cong (\text{msch})_E\) and \((\text{mring}) \cong \text{e-alg}\). The contravariant functor \(\text{Spec} : (\text{ring}) \to (\text{sch})\) extends to the mixed and twisted categories in a way which commutes with all the functors \(f, c_1, c_2, m, \overline{m}, \delta^1, \delta^*, \delta^s, \delta^t\) and \(\tau^*\).

**Proof.** This is a consequence of Proposition 8.3.2 and Proposition 8.2.5. \(\square\)

An important observation is that the functors \(m, \overline{m} : (\text{sch}) \to (\text{msch})\) are not essentially surjective. Just like in the general case (Definition 8.2.3), we will use the adjectives \textit{visible} resp. \textit{anti-visible} for those mixed schemes that are isomorphic to \(m(X)\) resp. \(\overline{m}(X)\), for some ordinary scheme \(X\). A mixed scheme that is not visible is called \textit{invisible}.

Since the Frobenius of a scheme acts trivally on the underlying topological space, a twister acts with orbits of length 1 or 2.

**Definition 9.1.3.** We will use the adjective \textit{blended} for twisted schemes where all orbits have length 1, i.e. the twister acts trivially on the underlying topological space.

In particular every twisted field must be blended. These objects are also known as \textit{fields with Tits endomorphism}: they are just pairs \((k, \theta)\) where \(k\) is a field and \(\theta\) an endomorphism such that \(x^{\theta^2} = x^p\) for every \(x \in k\). A mixed ring can never have a field as its underlying ordinary ring, since mixed schemes are never connected. This allows us to define unambiguously:

**Definition 9.1.4.** A \textit{mixed field} is a mixed ring \(m = (k, \ell, \kappa, \lambda)\) such that \(k\) and \(\ell\) are fields.

Equivalently, it is a mixed ring without twisted ideals in the sense of Definition 9.3.6.
9.2 Rational points and functor of points

A standard tool in algebraic geometry is the functor of points associated to a scheme, for instance see [SGA3, exp. I]. We apply it to our setting by using for their category $\mathcal{C}$ one of the categories $(\text{tsch})$ or $(\text{msch})$.

To begin, we observe that a notion of rational points is provided by Definition 1.2 in *op. cit.*: we define $\Gamma(X) = \text{hom}(1, X)$, where $1$ is a terminal object in $\mathcal{C}$. We will sometimes write $\Gamma_{\mathcal{C}}(X)$ for clarity. For $S$-objects $X$ and $T$ will also use the notation

$$\Gamma_{\mathcal{C}}(X/T) = \{ f \in \text{hom}_{\mathcal{C}}(T, X) | q_X \circ f = q_T \}.$$

Often we also denote $\Gamma(X/T) = X(T)$. We will also use the canonical identifications

$$\Gamma_{\mathcal{C}/S}(X) = \Gamma_{\mathcal{C}}(X/S)$$
$$\Gamma_{\mathcal{C}/S}(X/T) = \Gamma_{\mathcal{C}/T}(X \times_S T/T)$$

When it is clear that $X$ and $T$ are $S$-objects, we can also use the notation $X(T) = \Gamma_{\mathcal{C}/S}(X/T)$, so that the last equality reads $X(T) = X_T(T)$, where we also used the common index notation $X_T$ for the pullback $X \times_S T$. Finally, we note that in the examples we will sometimes write $\Gamma(R)$ or $\Gamma(R/S)$ where $R$ and $S$ are rings, which can be ordinary, twisted, or mixed. In such cases, we always mean $\Gamma(\text{Spec } R)$ and $\Gamma(\text{Spec } R/\text{Spec } S)$.

Let us now study the interaction between rational points of ordinary schemes, twisted schemes, and mixed schemes.

To obtain something recognizable, we need to assume the base-scheme behaves well. For twisted schemes, this means that the twister must be invertible; for mixed schemes we need to assume at least one of the mixing maps is epic.

---

2If there is no terminal object, one must rely on the Yoneda embedding to provide a non-representable terminal object; but we will not need that here.
Proposition 9.2.1. Let $\tilde{S} = (S, \Phi)$ be a twisted scheme with $\Phi$ invertible. Let $\tilde{X} = (X, g)$ be an $\tilde{S}$-scheme. Define $\alpha$ as follows:

$$\alpha : X(S) \to X(S) : x \mapsto g \circ x \circ \Phi^{-1}.$$ 

Then $\alpha$ is an involution on $X(S)$ and $\tilde{X}(\tilde{S}) = X(S)^\alpha$ its set of fixed points.

Proof. We have $(\alpha \circ \alpha)(x) = (g^2) \circ x \circ (\Phi^2)^{-1} = \text{Fr}_X \circ x \circ \text{Fr}_S^{-1} = x$ and $\Gamma(\tilde{X}/\tilde{S}) = \{ x \in X(S) \mid x \circ \Phi = g \circ x \} = X(S)^\alpha$. $\square$

In other words: the set of rational points of a twisted scheme over a base with invertible twister is the set of fixed points of an involution acting on the sets of rational points of an ordinary scheme. The application that we have in mind is where $S$ is the spectrum of a perfect blended field. In Section 10.3 we will see that this proposition generalizes the construction of the Suzuki-Ree groups over perfect fields. (Over non-perfect fields, there is actually nothing to show! More on this in Remark 10.3.2.)

Lemma 9.2.2. Let $\tilde{S} = (S_1, S_2)$ be a mixed scheme and $\tilde{X} = (X_1, X_2)$ an $\tilde{S}$-scheme. If $S_1$ is reduced, then $\Phi_{S_2}$ is epic.

Note that we now suppress the maps $\Phi_{X_1}$ etc. from the notation if we can, as we remarked just before Lemma 8.1.3.

Proof. The absolute Frobenius $\text{Fr}_{S_1} : S_1 \to S_1$ is the identity on the underlying topological space, so in particular it is an epimorphism of topological spaces. Furthermore, since $S_1$ is reduced, the induced map of structure sheaves $\text{Fr}_{S_1}^\sharp : \mathcal{O}_{S_1} \to \mathcal{O}_{S_1}$, which locally just raises to the $p$'th power, is injective—i.e. a monomorphism of sheaves of rings. From these two facts, we can conclude immediately that $\text{Fr}_{S_1}$ is an epimorphism of schemes as follows. Let $X$ be an arbitrary scheme and $f, g : S \to X$ two different morphisms of schemes such that $f \circ \text{Fr}_{S_1} = g \circ \text{Fr}_{S_1}$. Since at the level of topological spaces $\text{Fr}_{S_1}$ is epic, clearly $f$ and $g$ coincide as maps of topological spaces. This implies the equality $f_* \mathcal{O}_X = g_* \mathcal{O}_X$ as sheaves of rings over $S$ and this allows us to consider the pair of maps $f^\sharp, g^\sharp : f_* \mathcal{O}_X = g_* \mathcal{O}_X \to S$. Since $\text{Fr}_{X_1}^\sharp$ was monic and $\text{Fr}_{S_1}^\sharp \circ f^\sharp = \text{Fr}_{S_1}^\sharp \circ g^\sharp$, we may conclude that
Chapter 9. Twisting and mixing schemes

\[ f^\sharp = g^\sharp \] and this proves that \( f = g \), thus \( \text{Fr}_{S_1} \) is epic. Finally, since 
\[ \text{Fr}_{S_1} = \Phi_{S_2} \circ \Phi_{S_1}, \] 
so is \( \Phi_{S_2} \). \hfill \Box

**Proposition 9.2.3.** Let \( \tilde{S} = (S_1, S_2) \) be a mixed scheme \( \tilde{X} = (X_1, X_2) \) an \( \tilde{S} \)-scheme. Assume that \( \Phi_{S_2} \) is epic.

1. The following maps are injective
\[ g : X_1(S_1) \to X_1(S_2) : v \mapsto v \circ \Phi_{S_2} \]
\[ c_1 : \tilde{X}(\tilde{S}) \to X_2(S_2) : (u, v) \mapsto u \]

2. The following square is a pullback in (set):
\[ \begin{array}{ccc}
\tilde{X}(\tilde{S}) & \xrightarrow{c_1} & X_2(S_2) \\
\downarrow & & \downarrow f \\
X_1(S_1) & \xleftarrow{g} & X_1(S_2)
\end{array} \]
where \( f : X_2(S_2) \to X_1(S_1) : u \mapsto \Phi_{X_2} \circ u \).

3. We have: \( \tilde{X}(\tilde{S}) = f^{-1}(X_1(S_1)) \).

**Proof.**

1. Injectivity of \( g \) is trivial; for \( c_1 \), if \( (u, v) \) and \( (u', v) \) are two morphisms \( \tilde{S} \to \tilde{X} \) then \( u \circ \Phi_{S_2} = \Phi_{X_2} \circ v = u' \circ \Phi_{S_2} \) and so \( u = u' \).

2. Let \( u : S_1 \to X_1 \) and \( v = S_2 \to X_2 \) be a pair of morphisms of schemes. We claim that \( (u, v) \) is a morphism of mixed schemes \( \tilde{S} \to \tilde{X} \) if (and only if) the condition \( \Phi_{X_2} \circ v = u \circ \Phi_{S_2} \) is satisfied. In other words, we show that \( v \circ \Phi_{S_1} = \Phi_{X_1} \circ u \) is a consequence. Because \( \Phi_{S_2} \) is epic, this is equivalent to showing that
\[ v \circ \Phi_{S_1} \circ \Phi_{S_2} = \Phi_{X_1} \circ u \circ \Phi_{S_2}, \]
which is clear, since the left hand side is also
\[ v \circ \text{Fr}_{S_2} = \text{Fr}_{S_2} \circ v = \Phi_{X_1} \circ \Phi_{X_2} \circ v. \]

So we have
\[ \tilde{X}(\tilde{S}) = \{(u, v) \in X_1(S_1) \times X_2(S_2) \mid \Phi_{X_2} \circ v = u \circ \Phi_{S_2}\} \]
In other words, this is the claimed fibered product in the category of sets.

3. This is a reformulation of (2) with some abuse of language which takes (1) into account. □

**Remark 9.2.4.** A few notes.

1. The proposition is purely categorical and holds not just for schemes but it was placed in this section for the good interaction with Lemma 9.2.2.

2. The notation $f^{-1}(..)$ is slightly ambiguous in our situation. Since we have maps going from $S_1$ to $S_2$ and back, there are maps $X_2(S_2) \to X_2(S_1)$ and $X_2(S_1) \to X_2(S_2)$. So in principle, one has to be specific about how $X_2(S_2)$ sits inside $X_2(S_1)$ to avoid confusion: one could mean the image of the map $X_2(S_2) \to X_2(S_1)$, but also the image of the composed map $X_2(S_2) \to X_2(S_1) \to X_2(S_2) \to X_2(S_1)$ which sits deeper in $X_2(S_1)$.

3. In a typical situation $\tilde{S} = \text{Spec } m$ could be the spectrum of a mixed field $m = (k, \ell)$. By Lemma 9.2.2 the proposition applies and in fact it applies *mutatis mutandis* to the other component. So $\tilde{X}(m)$ is a subset of both $X_1(k)$ and $X_2(\ell)$.

4. An important situation where $\Phi_{S_2}$ is epic but $S_2$ is non-reduced is where $\tilde{S}$ is a non-reduced visible scheme so that in fact $\Phi_{S_2}$ is an isomorphism. For instance, take $\tilde{S} = m(\text{Spec } k(\varepsilon))$, where $k(\varepsilon)$ are the dual numbers over a field $k$.

5. Assume one *insists* on realizing the set $\tilde{X}(\tilde{S})$ with $S_1$-objects. Then one has to find suitable $S_1$-objects such that $Y(S_1) = X_1(S_2)$ and $Z(S_1) = X_2(S_2)$ such that the induced maps are indeed $f$ and $g$. Denoting for simplicity $\beta = \Phi_{S_2} : S_2 \to S_1$, we see that $Y = \beta_* \beta^* X_1$ and $Z = \beta_* X_2$ fit the bill. This implies that if we define a new $S_1$-scheme by $U = X_1 \times_{\beta, \beta^* X_1} \beta_* X_2$, then $U(S_1) = \tilde{X}(\tilde{S})$. This suggests the statement of Proposition 8.7.1: observing that $(c_2 f_* \tilde{X})(S) = \tilde{X}(f^* m)S$ gives away the second component of $f_* \tilde{X}$. A similar reasoning, starting from Remark 8.2.4.(iii) could perhaps lead to a generalization of that proposition for arbitrary morphisms $(\alpha, \beta) : (T, T') \to (S, S')$. 
9.3 Modules and sheaves

We want to introduce a notion of modules over twisted rings, and more generally sheaves of modules over twisted schemes. The main purpose is to define partial dimensions of a mixed scheme. The underlying idea is that if we have a scheme over, say, a mixed field \((k, \ell)\) we want to measure how much of it is defined over \(k\) and how much is defined over \(\ell\). In Chapter 10 we will construct certain mixed reductive groups; these are roughly speaking reductive groups where the long and short roots live over different halves of a mixed field. In Remark 10.4.2 we will explain how these partial dimensions count the number of dimensions determined by short and long roots.

Let us first recall the notion of a \(p\)-structure (also called Frobenius structure) before considering its ‘square root’.

**Definition 9.3.1.** A \(p\)-structure on a module \(M\) over a ring \(R\) of characteristic \(p\) is a map \(M \to M : x \mapsto x^{(p)}\) such that \(a^p x^{(p)} = (ax)^{(p)}\) and \((x + y)^{(p)} = x^{(p)} + y^{(p)}\), for \(a \in R\) and \(x, y \in M\).

Note that in particular, \(0 : x \mapsto 0\) is always a \(p\)-structure.

**Definition 9.3.2.** Let \(\tilde{R} = (R, f)\) be a twisted ring. Then a twisted \(\tilde{R}\)-module \(M\) is an \(R\)-module together with a map \(\psi : M \to f^* M\), where \(f^* M\) is \(M\), considered as an \(R\)-module through the multiplication \(a \circ x = f(a)x\).

In other words, it is a map \(\psi : M \to M\) which is semi-linear in the following sense:

\[\psi(ax + y) = a \circ \psi(x) + \psi(y) = f(a)\psi(x) + \psi(y)\]

Next we globalize these notions to obtain those of twisted sheaves.

**Definition 9.3.3.** A \(p\)-structure on a sheaf \(\mathcal{F}\) of modules over a scheme \((X, \mathcal{O}_X)\) with absolute Frobenius \(\text{Fr}\) is a is a morphism \(\langle p \rangle : \mathcal{F} \to \text{Fr}_* \mathcal{F}\).

**Definition 9.3.4.** A twisted sheaf of modules \((\mathcal{F}, \psi)\) over a twisted scheme \((X, \Phi)\) is a sheaf \(\mathcal{F}\) on \(X\) together with a morphism \(\psi : \mathcal{F} \to \Phi_* \mathcal{F}\).
For the convenience of the reader, let us break down this definition. For every open $U \subseteq |X|$, denote $\varphi(U) = \varphi^{-1}(U) = \overline{U}$. Then the twisted structure defines a morphism

$$\psi_U : \mathcal{F}(U) \to \mathcal{F}(\overline{U}),$$

of $\mathcal{O}_X(U)$-modules, where $\mathcal{F}(\overline{U})$ is considered an $\mathcal{O}_X(U)$-module through the map $\varphi^\sharp : \mathcal{O}_X(U) \to \mathcal{O}_X(\overline{U})$, and for every inclusion $V \subseteq U$, the obvious diagram commutes. For future reference, let us draw attention to the case of a mixed field.

**Definition 9.3.5.** A mixed vector space over a mixed field $m = (k, \ell, \kappa, \lambda)$ is a tuple $V_m = (V_k, V_\ell, \widehat{\kappa}, \widehat{\lambda})$, consisting of a $k$-vectorspace $V_k$, an $\ell$-vectorspace $V_\ell$ and a pair of semi-linear maps $\widehat{\kappa} : V_k \to V_\ell$, $\widehat{\lambda} : V_\ell \to V_k$, where semi-linear means that $\widehat{\kappa}(ax + y) = \kappa(a)\widehat{\kappa}(x) + \widehat{\kappa}(y)$ whenever $a \in k$, $x, y \in V_k$ and vice versa for $\lambda$. The partial dimensions are given by

$$\text{pardim } V_m = \left( \dim_k (V_k / \ker \widehat{\kappa}), \dim_\ell (V_\ell / \ker \widehat{\lambda}) \right)$$

An important class of twisted modules comes from twisted ideals:

**Definition 9.3.6.** A twisted ideal $a$ of a twisted ring $(R, f)$ is an ideal $a \subseteq R$ such that $f(a) \subseteq a$. A twisted sheaf $(\mathcal{J}, \psi)$ on a twisted scheme $(X, \psi)$ is a subsheaf of $\mathcal{O}_X$ such that the inclusion $\mathcal{J} \to \mathcal{O}_X$ respects twisters.

In other words, the twisted ideals are precisely those ideals for which $f$ induces a twisted structure on $R/a$ and a twisted sheaf of ideals provides the structure of a twisted scheme on the corresponding closed subscheme. In particular, observe that the structure sheaf $\mathcal{O}_X$ on a twisted scheme is a twisted sheaf, the twister being $\varphi^\sharp : \mathcal{O}_X \to \varphi_* \mathcal{O}_X$.

**Proposition 9.3.7.** The sheaf of differentials $\Omega_{X/S}$ on a twisted scheme $X$ over a base twisted scheme $S$ is canonically endowed with the structure of a twisted sheaf.

**Proof.** Taking differentials of the commutative square $\Phi_S \circ q_X = q_X \circ \Phi_X$ results in a map $d\Phi : \Omega_{X/S} \to (\Phi_X)_* \Omega_{X/S}$ as in [Stacks, Tag 01UV].
Remark 9.3.8. Originally, we had planned to define a twisted version of the tangent bundle $T X$. The ambition was to endow the tangent space $T_e G = T G \times_G S$ over the neutral element $e : S \to G$ of a mixed $S$-group scheme $G$ (see Definition 10.0.1) with the structure of a (to be defined) mixed Lie algebra $(L_1, L_2, \psi_1, \psi_2)$ which, in the visible case, would boil down to the Lie algebra endowed with the $p$-operation $(L, L, [p], 0)$.

Unfortunately, we couldn’t make this idea work. The problem is that the spaces of homomorphisms between twisted $\tilde{R}$-modules are not themselves twisted $\tilde{R}$-modules, and in particular the dual $M^\vee = \text{hom}(M, \tilde{R})$ of a twisted module $(M, \psi)$ is not itself a twisted module in a canonical way—unless $\psi$ is bijective. In other words, we lack internal hom objects and because of this, there is no obvious way to dualize the twisted cotangent bundle $(\Omega_X/S, d\Phi)$. We want to share several ideas which may contribute to resolve this difficulty, but so far we could not make any subset of them work to our satisfaction.

- The hom-sets are canonically endowed with a $p$-structure, perhaps we should accept that for a twisted object $X \in \text{Ob}(t\mathcal{C})$ its tangent bundle is an ordinary object $T X \in \text{Ob}(\mathcal{C})$.
- We can just accept that the cotangent sheaf is the fundamental object, and define a mixed Lie co-algebra directly without dualizing.
- The definition from [SGA3, exp. II, 3.1] relies on the inner hom in the category $\widehat{(\text{sch})} = \text{hom}((\text{sch})^{\text{op}}, (\text{set}))$. This works here too, but then there is no guarantee that the resulting functor is representable.
- If $\varepsilon : S[\varepsilon] \to S$ is a first order infinitesimal extension with section $\delta : S \to S[\varepsilon]$ then the tangent bundle can be defined as $T X = \varepsilon_* \varepsilon^* X$. Since the problem lies with the functor $\varepsilon_*$, we could define $T X = \varepsilon^* X$ instead. Instead of an $S$-linear diagram $X \to T X = \varepsilon_* \varepsilon^* X \to X$, we would obtain a diagram

$$
\begin{array}{cccccc}
X & \simeq & \delta^* \varepsilon^* X & \longrightarrow & T X = \varepsilon^* X & \longrightarrow & X \\
\downarrow & & \downarrow & & \downarrow & & \downarrow \\
S & \overset{\delta}{\longrightarrow} & S[\varepsilon] & \overset{\varepsilon}{\longrightarrow} & S
\end{array}
$$
So perhaps we should take such non-linearity for granted.

- Perhaps we should focus on mixed formal groups and ignore the Lie algebra altogether.

A working definition of such a mixed Lie algebra would be of great interest, because it could lead the way to a construction of algebraic objects, such that their automorphism groups are precisely the mixed groups that we will introduce in Chapter 10. (I.e. mixed versions of quadratic spaces, octonion algebras and Albert algebras.)

**Remark 9.3.9.** We have restricted ourselves to the study of what is strictly necessary for Section 10.4 and Remark 10.4.2 in particular, where we study some groups that we find particularly interesting. Nonetheless, it could be fruitful to redefine some of the common adjectives from algebraic geometry (connected, smooth) in the mixed setting. This could shed light on natural questions—consider a mixed affine scheme over a mixed field \((k, \ell)\); when do the partial dimensions add up to the dimension of each of the components?—and allow for an easy generalization of theorems that have been proven in a sufficiently ‘generic’—e.g. not relying on arguments about rational points over an algebraic closure—manner.

### 9.4 Examples

As before all rings are assumed to be of characteristic \(p\).

**Example 9.4.1.** Twisted and mixed rings.

1. A pair \(b = (k, \psi)\) where \(\psi^2(x) = x^p\) is a blended field. These are also known as fields with *Tits endomorphism* [DSW] or (for \(p = 2\)) *octogonal sets* [TW02, (10.11)].

2. Consider fields \(k, \ell\) such that \(k^p \subseteq \ell \subseteq k\). Then

\[
m = (k, \ell, \text{inc}_{k^p \to \ell} \circ \text{fr}_k, \text{inc}_{\ell \to k})
\]

is a mixed field; we leave it to the reader to verify that in fact every mixed field is of this form. As an \((\mathbb{F}_p \times \mathbb{F}_p)\)-algebra this is the twisted ring \((k \times \ell, (x, y) \mapsto (y, x^p))\) with the obvious structural morphism. The mixed field is visible in the extreme
Chapter 9. Twisting and mixing schemes

3. The pair $R = (\mathbb{F}_p[x,y], \varphi)$, where $\varphi : f(x,y) \mapsto f(y,x^p)$ is a twisted ring. It is not mixed or blended. Taking the tensor product with a blended field $b = (k, \psi)$ gives the twisted ring (and $b$-algebra)

$$R_b = R \otimes b = (k[x,y], f(x,y) \mapsto f^\psi(y,x^p)),$$

where $f^\psi$ means: apply $\psi$ to the coefficients of $f$. Taking the tensor product with a mixed field $m = (k, \ell, \kappa, \lambda)$ instead, we get a mixed ring (and $m$-algebra)

$$R_m = R \otimes m = (k[x,y], \ell[x,y], \hat{\kappa}, \hat{\lambda}),$$

where $\hat{\kappa} : f(x,y) \mapsto f^\kappa(y,x^p)$ and $\hat{\lambda} : f(x,y) \mapsto f^\lambda(y,x^p)$.

4. For every ring $K$, the mixing functor $m$ defines the visible mixed ring $m(K) = (K,K,\text{id},\text{id})$. For instance, $m(\mathbb{F}_p) = e$. If the Frobenius map $\text{fr}_K : K \to K : x \mapsto x^p$ is injective, and in particular if $K$ is reduced, then $m(K)$ is isomorphic to the mixed ring $(K^p, K, \text{inc}, \text{fr})$, the isomorphism being given by $(\text{fr}, \text{id})$. We may identify a mixed ring $M = (K,L,\kappa,\lambda)$ with a certain ring extension $K' \subseteq L' \subseteq K$ as we did for fields in item 2, provided that $\lambda$ is injective by taking $K' = K$ and $L' = \text{im} \lambda$. Recall from Remark 8.2.4 that a mixed ring $M = (K,L,\kappa,\lambda)$ is visible precisely when $\lambda$ is an isomorphism; this corresponds to the extreme situation where $L' = K'$.

5. If $b = (k, \psi)$ is a blended field then

$$m = b \otimes e = (k,k,\psi,\psi)$$

is a mixed field and the diagonal $\chi : b \to m : a \mapsto (a,a)$ provides an embedding of twisted rings. This gives in turn rise to a functor of base change $\chi^* : b_{-\text{alg}} \to m_{-\text{alg}} : A \mapsto A \otimes_b m$, which corresponds to the twixting functor introduced in Section 8.2, but taking slices over the object $b$.

6. If $M = (K,L,\kappa,\lambda)$ is a mixed ring then any $a \in L$ defines the new mixed ring $M' = (K[x]/(x^p - a^\lambda), L, \hat{\kappa}, \hat{\lambda})$ given by

$$\begin{cases}
\hat{\kappa} : f(x) \mapsto f^\kappa(a), \\
\hat{\lambda} : u \mapsto \lambda(u).
\end{cases}$$
We can denote this construction by $M' = (K(a^{\kappa^{-1}}), L)$; the same construction can be carried out with a family of elements $a_i \in L$ $i \in I$ and the ring $K[x_i, i \in I]/(x^p - a_i^\lambda | i \in I)$.

7. Consider field extensions $k/\ell$ such that $k^p \subseteq \ell$, and a pair of field extensions (or more generally étale algebras) $K/k$ and $L/\ell$ such that $K^p \subseteq L$. Then $(K, L)/(k, \ell)$ is an extension of mixed fields; moreover it is clear that every mixed field extension arises this way.

**Example 9.4.2.** Rational points of some of these examples:

1. It is clear that $\Gamma(b/b)$ is a singleton. Typically, one considers a field as the terminal object of its own slice category. This convention justifies the idea that $\Gamma(b)$ is a singleton.
2. Here too $\Gamma(m)$, an abuse of notation for $\Gamma(m/m)$, is a singleton.
3. Let us first compute $\Gamma(R_b/b)$.

\[
\Gamma(R_b/b) = \text{hom}_{\text{tsch}}(\text{Spec } b, \text{Spec } R_b) = \text{hom}_{\text{ring}}(R_b, b), = \{\alpha \in \text{hom}_{\text{ring}}(k[x,y], k) | \alpha \circ \varphi = \psi \circ \varphi\}.
\]

Any such $\alpha$ is fully specified by $\alpha(x) = x_0$ and $\alpha(y) = y_0$ and the condition says that $y_0 = \psi(x_0)$. So as a set, there is an identification $\Gamma(R_b/b) \cong k$.

Similarly, it is easy to verify directly that there is an identification $\Gamma(R_m/m) \cong m$, but let us verify this again with Proposition 9.2.3, using the notations $S_1 = \text{Spec } k$, $S_2 = \text{Spec } \ell$, $X_1 = \text{Spec } k[x,y] \cong \mathbb{A}^2_k$ and $X_2 = \text{Spec } \ell[x,y] \cong \mathbb{A}^2_\ell$. Then the proposition tells us that there is an identification of $\Gamma(R_m/m) = \tilde{X}(\tilde{S})$ and $f^{-1}(X_1(S_1))$, where $f$ is the induced map $f : X_2(S_2) \to X_1(S_2)$ and $X_1(S_1) \subseteq X_1(S_2)$ in the natural manner. In our case, there is a natural identification of $X_1(S_1)$ with $k \times k$ on the one hand and $X_2(S_2)$ and $X_1(S_2)$ with $\ell \times \ell$ on the other hand; moreover the inclusion $X_1(S_1) \hookrightarrow X_1(S_2)$ corresponds to the inclusion $k \times k \hookrightarrow \ell \times \ell$ and the map $f$ can be identified with $\ell \times \ell : (a, b) \mapsto (b^p, a)$.

Therefore $f^{-1}(X_1(S_1))$ corresponds to the subset of all $(a, b) \in$
\( \ell \times \ell \) such that \( b^p \in k \) and \( a \in k \), where the first condition is of course always satisfied since \( \ell^p \subseteq k \). So this corresponds precisely to the set \( k \times \ell \) and the inclusion \( \tilde{X}(S) \hookrightarrow X_2(S_2) \) is just the natural inclusion \( k \times \ell \hookrightarrow \ell \times \ell : (u, v) \mapsto (\kappa(u), v) \).

4. By fullness of the mixing functor \( m \), we have that

\[
\Gamma_{\text{mring}}(m(K)/m(k)) \simeq \Gamma_{\text{ring}}(K/k).
\]

The next examples concern group functors and group schemes. Because we have not yet introduced mixed group schemes (see Definition 10.0.1) we will focus on understanding the schemes and their sets of rational points and not emphasize the group structure.

**Example 9.4.3.** Mixing tori.

Let \( k \) be a field of characteristic 2 and let \( K = k[x]/(x^2 + x + \delta) = k(u) \) be a separable field extension of \( k \) of degree 2 with Galois group \( \langle \sigma \rangle \), where \( u^2 = u + 1 \). Whenever \( R \) is a \( k \)-algebra, there is an isomorphism of \( R \)-modules

\[
R_K = R \otimes_k K \cong R[x]/(x^2 + x + \delta) \cong R \oplus Ru,
\]

where \( u \) is just a formal symbol which satisfies the multiplication rule \( u^2 = \delta + u \). This means that we can freely think of \( R_K \) as the \( k \)-module \( R \oplus R \) endowed with a multiplication whenever this suits us. The involution on \( K \) extends to an involution on \( R_K \) given by

\[
R_K \rightarrow R_K : a + ub \mapsto (a + b) + ub.
\]

We will also think of \( R \) as a subset of \( R_K \) via the inclusion

\[
R \hookrightarrow R_K : x \mapsto x \otimes 1 = x + 0u.
\]

Let us now consider the following functors from the category of \( k \)-algebras to the category of sets; equivalently, these are presheaves on the category of affine schemes:

\[
\begin{align*}
\text{GL}_1 : R &\mapsto R^* \\
\text{GL}_1^\sigma : R &\mapsto \{ x \in R_K \mid xx^\sigma = 1 \} \\
T_2 : R &\mapsto R_K^*.
\end{align*}
\]
Our goal is to construct a mixed scheme for which the components correspond to $T_1 = \text{GL}_1 \times \text{GL}_1$ and $T_2$; the mixed scheme should then be thought of as some mixture of $T_1$ and $T_2$.

Let us start by defining a pair of $k$-morphisms between $T_1$ and $T_2$. In order to do so, it suffices to provide the components of these morphisms, considered as a natural transformation of functors:

\[
\begin{align*}
    f_R &: T_1(R) \to T_2(R) : (x, v) \mapsto xv \\
    g_R &: T_2(R) \to T_1(R) : y \mapsto (y^\sigma, y/y^\sigma)
\end{align*}
\]

To be completely precise, we must still verify that these indeed specify natural transformations. This means that for an arbitrary morphism $h : R \to R'$ of $k$-algebras, the corresponding diagram must commute.

So we must check that $h_K(x)h_K(v) = h_K(xv)$, where $h_K$ is the map $R_K \to R'_K : x + uy \mapsto h(x) + uh(y)$. This is clearly the case because $h$ is an algebra homomorphism, and we may verify the other square similarly.

The important point here is that these maps compose to the squaring operators as we will now verify. First, we need to compute the composition

\[
(\begin{array}{c}
    g_R \circ f_R : (x, v) \mapsto xv \mapsto (xx^\sigma vv^\sigma, xv/(xv)^\sigma) = (x^2, v^2)
\end{array})
\]

where we have used that $x^\sigma = x$ and $v^\sigma = v^{-1}$. For the other composition, we compute

\[
(\begin{array}{c}
    f_R \circ g_R : y \mapsto (yy^\sigma, y/y^\sigma) \mapsto y^2.
\end{array})
\]

We need one more ingredient, which is an isomorphism

\[
\vartheta : T_1 \to \Delta T_1.
\]
Chapter 9. Twisting and mixing schemes

Since the elements of the set $T_1(R)$, where $R$ is an arbitrary $k$-algebra, can be identified with the quadruples $(x', y', x, y) \in R^4$ for which $x'y' = 1$ and $x^2 + \delta y^2 + xy = 1$, and similarly for $\Delta T_1(R)$, except that we must replace $\delta$ by $\delta^2$, the map $\vartheta$ is completely determined by the following maps, one for every $k$-algebra:

$$\vartheta_R : T_1(R) \to \Delta T_1(R)$$
$$: (x', y', x, y) \mapsto (x', y', x + \delta y, y).$$

Note that these maps are well defined because

$$(x + \delta y)^2 + \delta^2 y^2 + (x + \delta y)y = x^2 + \delta y^2 + xy = 1$$

and functoriality is again trivial. So we obtain the following diagram:

$$T_1 \xrightarrow{f} T_2 \xrightarrow{\vartheta \circ g} \Delta T_1 \xrightarrow{\Delta f} \Delta T_2$$

It is now straightforward to verify that this determines a relative factorization, as defined in Definition 8.5.2, by directly computing the compositions $\theta \circ g \circ f$ and $\Delta f \circ \theta \circ g$. This means that we can use Corollary 8.5.5 to construct a mixed scheme, if we also have an absolute factorization of the base, i.e. a field extension $\ell/k$ for which $\ell^2 \subseteq k$. Corollary 8.5.5 then provides us with the following mixed scheme $\tilde{T}$ over the mixed field $m = (k, \ell)$:

$$\xymatrix{ T_2 & (T_1)_\ell \ar[l] \ar[r] & \Spec k & \Spec \ell \ar[l] }$$

Let us use Proposition 9.2.3 to compute the rational points $\tilde{T}(m)$ of this mixed scheme as a the pullback in (set) determined by the following diagram.

$$\xymatrix{ \tilde{T}(m) & T_2(\ell) \ar[l] \ar[r] & T_1(\ell) }$$
The bottom arrow in this diagram is just the functor $T_1$ applied to the inclusion $k \hookrightarrow \ell$. The right arrow is the map

$$g_\ell : T_2(\ell) \to T_1(\ell) : y \mapsto (yy^\sigma, y/y^\sigma).$$

Now we observe that $y/y^\sigma = yy^\sigma/(y^2)^\sigma$ and $y^2 \in k$ whenever $y \in \ell$. So rather than require that both $yy^\sigma \in k^\times$ and $y/y^\sigma \in K^\times$, one of these is sufficient. Furthermore, recall that $y \mapsto yy^\sigma$ is just the norm $N_{L/\ell} : L^\times \to \ell^\times : y \mapsto yy^\sigma$, where $L = K \otimes_k \ell$. With this information, we finally come to the following description of the set of $m$-rational points:

$$\tilde{T}(m) = \{ y \in T_2(\ell) \mid yy^\sigma, y/y^\sigma \in k \} = \{ y \in L^\times \mid N_{L/\ell}(y) \in k \}.$$

**Remark 9.4.4.** Let us verify some of these assertions with some computer algebra.

First note that these functors are represented by the following $k$-algebras:

- $O(GL_1) = k[x', y']/(x'y' - 1)$
- $O(GL_1^\sigma) = k[x, y]/(x^2 + xy + y^2\delta + 1)$
- $O(T_2) = k[x_1, x_2, y_1, y_2]/(x_1x_2 + \delta y_1y_2 + 1, (x_1 + y_1)(x_2 + y_2) + x_1x_2).$

Recall that the coordinate algebra of an affine scheme $\Delta(\text{Spec} A)$, where $A = k[x_1, \ldots, x_n]/(f_1, \ldots, f_m)$, is given by

$$A^{(p)} = k[x_1, \ldots, x_n]/(f_1^{(p)}, \ldots, f_m^{(p)}),$$

where $f_i^{(p)}$ is the polynomial $f_i$, with the $p$’th power map applied to each of its coefficients. In our case, this means that we can obtain the coordinate algebras of $\Delta T_1$ and $\Delta T_2$ by replacing every occurrence of $\delta$ by $\delta^2$.

By the Yoneda lemma, the $k$-morphisms $f$ and $g$ correspond to morphisms between these coordinate algebras, which are given by:

$$f^\sharp : O(T_2) \to O(T_1) : \begin{cases} 
  x_1 \mapsto x'x \\
  x_2 \mapsto y'x + y'y \\
  y_1 \mapsto x'y \\
  y_2 \mapsto y'y 
\end{cases}$$
$g^{\sharp} : \mathcal{O}(T_1) \rightarrow \mathcal{O}(T_2) :$

\[
\begin{align*}
x' & \mapsto x_1^2 + \delta y_1^2 + x_1 y_1 \\
y' & \mapsto x_2^2 + \delta y_2^2 + x_2 y_2 \\
x & \mapsto (x_1^2 + \delta y_1^2) \cdot (x_2^2 + \delta y_2^2 + x_2 y_2) \\
y & \mapsto y_1^2 \cdot (x_2^2 + \delta y_2^2 + x_2 y_2)
\end{align*}
\]

Finally, the map $\vartheta^{\sharp}$ is given by:

$\vartheta^{\sharp} : \mathcal{O}(T_1)^{(2)} \rightarrow \mathcal{O}(T_1) :$

\[
\begin{align*}
x' & \mapsto x' \\
y' & \mapsto y' \\
x & \mapsto x + \delta y \\
y & \mapsto y
\end{align*}
\]

To verify that we indeed obtain a relative factorization, we must check that in the corresponding diagram of $k$-algebra homomorphisms

$$
\mathcal{O}(T_2)^{(2)} \xrightarrow{(f^{\sharp})^{(2)}} \mathcal{O}(T_1)^{(2)} \xrightarrow{g^{\sharp} \circ \vartheta^{\sharp}} \mathcal{O}(T_2) \xrightarrow{f^{\sharp}} \mathcal{O}(T_1)
$$

the compositions $f^{\sharp} \circ g^{\sharp} \circ \vartheta^{\sharp}$ and $g^{\sharp} \circ \vartheta^{\sharp} \circ (f^{\sharp})^{(2)}$ are equal to the relative Frobenii. Let us use SAGE$^3$ to verify this.

In the following code, the variable $t$ plays the role of $\delta$ and is treated as a formal indeterminate over a field of characteristic 2. $^4$ We then define all rings as quotient rings of a polynomial ring $A_4$ in the 4 variables $v_1, w_1, v_2$ and $w_2$. We use $xx$ and $yy$ to denote $x'$ and $y'$, while the suffix $_p$ should be read as $(\ )^{(p)}$.

```sage
1 temp.<t>/uni2423=/uni2423PolynomialRing(GF(2))
2 k = FractionField(temp)
3 A4.<v1,w1,v2,w2> = PolynomialRing(k)
4 T2.<x1,y1,x2,y2> = A4.quotient_ring(
5 [v1*v2+t*w1*w2+1,(v1+w1)*(v2+w2)+v1*v2])
6 T1.<x,y,xx,yy> = A4.quotient_ring(
7 [v2*w2 + 1, v1^2+v1*w1+w1^2*t+1])
8 T2p.<x1p,y1p,x2p,y2p> = A4.quotient_ring(
```

$^3$Go to https://sagecell.sagemath.org/ for an easy to use web interface.

$^4$The reader may be surprised to see that we do not need the assumption that $x^2 + x + \delta$ is irreducible. This is because it is indeed irrelevant! All that is relevant is that $K$ is a separable $k$-algebra, but it does not have to be a field.
9.4. Examples

Next, we need to define all homomorphisms between these rings by specifying the images of the generators.

Executing this code will not only define the homomorphisms, but also verify that they are actually well defined. (For instance, try changing one of the $t$’s to $t^2$ and see what happens!) Finally, we need to check that the maps $f^\# \circ g^\# \circ \vartheta$ and $g^\# \circ \vartheta \circ (f^\#)^{(2)}$ are equal to the relative Frobenii in each case, i.e. that they square each of the generators. Sage will automatically coerce all variables to the appropriate quotient ring before applying a map to it and when testing equality, so we can just type the following code:

```python
for var in [v1, w1, v2, w2]:
    assert f(g(theta(var)))) == var^2
    assert g(theta(fp(var)))) == var^2
print("OK")
```

This will print OK if Sage could verify these identities, and throw an AssertionError otherwise.

On a sidenote, the fact that the compositions are equal to the squaring operators amounts to verifying the following compositions:

$$f^\# \circ g^\# : \mathcal{O}(T_1) \rightarrow \mathcal{O}(T_1) : \begin{cases}
x' \mapsto x'^2 \\
y' \mapsto y'^2 \\
x \mapsto x^2 + \delta y^2 \\
y \mapsto y^2
\end{cases}$$
$g^* \circ f^* : \mathcal{O}(T_2) \to \mathcal{O}(T_2) : \begin{cases} x_1 \mapsto x_1^2 + \delta y_1^2 \\ x_2 \mapsto x_2^2 + \delta y_2^2 \\ y_1 \mapsto y_1^2 \\ y_2 \mapsto y_2^2 \end{cases}$

We will leave it to the reader to verify this with Sage by computing for instance $f(g(x)) - x^2 - ty^2$.

**Example 9.4.5.** Mixing an adjoint with a simply connected group.

Consider the $k$-group functors which correspond to the adjoint and simply connected split groups of type $A_{p-1}$, i.e.

$$
\begin{align*}
\text{SL}_p &= A_{p-1}^{\text{sc}} : K \rightsquigarrow \text{SL}_p(K) \\
\text{PGL}_p &= A_{p-1}^{\text{ad}} : K \rightsquigarrow \text{PGL}_p(K),
\end{align*}
$$

where $K$ denotes an arbitrary $k$-algebra. Recall that the set $\text{SL}_p(K)$ denotes the set of matrices $(x_{ij}) \in K^{p \times p}$ with determinant $\det(x_{ij}) = 1$. Recall also that set $\text{PGL}_p(K)$ denotes the set of invertible matrices $[x_{ij}] \in K^{p \times p}$ up to an equivalence relation, which expresses that $[x_{ij}] = [y_{ij}]$ if for every $n \geq 0$ and every homogeneous polynomial $f$ of degree $np$ the relation

$$
\frac{f(x_{ij})}{\det(x_{ij})^n} = \frac{f(y_{ij})}{\det(y_{ij})^n}
$$

holds. (In particular $[x_{ij}] = [\lambda x_{ij}]$ for every $\lambda \in K^\times$; if $K$ is a field this completely determines $\text{PGL}_p(K)$.) These functors are representative and therefore define affine $k$-group schemes. We can now define a morphism between these group schemes by defining natural transformations between the corresponding functors:

$$
\begin{align*}
\text{SL}_p(K) &\longrightarrow \text{PGL}_p(K) \longrightarrow \text{Fr}^* \text{SL}_p(K) \\
(x_{ij}) &\longmapsto [x_{ij}] \longmapsto (x_{ij}^p)/\det(x_{ij})
\end{align*}
$$

It is not trivial that the second map is well defined. To see this we must observe that when $[x_{ij}] = [y_{ij}]$ then $(x_{ij}^p)/\det(x_{ij}) = (y_{ij}^p)/\det(y_{ij})$ by definition, and also that

$$
\det \left( \frac{x_{ij}^p}{\det(x_{ij})} \right) = \det(x_{ij})^{-p} \det(x_{ij}^p) = 1,
$$
9.4. Examples

so that the image of \([x_{ij}]\) really has determinant = 1.

These maps now define a relative factorization, as defined in Definition 8.5.2. So if we consider an absolute factorization of the base, i.e. a field extension \(\ell/k\) such that \(\ell^p \subseteq k\), then we can apply Corollary 8.5.5, which provides the following mixed scheme \(\tilde{G}\) over the mixed field \(m = (k, \ell)\):

\[
\begin{align*}
(SL_p)_k & \longrightarrow (PGL_p)_\ell \\
\downarrow & \quad \downarrow \\
\text{Spec } k & \longrightarrow \text{Spec } \ell.
\end{align*}
\]

Let us use Proposition 9.2.3 to compute \(\tilde{G}(m)\) as a pullback in (set) determined by the following diagram.

\[
\begin{align*}
\tilde{G}(m) & \longrightarrow SL_p(\ell) \\
\downarrow & \quad \downarrow \\
PGL_p(k) & \longrightarrow PGL_p(\ell),
\end{align*}
\]

This means that we must compute the inverse image of \(PGL_p(k)\) under the map \(SL_p(\ell) \rightarrow PGL_p(\ell)\). These are just the \(p \times p\)-matrices \((x_{ij}) \in SL_p(\ell)\) such that \([x_{ij}] = [y_{ij}]\) with all \(y_{ij} \in k\), in other words \(\lambda x_{ij} \in k\) for some \(\lambda \in \ell^\times\). Clearly, for this to hold it is necessary and sufficient that \(f(x_{ij}) \in \ell\) holds for all monomials of degree \(p\). In other words, the rational points are described by the ordinary equation \(\det(x_{ij}) = 1\) together with the ‘mixing equations’ \(f(x_{ij}) \in \ell\), where \(f\) runs through the monomials in the \((x_{ij})\) of degree \(p\).

Remark 9.4.6. The situation in Example 9.4.5 occurs because the kernel \(\mu_p \leq SL_p\) from the short exact sequence

\[
1 \rightarrow \mu_p \rightarrow SL_p \rightarrow PGL_p \rightarrow 1
\]

is an infinitesimal subgroup scheme of height 1, i.e. \(\mu_p\) is contained in the kernel of the relative Frobenius on \(SL_p\). By [CGP15, A.7.14] this can be traced back to the fact that the restricted Lie algebra \(sl_p = \text{Lie}(SL_p)\) has an ideal \(z\) of scalar matrices; see also [Bor91, 17.5.2].
So it is the failure of the isomorphism between $\mathfrak{pgl}_p = \mathfrak{gl}_p/\mathfrak{z}$ and $\mathfrak{sl}_p$ in characteristic $p$ which causes this behaviour and this ultimately comes down to the observation that the identity matrix has trace 0.

More generally, whenever we have an algebraic group with an infinitesimal normal subgroup scheme of height 1, such as

\[ 1 \to \mu_\ell \to \text{SL}_p \times \text{GL}_1 \to \text{GL}_p \to 1, \]

we can use this to construct a mixed scheme in the same manner. This mixed scheme is then actually a mixed group scheme and we can apply a Weil restriction from the mixed field $m = (k, \ell)$ back to the ordinary field $k$ to obtain a $k$-group scheme. It would be interesting to see whether the resulting $k$-group schemes are necessarily pseudo-reductive if both components are; for instance Example 9.4.5 should come down to the construction in [Tit92, §4]. It would also be interesting to characterize the class of pseudo-reductive groups that can be obtained in this fashion. Note that Theorem 10.5.1 will state precisely that the exotic pseudo-reductive groups arise in this fashion, by using a very special isogeny as starting point, see Section 10.5 for details.
Twisting and mixing groups

“Ah—so the Suzuki-Ree groups are not $F$-points of an algebraic group over $F$ after all then, I guess. I guess that the endomorphism $\varphi$ cannot be used to define descent data as required. So weird…”

Marty Weissman$^1$

We will now explain how various classes of exotic groups can be integrated into our theory. Continuing our discussion about twisted and mixed schemes, we will define twisted and mixed group schemes by recycling Definitions 2.1.1. and 2.1.2. from [SGA3]:

**Definition 10.0.1.** A twisted group scheme is a group object in $(\text{tsch})$; a mixed group scheme is a group object in $(\text{msch})$.

10.1 Informal statement of the theorems

Let us first *informally* state our main theorems and provide some context.

**Theorem.** All twisted abstract groups arise as rational points of twisted group schemes.

This is an informal statement because the notion of a twisted abstract group is not well defined in the literature. Rather, there is a list of examples which are referred to as twisted Chevalley groups. We attempt to tell the full story in Appendix A, let us now just sketch an overview of known twisted groups, a list which includes

$^1$https://mathoverflow.net/questions/33842

179
in the perfect and in particular the finite case, the Suzuki [Suz60] and Ree [Ree61b; Ree61c] groups. Although more ad-hoc constructions have been found by Wilson [Wil13], the construction by Ree and the exposition by Carter [Car72] remain the golden standard. With much of the research being focussed on the finite case, it is less widely known that this construction was extended to imperfect fields by Tits [Tit61]. All subsequent research into these groups over imperfect fields is closely related to the theory of Moufang buildings or Moufang sets (a substitute for the Moufang buildings of rank 1). To be precise: $^2B_2$ and $^2G_2$ admit a split BN-pair of rank 1, so correspond to Moufang sets; and $^2F_4$ admits a BN-pair of rank 2 and appears in the classification of Moufang polygons [TW02, (41.21.v)] under the guise of the Moufang octagons—although there they are grouped together with the mixed groups. As far as forms of these twisted groups are concerned, we think that forms of $^2F_4$ of relative rank 1 (i.e. Moufang sets) have been investigated in [MM06]; the main result of [DSW] states in some sense that twisted descent commutes with Galois descent. Of anisotropic forms of $^2B_2$, $^2G_2$ and $^2F_4$, there is no trace in the literature, presumably because there is no geometric structure attached to them.

Because of all this, we interpret this theorem as speaking only about split twisted groups over perfect fields. See Theorem 10.3.1 for a precise statement and proof.

**Theorem.** All mixed abstract groups arise as rational points of mixed group schemes.

Our main reference here is Tits’s 1974 lecture notes on buildings [Tit74, (10.3.2)], where he constructs classes of abstract groups that we think of as split mixed groups as also suggested by [TW02]; this construction works over an arbitrary field of the appropriate characteristic, although it becomes more interesting if the field is not perfect. These groups are adjoint groups; other isogeny types are not explicitly mentioned.

For forms of these groups, however, the literature is quite confusing. In [TW02, Ch. 41] the buildings related to forms of mixed BC$_n$ are swept under the rug of the classical buildings so they don’t show
up directly. Somewhat further in [TW02, (41.20)], Weiss lists the Moufang spherical buildings which are neither classical nor algebraic and it is suggested that they are associated to “\((K,k)\)-forms” of these split mixed groups but these are not further defined. In fact, the main reason Weiss considers these forms is his discovery of an exotic class of Moufang quadrangles which go by the name \(\text{mixed quadrangles of type } F_4\). The situation is further confused since the octagons (related to twisted groups) are on the list. Finally, in [CD15], a certain Moufang set is constructed which we suspect to be a form of mixed \(F_4\) which arises by mixing together a split \(F_4\) and one of relative rank 1.

Because of all this, we interpret these theorems as speaking only about the split mixed groups. See Theorem 10.4.1 for a precise statement and proof.

In [CGP15, (Ch. 7)] Tits’s constructions are revisited by Conrad, Gabber and Prasad using an alternative approach which is ‘well suited to working with arbitrary \(k\)-forms’, but the context is not that of groups related to buildings. Rather, the subject of their study is a closely related class of algebraic groups whose construction relies on the existence of a particular type of isogeny. Our final theorem says how these groups are related to the mixed algebraic groups:

**Theorem.** The exotic pseudo-reductive groups [CGP15, (8.2.2)] are Weil restrictions of (reductive) mixed group schemes.

See Theorem 10.5.1 for a precise statement and proof.

### 10.2 Mixed groups over visible fields

First we want to construct some mixed algebraic groups; these groups are reductive and in fact even semi-simple, although we will refrain from defining those notions in the mixed case.

We will show existence of some groups \((G_1, G_2, \varphi_1, \varphi_2)\) over visible fields \(m(k) = (k, k, \fr_k, \id_k)\), relying on Proposition 8.5.4 for a construction from a relative factorization

\[
\Fr_{G_2/k} : G_2 \xrightarrow{\varphi_2} G_1 \to G_2 \times_{\fr_k} \Spec k = \Fr_k^* G_2 = \Delta G_2.
\]
Most often, $G_2$ is smooth over $k$ and therefore it is reduced. But then the Frobenius $\text{Fr}_{G_2}$ is an epimorphism and this implies that $\varphi_2$ is epic as well (Lemma 9.2.2). Since the kernel of $\varphi_2$ is contained in the kernel of the relative frobenius $\text{Fr}_{G_2/k} : G_2 \to \Delta G_2$, we can use Borel’s technique [Bor91, §17] of taking the quotient of a group by a Lie subalgebra to construct all mixed $k$-groups $(G_1, G_2)$ for a fixed smooth $G = G_2$ as follows.

Let $\mathfrak{g}$ denote the Lie-algebra of $G$ and $\mathfrak{h}$ a restricted subalgebra which is Ad-invariant. Then there is a $k$-group $G/\mathfrak{h}$ and a $k$-isogeny $\pi : G \to G/\mathfrak{h}$ such that its differential factors as

$$d\pi : \text{Lie}(G) = \mathfrak{g} \to \mathfrak{g}/\mathfrak{h} \to \text{Lie}(G/\mathfrak{h}).$$

By taking the quotient of $G/\mathfrak{h}$ with $\text{im}(d\pi)$ we obtain a map $\overline{\pi} : G/\mathfrak{h} \to \text{Fr}^*G$.

Note that the partial dimensions of the resulting mixed object are by construction given by $(\dim \mathfrak{h}, \dim \mathfrak{g} - \dim \mathfrak{h})$. By Lemma 8.5.3, these isogenies correspond to relative factorizations in the sense of Definition 8.5.2; applying Proposition 8.5.4 then yields the mixed object $(G/\mathfrak{h}, G, p \circ \overline{\pi}, \pi)$ where $p$ is the (non $k$-linear) projection $\text{Fr}^*G \to G$.

**Proposition 10.2.1.** To every column of the following table corresponds a mixed group scheme.

<table>
<thead>
<tr>
<th>$p$</th>
<th>2</th>
<th>2</th>
<th>2</th>
<th>3</th>
<th>2</th>
<th>2</th>
<th>2</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X$</td>
<td>$\mathbb{B}^{sc}_n$</td>
<td>$\mathbb{C}^{sc}_n$</td>
<td>$F_4$</td>
<td>$G_2$</td>
<td>$\mathbb{B}^{ad}_{2n}$</td>
<td>$\mathbb{C}^{ad}_{2n}$</td>
<td>$\mathbb{B}^{ad}_n$</td>
<td>$\mathbb{C}^{ad}_n$</td>
</tr>
<tr>
<td>$Y$</td>
<td>$\mathbb{C}^{sc}_n$</td>
<td>$\mathbb{B}^{sc}_n$</td>
<td>$F_4$</td>
<td>$G_2$</td>
<td>$\mathbb{C}^{sc}_n$</td>
<td>$\mathbb{B}^{sc}_{2n}$</td>
<td>$\mathbb{C}^{ad}_n$</td>
<td>$\mathbb{B}^{ad}_n$</td>
</tr>
</tbody>
</table>

More precisely: let $k$ be a field of characteristic $p$; let $X$ denote a semi-simple $k$-group of that type. Then there exists a $k$-group $Y$ of indicated type such that there is a mixed $k$-group $M_{XY}$ such that $M_{XY} = (Y, X, p \circ \overline{\pi}, \pi)$.

**Proof.** The simply connected columns (in particular the $F_4$ and $G_2$ columns) are dealt with by the preceding discussion and [CGP15, (7.1.3–5)] where the very special isogenies are constructed as quotients of $X$ with a Lie-subalgebra. Since the very special isogenies send the centers to the centers, we immediately get the similar result for the adjoint groups.
Moreover, following the notation of loc. cit (7.1.2) in the cases $X = B_{2n}^{sc}$ or $C_{2n}^{sc}$ the irreducible submodule $\mathfrak{z} = \text{Lie}(Z)$ is contained in the kernel of the very special isogeny (on a separable closure). So the very special isogeny $\pi$ kills the center, and thus factors through the adjoint group $X^{\text{ad}} = X/Z(X)$ and we obtain a diagram of epimorphisms

$$
\begin{array}{ccc}
X^{\text{sc}} & \xrightarrow{\pi' = p \circ \pi} & Y^{\text{sc}} \\
p_1 \downarrow & & \downarrow \alpha \\
X^{\text{ad}} & & 
\end{array}
$$

such that $\alpha \circ p_1 \circ \pi'$ and $\pi' \circ \alpha \circ p_1$ are the absolute Frobenius on the respective objects. Since $p_1$ is epic, we have that $p_1 \circ \pi' \circ \alpha = \text{Fr}$ iff $p_1 \circ \pi' \circ \alpha \circ p_1 = \text{Fr} \circ p_1 = p_1 \circ \text{Fr}$, which is clearly the case and this gives the isogenies between simply connected $B$ and adjoint $C$ types and vice versa.

\begin{proof}
\end{proof}

**Remark 10.2.2.**

1. A base change will provide us now with a large number of mixed groups over arbitrary mixed fields. But, as we noted earlier in Remark 8.5.6, there is no reason to believe that every mixed group can be realised in this manner. In fact, the next bullets show that one can never obtain groups of type $B_{2n}/C_n$ with the group of type $C_n$ non-split by base changing one of the special isogenies between groups of types $B_n \to C_n$ and it seems likely that the same observation is true with the roles $B$ and $C$ swapped. So although there could a priori certainly exist mixed groups of type $B/C$ where both components are non-split, they do not arise via base change from a visible field.

2. The cases $X_n = B_{2n}^{\text{ad/sc}}$ and $Y_n = C_{2n}^{sc}$ correspond to a classical construction, see [CGP15, (7.1.6)]. One starts from a defective non-degenerate quadratic form on a $2n + 1$-dimensional space $V$. The groups of automorphisms of the quadratic form preserves the 1-dimensional radical $V^\perp$; therefore the automorphism group $\text{Aut}(V, q) = \text{SO}(q)$ acts on the space $V/V^\perp$ endowed with non-degenerate alternating form $\overline{B}(\overline{x}, \overline{y}) = q(x + y) - q(x) - q(y)$. This provides a morphism $\text{SO}(q) \to \text{Sp}(\overline{B}_q)$, which is a morphism between groups of types $B_{2n}^{\text{ad}} \to C_{2n}^{sc}$. 


3. However, note that the resulting symplectic group is always a split group, since every non-degenerate alternating form can be reduced to the standard form $\sum_i x_ix'_i$. Since [CGP15] claims that the isogeny $\pi: C_n \to Fr^*B_n$ is unique, it should be the case that $Fr^*SO_{2n+1}(q)$ is always a split group. This is indeed the case, since

$$x_0^2 + \sum_{i=1}^n (a_ix_i^2 + x_ix'_i + a'_ix'_i^2) \simeq k^{1/2} (x_0 + \sum_{i=1}^n (\sqrt{a_i}x_i + \sqrt{a'_i}x'_i))^2 + \sum_{i=1}^n x_ix'_i.$$ 

4. The construction for $G_2$ requires more specialized knowledge. The issue is that in characteristic 3 the adjoint representation of dimension 14 is not irreducible, but contains the standard representation of dimension 7; this provides an ideal in the Lie algebra.

To see how this happens, we recall that any group of type $G_2$ can be realized as automorphism group $\text{Aut}(O)$ of an octonion algebra $O$ endowed with a binary product, a norm $q$ and unit 1. The corresponding Lie algebra arises as derivations of the octonion algebra: $g_2 = \text{Der}(O)$. An octonion algebra is alternative, this means that the associator $(x,y,z) = (xy)z - x(yz)$ is an alternating trilinear map. Therefore if we define a new product $[xy]$ on $O$ by setting $[xy] = xy - yx$, then clearly $[xx] = 0$ and moreover we have the identity

$$[[x_1x_2]x_3 + [x_2x_3]x_1 + [x_3x_1]x_2 = \sum_{\sigma \in \text{Sym}(3)} (-1)^\sigma (x_{\sigma(1)}, x_{\sigma(2)}, x_{\sigma(3)}) = 6(x_1, x_2, x_3).$$

Therefore in characteristic 3, the Jacobi identity holds and we have endowed $O$ of dimension 8 with the structure of a Lie algebra which we denote with $\mathfrak{O}$. Of course $I = \langle 1 \rangle$ is an ideal in $\mathfrak{O}$ and the quotient $\mathfrak{V} = \mathfrak{O}/I$ is a 7-dimensional Lie-algebra. (It is actually the Lie algebra $\mathfrak{psl}_3$.) The adjoint representation of this Lie algebra provides a map

$$\mathfrak{V} \to \text{Der}(\mathfrak{V}),$$
The proof is completed by showing that \( \text{Der}(\mathfrak{W}) = \text{Der}(\mathfrak{O}) \). Of course there is a map \( \text{Der}(\mathfrak{O}) \to \text{Der}(\mathfrak{W}) \). A map in the other direction can be constructed by exploiting the orthogonal decomposition of quadratic spaces \( \mathfrak{O} = \langle 1 \rangle \perp 1^\perp \) and defining a product \( \times \) on \( 1^\perp \) by \( a \times b = ab - q(ab, 1)/q(1, 1)1 \), i.e. by projecting the octonion product back to \( 1^\perp \). Since every derivation of \( \mathfrak{O} \) sends \( 1 \) to \( 0 \), this construction allows one to extend a derivation on \( (1^\perp, \times) \) to a derivation on \( \mathfrak{O} \). The final observation is then that \( (1^\perp, \times) \) is (up to a multiple) equal to \( (\mathfrak{W}, [, [,]]). \)

5. We suspect that there exists a parallel construction for \( F_4 \) along the following lines. The issue is again that the standard representation of dimension 26 is contained in the adjoint representation of dimension 52, providing an ideal in the Lie algebra. Any group of type \( F_4 \) can be realized as automorphism group \( \text{Aut}(A) \) of an Albert algebra \( A \) of dimension 27 endowed with a quadratic operator \( U \), a norm \( q \) and unit \( 1 \). An Albert algebra is a quadratic Jordan algebra, this means that if \( p = 2 \) it also has the structure of a restricted Lie algebra by defining a Lie bracket and \( p \)-operation as follows:

\[
[x, y] = (U_{x+y} - U_x - U_y)1 \quad \text{and} \quad x^{[2]} = U_x 1.
\]

For a proof, see [Jac69, p.30, §1.4]; the important equation is QJ20 on page 24 which states essentially that

\[
[x^{[2]}, y] = [x, [x, y]] + 2U_{xy},
\]

and thus if \( p = 2 \) we have \( \text{Ad}(x^{[2]}) = \text{Ad}(x)^2 \) from which the Jacobi identity follows by linearization. From here on, one can concoct an argument which parallels the argument we sketched here for \( G_2 \).

6. The table from Proposition 10.2.1 does not cover all possibilities with both groups semi-simple; for instance see Example 9.4.5 for type \( A_{p-1} \) in characteristic \( p \), and there are further possibilities for types \( B, C, D \) and \( E_7 \) in characteristic 2 and \( E_6 \) in characteristic 3. We do believe that the table covers all non-central possibilities.
10.3 Twisted groups: the Suzuki–Ree groups

Theorem 10.3.1. Let $\mathbb{2}X_n(k, \theta)$ be a twisted Chevalley group, as defined by [Car72, Ch. 13 §4]. Then there is a twisted group scheme $\tilde{X}$ and a blended field $b$ such that $\mathbb{2}X_n(k, \theta) \cong \tilde{X}(b)$.

Proof. It is well known that there is an endomorphism of $\mathbb{F}_p$-groups $g : X_n \to X_n$ which extends the graph automorphism of the Chevalley group $X_n(\mathbb{F}_p) \to X_n(\mathbb{F}_p)$. If $\theta \in \text{Aut}(k)$ is chosen so that $fr_k \circ \theta^2 = \text{id}_k$ then the twisted group is defined as the group of fixed points of $g\theta$ acting on $X_n(k)$. On the other hand, we get a twisted group $\tilde{X} = (X_n, g)$ over the twisted field $(\mathbb{F}_p, \text{id})$ and by Proposition 9.2.1 we have $\tilde{X}(b) \cong (X_n(k))^{g\theta} = \mathbb{2}X_n(k, \theta)$.

Remark 10.3.2. The twisted groups in [Car72] are only defined over perfect fields and therefore the previous theorem is only meaningful in that case. The restriction to perfect fields is not a shortcoming of the theory; rather it is a shortcoming of the construction of the Suzuki-Ree groups as groups of fixed points of an involution. Close examination of Tits’s alternative definition [Tit61, p. 66, “$\alpha_\pi(g) = \alpha_\sigma(g)$”]—which also works over non-perfect fields—shows that he defines the twisted groups by the equation $xe\Phi = gx$, i.e. as groups of rational points of a twisted group scheme. So there is simply nothing to prove in that case and it is completely trivial that they are also groups of rational points of twisted group schemes. So in some sense our proof simply comes down to showing that Tits’s non-perfect definition indeed generalizes the definition [Car72] more commonly accepted as standard.

10.4 Mixed groups and buildings of mixed type

Theorem 10.4.1. Let $X_n(k, \ell)$ be a split mixed group as defined by [Tit74, (10.3.2)]. Then there exists a mixed group scheme $\tilde{X}$ and a mixed field $m$ such that $X_n(k, \ell) \cong \tilde{X}(m)$.

Proof. Step 0. Let us review the construction of the split mixed
groups from [Tit74, (10.3.2)] in the notation from loc. cit.. One starts with the adjoint $k$-split simple algebraic group $X_n$ defined over the field $k$. If we choose a maximal $k$-split torus $T$, then there is a corresponding root system $\Phi = \Phi_\geq \cup \Phi_\leq$ consisting of a set of short roots and long roots. For each root $r \in \Phi$, there is an $k$-unipotent subgroup $U_r$ upon which $T$ acts, one for each $r \in \Phi$ together with an isomorphism $u_r : G_a \to U_r$. Tits then defines the mixed group by providing the following set of generators:

$$X_n(k, \ell) = \langle T(k, \ell) \cup \{U_r(k) \mid r \in \Phi_\geq\} \cup \{U_r(\ell) \mid r \in \Phi_\leq\} \rangle \subseteq X_n(\ell),$$

where we have

$$T(k, \ell) = \{t \in T(\ell) \mid r(t) \in k \text{ if } r \in \Phi_\geq\} \subseteq T(\ell).$$

We know from [CGP15, (7.1.1)] that if we express a long root as a linear combination of fundamental roots, the coefficients of the short roots are all divisible by $p$. Therefore we can also define

$$T(k, \ell) = \{t \in T(\ell) \mid r(t) \in k \text{ if } r \in \Delta_\geq\} \subseteq T(\ell),$$

where $\Delta$ is a system of fundamental roots and $\Delta_\geq = \Delta \cap \Phi_\geq$. In rough terms: $X(k, \ell)$ arises from $X(\ell)$ by restricting the long roots to the smaller field, both for the root subgroups and for the torus.

**Step 1.** We will show that $X_n(k, \ell)$ can be constructed as follows. Let $Y_n$ denote the dual group, i.e. $Y_n = B_n^{ad}$ if $X_n = C_n^{ad}$ etc and $\pi : X_n \to Y_n$ the corresponding very special isogeny between adjoint groups from Proposition 10.2.1 and let us also assume a choice of maximal torus, roots and fundamental roots in both groups, denoted by $T$, $\Phi$, $\Delta$ and $T$, $\Phi$, $\Delta$ with a bijection $\Phi \to \Phi : r \mapsto \tau$ as constructed in [CGP15, (7.1.5)]. Then there are maps

$$\begin{array}{ccc}
X_n(\ell) & \xrightarrow{f=\pi_\ell} & Y_n(\ell) \\
\downarrow & & \downarrow \\
X_n(k) & \xleftarrow{Y_n(inc)} & Y_n(\ell)
\end{array}$$

We claim that $X_n(k, \ell) = f^{-1}(Y_n(k))$, in other words, let $x \in X_n(\ell)$ be arbitrary, then we claim that

$$x \in X_n(k, \ell) \iff f(x) \in Y_n(k). \quad (10.1)$$
Step 1a. Let us first show that
\[ t \in T(k, \ell) \iff f(t) \in \overline{T}(k) \]

Since the group is adjoint, we can rely on the isomorphism
\[ T(\ell) \to \prod_{r \in \Delta} \text{GL}_1(\ell) : t \mapsto (r(t))_{r \in \Delta} \]
and a similar isomorphism for $\overline{T}$. The very special isogeny $f = \pi_\ell$ induces the Frobenius on the $\text{GL}_1$ corresponding to short fundamental roots and the identity on the long roots. This means that an element $(r(t))_{r \in \Delta}$ is sent to $(r(t)^{s_r})_{r \in \Delta}$, where $s_r = p$ if $r$ is short and $s_r = 1$ if $r$ is long. Therefore the value on the long roots in $\Delta$ automatically ends up in $\ell^p \subseteq k$. Therefore the condition $f(t) \in \overline{T}(k)$ states that $r(t)$ must be contained in $k$ whenever $\tau$ is short and thus $r$ is long, i.e. $t \in T(k, \ell)$.

Step 1b. We can now show the equivalence (10.1). One implication is clear, since $f(X_n(k, \ell)) \subseteq Y_n(k)$ as is easily seen by evaluating $f$ on each of the generators. For the other implication, we consider an arbitrary $x \in X_n(\ell)$. The normal form (see [Hum75]) of $x$ is of the following form:

\[ x = \prod_{r \in \Phi} u_r(x_r) \cdot n(\sigma)t \cdot \prod_{r \in \Phi_\sigma} u_r(y_r) \]

We can use [CGP15, (7.1.5)] to compute $f(x)$:

\[ f(x) = \prod_{\tau \in \Phi} \overline{u}_\tau(x_{\tau}^{s_\tau}) \cdot n(\sigma)f(t) \cdot \prod_{\tau \in \Phi_\sigma} \overline{u}_\tau(y_{\tau}^{s_\tau}), \]

where $s_\tau$ is the same number from earlier. Expressing that $f(x) \in Y_n(k)$, relying in particular on the uniqueness of this normal form and the observation that $\Phi_\sigma = \Phi_\overline{\tau}$, we get the conditions

\[ x_{c}^{s_c}, y_{c}^{s_c} \in k \text{ and } f(t) \in \overline{T}(k), \]

where $\overline{T}$ is a maximal torus in $Y_n$. Recalling that $\ell^p \subseteq k$ and by relying on step 1a for the condition on $f(t)$, we conclude that that $x$ is indeed generated by Tits’s generators.
Step 2. We now use the same data, i.e. the isogeny $\pi : X_n \to Y_n$ and the field extension $\ell/k$ together with Corollary 8.5.5 to obtain a mixed group scheme $\tilde{G} = ((Y_n)_k, (X_n)_\ell)$ over the mixed field $m = (k, \ell)$. Since a field is certainly reduced as a scheme, we may apply Proposition 9.2.3 to compute its set of rational points:

$$\tilde{G}(m) = f^{-1}(Y_n(k)).$$

Here $Y_n(k)$ is considered a subset of $Y_n(\ell)$ in the natural way and $f : X_n(\ell) \to Y_n(\ell)$ is the map induced by $\pi$. In Step 1 we have proven that

$$f^{-1}(Y_n(k)) = X_n(k, \ell),$$

and therefore Tits's mixed group is indeed realized as rational points of the mixed group scheme $\tilde{G}$. 

**Remark 10.4.2.** The following remarks are some expectations that we have, but that will require some future work to verify.

1. Continuing the notation from the proof, we let $\Phi = \Phi_\succ \cup \Phi_<$ be a root system of type $X_n$ with fundamental system $\Delta = \Delta_\succ \cup \Delta_<$. Note that $|\Delta| = n$; we also denote $r = |\Phi|, n_\succ = |\Delta_\succ|$ etc. Since the dimension of the corresponding algebraic group of type $X_n$ is $r + n$, this is also the dimension of each of the components of the mixed group scheme $\tilde{X}$ over $(k, \ell)$. Nonetheless, the partial dimensions are given by $(r_\succ + n_\succ, r_\prec + n_\prec)$.

2. Our original motivation for studying mixed schemes—and not just rings or affine schemes—was that the homogeneous spaces of a mixed group will be mixed schemes. More specifically we can start from a the mixed group $(G, G')$ and take a pair of Borel subgroups $B \subset G, B' \subset G'$ in such a way that the mixing maps send $B$ into $B'$ and conversely $B'$ into $B$. This allows one to construct a mixed scheme $G/B \leftrightarrow G'/B'$ where both components have dimension dimension $r/2$ (since $\dim B = n + r/2$) but partial dimensions $(r_\succ/2, r_\prec/2)$. Taking rational points of this scheme over $(k, \ell)$, we obtain a set which can be identified with the flag complex of the mixed building.

3. Moreover, let us consider a set $\Gamma \subset \Delta$ of fundamental roots and denote by $s$ the number of roots in the induced subrootsystem.
To the choice of $\Gamma$ correspond parabolic subgroups $B \subseteq P \subseteq G$ and $B' \subseteq P' \subseteq G'$ of dimension $n + (r + s)/2$ and corresponding schemes $G/P$ and $G'/P'$ of dimension $(r - s)/2$. These subgroups give rise to a mixed scheme $G/P \leftrightarrow G'/P'$ of partial dimensions $(\frac{r-s}{2}, \frac{r+s}{2})$. Typically the set $\Gamma$ is chosen as large as possible to obtain structures of reasonable dimension that are studied by algebraic or incidence geometers.

4. In particular, if $S$ contains all the long (or short) fundamental roots, one of the partial dimension will be 0. For instance, we could start from a root system of type $B_n$ and take the subset $\Gamma \subset \Delta$ consisting of all $n - 1$ long fundamental roots. If we note that there are precisely $2n$ short roots, we expect to find a mixed scheme of partial dimension $(0, n)$: the mixed quadric. It is very easy to describe this thing—or at least an affine part of it—more explicitly over a mixed field $(k, \ell, \kappa, \lambda)$: we define the $k$-algebra $Q$ and $\ell$-algebra $W$ by

$$Q = k[x_0, x_1, \ldots, x_n]/(x_0^2 - q(x_1, \ldots, x_n)),$$

$$W = \ell[y_1, \ldots, y_n],$$

where $q$ is a non-degenerate quadratic form, e.g. the hyperbolic one. We then define maps between both algebras extending $\kappa$ and $\lambda$ as follows:

$$\tilde{\kappa} : Q \to W : \begin{cases} x_i \mapsto y_i^2 & \text{for } i \geq 1 \\ x_0 \mapsto q^\kappa(y_1, \ldots, y_n) \end{cases}$$

$$\tilde{\lambda} : W \to Q : y_i \mapsto x_i,$$

where $q^\kappa$ simply applies $\kappa$ to the coefficients of $q$. It is easily verified that this defines a mixed ring $(Q, W, \tilde{\kappa}, \tilde{\lambda})$ and therefore a mixed affine scheme. Since $\tilde{\kappa}$ sends all variables to squares, the differential vanishes and the mixed object has indeed partial dimensions $(0, n)$. 

10.5 Exotic pseudo-reductive groups

In [CGP15] the authors provide a structure theory for pseudo-reductive groups over a base field \( k \). In rough terms the outcome is that almost every pseudo-reductive group arises from a standard construction, which has as its starting point a Weil restriction \( R_{k'/k}(G) \) of a reductive \( k' \)-group \( G \) through a purely inseparable field extension \( k'/k \). Some exotic examples, introduced in Chapter 7 of op.cit. do not follow this pattern and require a more elaborate construction. Our next theorem states that these exotic groups, actually do fit this pattern, albeit within the category of mixed group schemes. In fact the so called basic exotic groups can be thought of as \( G = R_m/k(\tilde{G}) \), where \( \tilde{G} \) is a reductive mixed group over the mixed field \( m \) and \( R_m/k \) denotes the mixed version of the Weil restriction. An arbitrary exotic group arises by further Weil restriction of a basic exotic group; of course this is also a Weil restriction.

**Theorem 10.5.1.** Let \( \mathcal{G} \) be an exotic pseudo-reductive group as defined by [CGP15, (8.2.2)]. Then there is a mixed group scheme \( \tilde{G} \) over a mixed algebra \( M \) such that \( \mathcal{G} = c_2 R_{M/k}(\tilde{G}) \).

**Proof.** Let us review the construction of the basic exotic groups from [CGP15, (7.2.3)] in the notation from loc. cit.: one starts with a very special isogeny of \( k \)-groups \( \pi : G \rightarrow \overline{G} \) and a purely inseparable field extension \( \ell/k \) of finite degree such that \( \ell^p \subset k \). Base changing \( \pi \) to \( \ell \) followed by Weil restriction back to \( k \) gives a map

\[
f : R_{\ell/k}G_{\ell} \rightarrow R_{\ell/k}\overline{G}_{\ell}.
\]

On the other hand, the unit of the adjunction between base change and Weil restriction provides a map \( \overline{G} \rightarrow R_{\ell/k}\overline{G}_{\ell} \). Then one defines \( \mathcal{G} = f^{-1}(\overline{G}) \). Changing the notation to our own, we denote \( \beta : \text{Spec } \ell \rightarrow \text{Spec } k \) so that \( f = \beta_*\beta^*\pi \) and

\[
\mathcal{G} = \overline{G} \times_{\beta_*\beta^*\overline{G}} \beta_*\beta^*G.
\]

Now, let us start from the same very special isogeny and field extension \( \ell/k \), and construct the mixed group \( \tilde{G} = (\overline{G}, \beta_*G, \ldots) \) as in
Chapter 10. Twisting and mixing groups

Corollary 8.5.5. Applying the formula from Proposition 8.7.1 we find

$$R_{m/k}(\tilde{G}) = h_*(\tilde{G}) = (\overline{G}, \overline{G} \times_{\beta_*\beta^*G} \beta_*\beta^*G, \ldots)$$

where $h: \text{Spec}(k, \ell) \to \text{Spec} k$ is the corresponding extension of mixed fields. Note that the second component is indeed $\mathcal{G}$.

Taking the product of such field extensions, we see that every $K$-group $\mathcal{G}$ for a non-zero finite reduced $k$-algebra $K$ whose fibers are basic exotic pseudo-reductive groups can be realised as $c_2 R_{M/K}(\tilde{G})$.

Applying the Weil restriction $R_{K/k}$ to this, we find that the exotic group $G = R_{K/k}(\mathcal{G})$ as in [CGP15, (8.2.2)] can be realized as $G = c_2 R_{M/k}(\mathcal{G})$.

Remark 10.5.2. 1. Although the catch phrase in our abstract states that exotic pseudo-reductive groups are Weil restrictions of mixed groups, we now see that this is not entirely correct. They arise from a Weil-mixtor restriction: first a Weil restriction $R_{M/k}$ and then a mixtor restriction $c_2$—see Remark 8.4.2.

2. The dimension of a semi-simple group is completely determined by combinatorial data coming from a root system. If we apply an inseparable Weil restriction, we find a pseudo-reductive group with dimension determined by combinatorial data of the original group, and the degree of the restriction morphism. For the exotic pseudo-reductive groups on the other hand, there is a formula (7.2.1) in [CGP15] which states

$$\dim \mathcal{G} = (r_> + n_>) + (r_< + n_<)[k': k].$$

Since the corresponding mixed group has partial dimensions (see Remark 10.4.2)

$$(r_> + n_>, r_< + n_<)$$

and a Weil restriction proceeds along an extension of mixed fields with degrees $(1, [k': k])$ we have this same separation of the dimension into combinatorial data of the original group and degree of a morphism.
3. If $[k' : k]$ is infinite, the mixed group still exists: all the infiniteness is gobbled up by the mixing maps but the structure morphism is still of finite type. But in that case taking a Weil restriction becomes problematic.

4. Initially we had hoped to also state and prove a theorem which states that Weil restrictions of mixed reductive groups are always pseudo-reductive. It seems conceivable that the proof of the corresponding statement from [CGP15, 1.1.10] will generalize to the mixed setting, but only after a thorough study of mixed versions of the typical adjectives from algebraic geometry such as smooth and connected. For instance, one must first study a mixed notion of smoothness and prove a mixed version of the infinitesimal criterion. Perhaps this will eventually allow to circumvent some of the difficulties encountered by Conrad–Gabber–Prasad by starting from an arbitrary pseudo-reductive group, constructing its parental mixed reductive group directly and deducing the structure theory of pseudo-reductive groups from there, similar to how the current structure theory works away from characteristic 2 and 3.

5. The exotic groups are not the only strange cases which appear in the theory of Conrad, Gabber and Prasad: in characteristic 2, there are other esoteric constructions. Nonetheless, these constructions are also related to mixed buildings, and admit a corresponding twisted group, a generalized Suzuki group. (See Section A.3 for some details.) This gives us some hope that these groups too fit into our framework. We see two observations which could be relevant. The first observation is that mixed algebraic groups related to regular but defective quadratic forms of defect $\geq 2$ are likely of interest [Car72, §1.6]. The second observation is that there could be mixed reductive groups over an invisible field which do not arise via base change from a mixed reductive group over a visible field, as we explained in Remark 10.2.2. Such groups could not be constructed as exotic groups because the presence of non $k$-linear maps cannot be circumvented so easily; this could explain why some of the constructions in the later chapters of [CGP15] are so indirect.
Appendices
Theorem. Every thick irreducible spherical building of rank two satisfying the Moufang condition as well as every thick irreducible building of rank greater than two is classical, algebraic, or mixed.

Jacques Tits and Richard Weiss

We will now provide a historical overview of people and facts related to mixed and twisted groups, with an attempt at being complete. Because the story is so heavily intertwined with the mathematical works of Jacques Tits, we will also try to tell his story, albeit with no attempt at being complete.

A.1 1950–1960

In 1955, Claude Chevalley publishes his Tōhoku paper where he shows that each of the known semi-simple Lie-algebras, known from the Killing-Cartan classification, gives rise to a class of groups, which can be defined over an arbitrary field: the Chevalley groups. The next few years Chevalley leads the Séminaire Cartan-Chevalley in 1955–56 and the Séminaire Chevalley in 1956–58, where the foundations for modern scheme theory and algebraic group theory are laid.

In 1957, Rimhak Ree [Ree57] shows that the Chevalley groups of types A, B, C, D correspond to the classical linear, symplectic and some of the orthogonal groups over the corresponding fields as one would

1[TW02, p. 4]
expect; he also shows that the groups of type $G_2$ are isomorphic to
groups defined much earlier by Dickson [Dic01; Dic05].

This raised the question whether the unitary groups could be seen
as variations on this theme. In 1959, Robert Steinberg [Ste59] pub-
lishes an article with that title presenting a new construction: he
observes that in the cases $A_n$, $D_n$ and $E_6$, the Dynkin diagram has an
automorphism which gives rise to an automorphism of the Chevalley
group. In combination with an automorphism of the underlying field
this gives interesting involutory automorphisms whose fixed points
are classes of simple groups now known as *Steinberg groups* $^2A_n$, $^2D_n$,
$^3D_4$ and $^2E_6$, where the former two provide the unitary groups and
the missing orthogonal groups, and latter two were new.

Quite a different perspective comes from Jacques Tits, who is looking
into generalizations of the *fundamental theorem of projective geometry.*
This theorem says that every permutation of the points of projective
space which is incidence preserving—i.e. sends lines to lines—is in-
duced by a semi-linear map on the underlying vector space. In other
words: combinatorial axioms for a projective space characterize the
Lie groups of type $A_n$. The program that Tits had begun pursuing
in the early 50s was to provide axiomatic systems of points and lines
which characterize the other Lie groups in a similar manner. In 1953
[Tit53], in some of his earliest work, he investigates the real octonion
plane, polarities of this plane, and the related Lie groups $F_{4(-52)}$
$F_{4(-20)}$, $E_{6(-26)}$ and $E_{7(-25)}$. A skeptical reviewer, named Chevalley,
spots a mistake in a proof which Tits fixes in a follow-up article
[Tit54] in 1954 where also the announced construction of $E_{7(-25)}$ is
given. (The latter groups make a come-back appearance in [Tit74]
under the guise of the *non-embeddable polar spaces.*) In 1955 [Tit55]
and 1956 [Tit56] Tits publishes two long studies about homogeneous
spaces of Lie groups, which can be interpreted as a precursor to his
theory of buildings. One of his early achievements is a description of
the (split) group $E_6$ for the Séminaire Bourbaki [Tit58] as the auto-
morphism group of some sort of ‘plane’—a parapolar space in later
terminology. Investigating polarities in this $E_6$-‘plane’ lead him to the
independent discovery of the groups of type $^2E_6$ over the reals.
A.2 1960–1970

In 1960, Michio Suzuki is investigating a class of groups named Zassenhaus groups. A Zassenhaus group is a permutation group which (i) acts doubly transitively, (ii) with only the identity fixing 3 elements and (iii) without a regular normal subgroup—the latter case only excludes some degenerate cases such as the Frobenius groups. Suzuki notes that a Zassenhaus group of odd degree is simple so he was more than interested in classifying them.

According to Suzuki [Suz60], it had been conjectured by Feit and Hasse that the only examples were the groups $\text{SL}(2, 2^n)$, but in loc. cit. he reports the discovery of a new class of Zassenhaus groups of odd degree. He constructs the groups $G(q) = \text{Sz}(q)$ as subgroups of $\text{GL}_4(q)$ generated by certain matrices, where $q$ is an odd power of 2. About these groups, he writes: The series of groups $G(q)$ gives, therefore, the second infinite series of simple groups which are not of Lie type. Nonetheless, further in the article he also notes that his generators leave a bilinear form invariant, so they are also subgroups of $\text{Sp}_4(q) = \text{B}_2(q)$. (In 1962 [Suz62] Suzuki would classify the odd degree Zassenhaus groups and show that the Suzuki groups complete the list, and in 1964 he would also classify the even degree Zassenhaus groups.)

Later that year, Ree realizes that Suzuki’s groups are in fact closely related to the Chevalley groups of type $B_2 = C_2$. If $k$ is a field and $\theta : k \to k$ an automorphism such that $\theta(\theta(x)) = x^2$, he could use this data to construct an involutory automorphism of $\text{Sp}_4(k)$ such that the fixed subgroup is precisely $\text{Sz}(q)$. Repeating the procedure for the Chevalley groups of types $F_4$ and $G_2$ he constructs what are now known as the large Ree groups $^2F_4$ [Ree61a; Ree61b] (for $p = 2$) and small Ree groups $^2G_2$ [Ree60; Ree61c] (for $p = 3$).

By 1961 Tits too has turned his attention to algebraic groups and he reports on a geometric approach to the simple groups of Suzuki and Ree for the Séminaire Bourbaki [Tit61]. A thorough treatment of the Suzuki groups was later also published in [Tit62a]. Tits’s work is an interesting variation on his earlier work on polarities: the ‘polarity’ of a plane with itself has to be replaced with what he calls une sorte
de dualité between two different varieties, embedded in $\mathbb{P}^3$ and $\mathbb{P}^5$ in the case of $^2B_2$ and embedded in $\mathbb{P}^6$ and $\mathbb{P}^{13}$ in the case of $^2G_2$. These varieties are actually the homogeneous spaces $G/P_1$ and $G/P_2$, where $G$ is the algebraic group of the corresponding type and $P_1$ and $P_2$ are the two classes of maximal parabolic subgroups containing a common Borel subgroup $B$. In later terminology the rational points of these varieties can be identified with the points and lines of a generalized quadrangle or hexagon, where the incidence relation can be read off from the flag variety. In other words, applying the functor $\Gamma : X \rightsquigarrow X(k)$ to the left diagram below, which is a diagram of schemes, we obtain the right diagram, which is a diagram of sets and can be regarded as an edge colored graph (then $\mathcal{F}$ are the vertices and $\mathcal{P}$ and $\mathcal{L}$ are the two colors) or an incidence geometry (then $\mathcal{P}$ and $\mathcal{L}$ are points and lines and $\mathcal{F}$ the incidence relation) but in either case it encodes a Moufang building of rank 2.

\[
\begin{array}{ccc}
G/B & \xrightarrow{\Gamma} & \mathcal{F} \\
G/P_1 & \xrightarrow{} & \mathcal{P} \\
G/P_2 & \xleftarrow{} & \mathcal{L}
\end{array}
\]

Tits must have felt that something remarkable was going on: his geometric construction provides maps which compose to the Frobenius, rather than to the identity. For a perfect field $k$, he could think of his construction as a polarity of the geometry with points and lines given by $((G/P_1)(k), (G/P_2)(k))$, but since Tits had observed that he could also make the construction of the Suzuki and Ree groups over imperfect fields, he chose to phrase it rather carefully as some sort of duality.

In the terminology that we introduced in this work and with the benefit of hindsight, we could say that Tits was looking at the mixed analogon

\[
\begin{array}{ccc}
G/B & \xhookrightarrow{} & G'/B' \\
G/P_1 & \xhookrightarrow{} & G'/P_2' \\
G/P_2 & \xhookleftarrow{} & G'/P_1'
\end{array}
\]
constructed from the mixed group \((G, G')\) over a mixed field \((k, \ell)\), with \(P_1, P_2, P'_1\) and \(P'_2\) appropriate parabolic subgroups. By taking rational points of this diagram of mixed schemes, we obtain again an incidence geometry. But in the era before incidence geometry, one would want to realise these mixed rational points as subsets of varieties, which can be done as in Proposition 9.2.3. For both the points and lines, one of the partial dimensions will be 0, as we explained in Remark 10.4.2. This suggests it is easier to realise the points by embedding them in a \(k\)-variety, whereas it is easier to realise the lines by embedding them in an \(\ell\)-variety! But at the time, this would have been impossible to see, because \(k \cong \ell\) and \(G \cong G'\).

Over the next few years, Tits drastically picks up the pace and makes many important contributions to group theory and geometry. At first this is often over perfect fields only, but after Alexander Grothendieck proved his deep theorem on the existence of maximal tori over arbitrary fields [SGA3, exp. XIV] in 1964, this restriction can be lifted. Let us just mention some of these developments: in [BT65] Borel and Tits provided their structure theory for reductive groups; in [Tit62b; Tit64] Tits initiates the theory of groups with a \(BN\)-pair; in [Tit66] he provides a structure theory for semi-simple groups in terms of their Tits index and anisotropic kernel; in collaboration with François Bruhat, he investigates the structure of algebraic groups over local fields. Meanwhile, he works on lecture notes on the theory of buildings. A first preprint of these notes appears in the fall of 1968 and must have circulated widely in the years thereafter, but the notes weren’t formally published until 1974 [Tit74].

The first time a mixed group makes an explicit appearance in the literature is, as far as we can tell, in Steinberg’s Yale lecture notes from 1967–1968 [Ste68] on groups of Lie type, in the following remark on page 153:

If \(k\) is not perfect and \(\varphi : G \to G\) then \(\varphi G\) is the subgroup of \(G\) in which \(X_\alpha\) is parametrized by \(k\) if \(\alpha\) is long and \(k^p\) if \(\alpha\) is short. Here \(k^p\) can be replaced by any field between \(k^p\) and \(k\) to yield a rather weird simple group.
It is probably not a coincidence that Tits is also in Yale around that time. In fact, Hendrik Van Maldeghem has suggested to us that the first time a mixed group (or variety) was observed in the wild may have been in Tits’s unpublished classification of Moufang hexagons. We could not date this classification precisely but given that Tits introduced generalized polygons as early as 1959 and that they were probably the first class of buildings he seriously investigated, it seems very plausible that it were indeed the hexagons which lead Tits to these groups for the first time.

One year earlier, in his own Yale 1966–1967 lecture notes on algebraic groups, Tits had investigated unipotent groups in positive characteristic. The lecture notes were never formally published until they appeared as appendix B1 of his collected works in 2014, although the results on unipotent groups had appeared earlier in the works of Oesterlé and, in a revised form, in appendix B of [CGP15]. Although we found no written evidence of this hypothesis, we believe it is likely that Tits thought that a thorough study of algebraic groups in positive characteristic could lead to a more satisfying explanation for why mixed groups are required to complete his classification of buildings.

A.3 1970–1990

The 70’s and early 80’s are the golden years for the classification of finite simple groups. While the mixed groups and mixed buildings begin gathering dust, the twisted groups, at least the finite ones, are an important part of the classification and as such well known and studied by group theorists. In particular we should mention that the characterization of the (small) Ree groups proved to be one of the hardest steps in the classification: it cost John G. Thompson three difficult papers [Tho67; Tho72; Tho77] in 1967, 1972 and 1977 to reduce it to a number-theoretic problem which is solved by Enrico Bombieri [Bom80] in 1981 in an dazzling application of elimination theory; the reviewer remarks that ordinary mortals such as the present reviewer are overawed by the author’s tour de force.

Also the representation theory of these twisted groups is studied
thoroughly. (We also noted a strange occurrence of the mixing functor \( m \), introduced in Section 8.2 in Pierre Deligne and George Lusztig’s work [DL76, §11] on representations for finite groups of Lie type where they remark that their approach to the Suzuki and Ree groups works equally well for “groups of the form \( G = G_1 \times G_1 \) with \( F'(x, y) = (F(y), x) \)” but it is unclear to us what the significance of this is.)

During that period, from the early 70s to the late 80s, with many researchers focusing on finite groups, and Tits himself lecturing at the Collège de France about sporadic groups in 1976–1977, and about the monster group in 1982–1983, 1985–1986 and 1986–1987, it could seem that not much is happening in the theory of buildings and algebraic groups. But at the same time Tits is actually working on the classification of Moufang polygons, on affine buildings, on Kac-Moody groups and algebras and twin buildings. We will not go into all these developments but focus on the Moufang polygons, since these are most relevant to our story.

As we mentioned earlier, we suspect that Tits had completed the classification of Moufang triangles, hexagons and octagons quite early, perhaps in the early 70s. By 1974, he finally publishes his lecture notes\(^2\), where he classifies (spherical) buildings in rank \( \geq 3 \). In this classification other mixed buildings pop up, namely those related to groups of type \( F_4 \) and those to groups of type \( B_n \) and \( C_n, n \geq 3 \). These groups—together with the \( G_2 \)-variant—are precisely the groups for which we show in our Theorem 10.4.1 that they arise as groups of rational points of a mixed group scheme. We note however, that the \( B/C \)-class admits further generalization to groups which are defined over a pair of fields \( K, L \) and an additional \( K \)-vectorspace contained in \( L \). We have not yet related these groups to our own work; the only insight that we have to offer here is that they are probably related to defective quadratic forms (see Remark 10.5.2.5).

So by 1974 all Moufang buildings of rank \( \geq 2 \) are classified except for the the Moufang quadrangles. Tits publishes a preprint with some thoughts on the subject around 1976 (it is referenced in Van Maldeghem’s book on generalized polygons [VM98, §3.4.2]) and it

\(^2\)Perhaps publication was delayed because he wanted to include a classification of Moufang polygons, but he never got around the case of the quadrangles
seems that after this Tits did not touch the subject in the next 20 years. One interesting feature is that since $B_2 \cong C_2$, the buildings with a pair of fields and a vector space from the previous paragraph, can be generalized to a ‘doubly exotic’ class of Moufang quadrangles defined over a pair of fields together with a pair of vector spaces over each field, contained in the other field. The class is also notable because there exists a twisted variant which generalizes the Suzuki-groups $^2B_2(k, \theta)$. (These groups are hinted at in [VM98, §7.6] and more explicitly studied by Van Maldeghem in 2007 [VM07]. They do not appear explicitly in Tits’s overview of Moufang sets in his 1999–2000 lecture notes.) It is noteworthy that there is no analogon for the case of $G_2$ and hexagons in characteristic 3.

Another interesting development from that time is a program, proposed by Francis Buekenhout [Bue79], to study and eventually classify sporadic groups by associating certain diagram geometries to them—some sort of generalizations of buildings. To some extent, geometric ideas do play an important role in the proof of the classification of finite simple groups, but these recognition theorems can only be applied deep into the proof, after a very difficult group theoretical analysis and case distinction. Even though Buekenhout’s program gets largely outpaced by the rapid developments in finite group theory, it marks the beginning of research in pure incidence geometry, with (algebraic or finite) groups coming in a posteriori or not at all.

With the end of the classification announced by Daniel Gorenstein in 1983—perhaps prematurely so—there is a definitely a renewed interest in the ideas surrounding the theory of buildings and with the appearance of textbooks such as [Bro89] and [Ron89] the subject also becomes more accessible to newcomers.

**A.4 1990–2000**

Most of Tits’s later research interests can only be found in his Résumés des Cours au Collège de France 1973–2000. Of particular interest to our story are the 1991–1992 and 1992–1993 courses on algebraic groups in positive characteristic with a focus on inseparable
phenomena and pseudo-reductive groups. One could consider it Tits’s metastrategy for doing mathematics to collect all the examples and then study their common features and eventually weave them into an elegant theory; as far as pseudo-reductive groups are concerned it seems that—with the benefit of hindsight and relying on [CGP15]—around that time Tits is in the process of constructing all the examples but a few crucial constructions are still missing. It is remarkable (and a bit unfortunate for us!) that although Tits’s constructions are very reminiscent of the mixed buildings he discovered decades earlier, he never makes the connection explicit. We can only guess why—our own guess is of course that Tits was unsatisfied with the idea that these pseudo-reductive groups were the ultimate culprit responsible for the mixed buildings, but the reader can make up his own story. After Tits’s lectures the subject would lay dormant again for many years.

In his 1994–1995 lectures then, Tits returns to the classification of Moufang polygons. Relying on his own unpublished work, he proposes a strategy to carry out the classification, lists the known types, and conjectures there are no others. To our own surprise, a definition in his 1994–1995 lecture notes §3 speaks of a pair of fields with maps $\kappa : K \to L$ and $\lambda : L \to K$ such that the compositions are the square operators. It is a subtle change in point of view that seems to have gone unnoticed by subsequent authors: although Richard Weiss recollects that Tits expressed a certain fondness of the symmetry between $K$ and $L$ on many occasions, it is an observation which seemed hard to exploit.

Tits’s lectures clearly worked inspiring because by 1996–1997 Weiss actually manages to complete this classification *faisant preuve d’une virtuosité technique remarquable*\(^3\) as Tits puts it on the first page of his 1999–2000 lecture notes. Around February 1997\(^4\) and much to Tits’s surprise, Weiss discovers a new and highly exotic class of Moufang quadrangles. Weiss recollects that at first, Tits was somewhat sceptical about the discovery, but he became very enthusiastic about it later on; in fact he decides to lecture about it at the Collège later

---

\(^3\)‘demonstrating a remarkable technical virtuosity’

\(^4\)According to Norbert Knarr in his review for [MV99]
that year. After the end of the course, in which Weiss’ discovery had been presented, Hendrik Van Maldeghem and Bernhard Mühlherr [MV99] find a way to realise these quadrangles as fixed buildings associated to a ‘Galois-like’ involution on the mixed buildings of type $F_4$ and so the class of quadrangles was dubbed *mixed quadrangles of type* $F_4$ when the classification of Moufang polygons appeared in print [TW02]. According to the Résumé de Course 1997–1998, Tits gives six lectures on the subject *exotic groups and Galois cohomology* where he adapts the notion of Tits index and anisotropic kernel to give a *Galoisian* proof of the existence of these quadrangles, unfortunately no further details are given. We asked around for a bit and although no-one could tell us for sure what happened during these seminars, it seems unlikely that Tits actually got to the Galois descent part.

We speculate that in our terminology, these *Weiss quadrangles* arise by mixing together two groups of type $F_4$ of relative rank 1. This would explain why these quadrangles can only exist in the exotic situation of a pair of fields $k^2 \subsetneq \ell \subsetneq k$ with strict inclusions: if one of the mixing maps is linear this would split on of the quadratic forms which underly the anisotropic kernel of type $B_3$ as in Remark 10.2.2.3 and then one of the components would be of relative rank 4. It is also required that both components are isomorphic in order for the group to have a chance to admit twisted descent, which appears to be the case sometimes, as we explain below.

Tits’s last set of lecture notes dates from 1999–2000. Inspired by the success of his lectures on Moufang polygons, he lectures on *groups of rank 1 and Moufang sets*. He clearly has some hope that at some point a classification may be achieved, although we add that to this date, most experts believe that this is still far out of reach. The final section is titled *immeubles de Moufang de rang 1 (suite mais non fin)* (Moufang buildings of rank 1—sequel but not the end).

### A.5 2000–2016

Around the year 2000, Tits retires from public mathematical life but there are many other mathematicians ready to take up the baton.
Nonetheless it seems that no one could oversee the fields of incidence geometry and (pseudo-)reductive algebraic group theory the way Tits could and as a result developments in both fields occurred more or less independently from that point onwards. We distinguish two major developments that are relevant for our history of twisted and mixed groups.

A first development arose in the process of collecting all the examples of twisted and mixed groups. Since no further examples were to be expected for rank \( \geq 2 \), the innovations concern the rank 1 case. In 2006 [MM06], Mühlherr and Van Maldeghem find new examples of Moufang sets, arising by some sort of Galois descent from the mixed quadrangles of type \( F_4 \). Since these quadrangles are themselves some sort of twisted \( F_4 \)-buildings, these Moufang sets are sometimes called doubly twisted Moufang sets of type \( F_4 \). They stand out because, together with the small Ree groups \(^2G_2\), they are the only Moufang sets with root groups of nilpotency class 3, rather than 1 or 2. Later, in [DSW], Tom De Medts, Yoav Segev and Richard Weiss show that these groups can also be obtained starting from groups of type \(^2F_4\) related to the Moufang octagons, resulting in a ‘commuting diagram’ of groups or geometries as depicted below. In fact this diagram commutes in a strong sense: every doubly twisted group of type \( F_4 \) can be obtained via either route. We suspect that in our terminology, these doubly twisted groups arise by taking rational points of twisted group schemes which fit into the diagram below; if something like this is true the main result of [DSW] can be paraphrased by stating that Galois descent commutes with twisted descent (as introduced in Section 8.3). The diagram under consideration is then the diagram of twisted and mixed groups, defined over the twisted and mixed fields that are depicted on the left.

\[
\begin{align*}
(K, L) &\quad (k, \ell) &\quad \text{mixed } F_{4,4} \\
(K, \theta) &\quad (k, \theta) &\quad \text{mixed } F_{4,1} &\quad \text{mixed } F_{4,1}
\end{align*}
\]

In these diagrams an arrow signifies that its target arises from its source via some sort of descent—which can be twisted descent \((T)\) or
Galois descent \((G)\). Conversely, one may extend scalars from a group at the target of an arrow create an group at its source. In particular, if \(X\) is a group of type \(\mathbb{G}_2\), it give rise to a group of type \(\mathbb{G}_2\), i.e. one of the large Ree groups, via the functor \(X \sim X \otimes (k, \theta) (K, \theta)\). On the other hand, it also gives rise to a mixed group of type \(\mathbb{F}_4\), via the functor of base change \(X \sim X \otimes (k, \theta) (k, k, \theta, \theta)\). In particular, it gives rise to a mixed split group of type \(\mathbb{F}_4\) in two ways, but these are isomorphic because the extension of scalars satisfies the property

\[
(X \otimes (K, \theta)) \otimes (K, K) = X \otimes (K, K) = (X \otimes (k, k)) \otimes (K, K).
\]

Taking the corresponding geometries, we obtain the following diagram, which depicts Moufang buildings of rank 4, 2, 2 and 1.

\[
\begin{array}{ccc}
\text{mixed } \mathbb{F}_4 \text{ building} & \sim & \text{Weiss quadrangle} \\
\text{Ree–Tits octagon} & \sim & \text{doubly twisted } \mathbb{F}_4 \\
(G) & \sim & (T) \\
\end{array}
\]

This could make one of the main results of \([DSW]\) far easier to obtain, but one must take into account that the difficulty has now shifted to proving that these geometries indeed arise from mixed groups and twisted groups and that ‘descent in buildings’ indeed corresponds to descent in the category of twisted schemes in these cases. These are themselves very non-trivial statements and our main results so far have only investigated the arrow from the mixed \(\mathbb{F}_4\) building to the Ree-Tits octagon, everything else is still a hypothesis.

An important further development is the trilogy \([Wei03; Wei09; MPW15]\), with the first two monographs authored by Weiss and the last one by Bernhard Mühlherr, Holger Petersson and Richard Weiss. The first two books aim to provide proofs of Tits’s classification theorems for spherical and affine buildings, accessible to a wide audience and without invoking existence theorems from algebraic group theory which ultimately rely on root system computations, inspired by Lie theory. The final book completes this decoupling of algebraic groups
and buildings by providing a purely combinatorial descent theory for buildings (amongst other things). This provides a solid foundational background in which one can study the descent diagram for $F_4$-geometries that we drew earlier and in particular it provides a far reaching generalization of (what we presume was) Tits’s Galoisian proof of the existence of the mixed quadrangles of type $F_4$ and the construction in [MV99].

In 2015, Elizabeth Callens and Tom De Medts [CD15] find Moufang sets related to groups of type $F_4$ and relative rank 1. We speculate that these groups are mixed groups which arise from mixing two groups of type $F_4$, one split and one of relative rank 1. In particular such groups cannot admit twisted descent since the components are never isomorphic, as we saw in Corollary 8.3.7. It is remarkable that these easier groups were only discovered later: one reason is probably that incidence geometric intuition becomes frail in low rank, because this means the geometry has only points and no lines or subspaces of higher dimensions. In rank 0, the geometry is empty and then such intuition completely breaks down. So although we are now encouraged to search for, say, anisotropic groups of type $^2F_4$—by applying twisted descent to anisotropic mixed groups of type $F_4$—they haven’t been described thus far.

A second independent development that took place around the same time was the development of a structure theory and classification theorems for pseudo-reductive algebraic groups. This work appeared in the monographs [CGP15] by Brian Conrad, Offer Gabber and Gopal Prasad and [CP16] by Conrad and Prasad. (Bertrand Rémy wrote an accessible exposition with some of the key ideas of the first book [Rém11] and there is also survey of both books by Conrad and Prasad [CP17].) The outcome of this classification effort is that with few exceptions pseudo-reductive groups arise through a standard construction which requires as input the Weil restriction of a reductive group (amongst other things). The only exceptions arise in characteristic 2 and 3; an important class, which includes all characteristic 3 exceptions is the class of exotic groups. Our Theorem 10.5.1 states precisely that these exotic groups arise as Weil restrictions too, but starting from mixed reductive groups. It should be noted that this
theory encounters many difficulties beyond these exotic groups too, which we have not related to our theory of mixed group schemes yet. Some of the difficulties are certainly related to the mixed groups associated to the general class of mixed groups of type $B_n/C_n$ associated to a pair of fields together with a vector space, that we mentioned earlier—and the extra complication when $n = 2$ with two vector spaces. Actually in characteristic 2, the first edition of [CGP15] often assumes the situation of a base field $k$ such that $[k : k^2] ≤ 2$ to avoid having to deal with these situations. This shortcoming is absent in the second edition where new ideas lead to a complete theory, regardless of $[k : k^2]$. 


"Mathematics is a process of staring hard enough with enough perseverance at the fog of muddle and confusion to eventually break through to improved clarity. I’m happy when I can admit, at least to myself, that my thinking is muddled, and I try to overcome the embarrassment that I might reveal ignorance or confusion."

Bill Thurston

B.1 m-categories

After reading through Chapter 8, the reader may feel encouraged to consider other, similar constructions. We now have $F_\sqrt{p}$, but what about $F_{3\sqrt{p}}$ or $F_{6\sqrt{p}}$? Maybe we can take a limit here or there and consider $F_{\infty\sqrt{p}}$? This is not directly relevant in the study of groups related to reductive groups, because of the contraints imposed by the combinatorics of root systems, but conceivably such constructions—in particular in conjunction with Weil restrictions—could produce other interesting objects.

In this Appendix we will briefly suggest a general approach to such constructions. Exclusively for this section, we will denote composition by concatenation in diagrammatical order, i.e. we write $fg$ for $g \circ f$.

Let us first make explicit some observations that were just below the surface throughout Chapter 8.

1https://mathoverflow.net/users/9062/bill-thurston
1. It seems natural to study a category of pairs \((\mathcal{C}, F)\) where \(\mathcal{C}\) is a category and \(F\) an endomorphism of \(\text{id}_\mathcal{C}\), with arrows between such pairs being functors \(H : \mathcal{C} \to \mathcal{D}\) such that \(H(F_x) = G_{H(x)}\) for every object \(x \in \mathcal{C}\)—in more technical terms, the whiskerings \(F \triangleleft H = H \searrow G\) agree. (See Proposition 8.2.5 or Definition 8.6.5.) For instance, consider the walking endomorphism \(\mathcal{N}\); this is a category with a single object \(\bullet\) and the monoid \(\mathbb{N}\) as endomorphisms \(\text{End}_\mathcal{N}(\bullet)\). Its identity functor has an endomorphism which is called the step \(s : 1_\mathcal{N} \to 1_\mathcal{N}\), defined by

\[
\begin{array}{c}
\bullet \\ \downarrow 1 \\
\bullet
\end{array}
\]

\(\begin{array}{c}
n \\
\downarrow 1 \\
\end{array}\)

This provides us with such a pair \((\mathcal{N}, s^n)\) for every natural number \(n\). We can then verify that

\[
\text{hom}((\mathcal{N}, s^0), (\mathcal{C}, F)) \simeq \mathcal{C}^F
\]

\[
\text{hom}((\mathcal{N}, s^1), (\mathcal{C}, F)) \simeq \mathcal{C}
\]

\[
\text{hom}((\mathcal{N}, s^2), (\mathcal{C}, F)) \simeq t\mathcal{C},
\]

where \(\mathcal{C}^F\) is the full subcategory of objects \(X\) such that \(F_X = \text{id}_X\) and \(t\mathcal{C}\) the twisted category from Definition 8.1.1. (Perhaps this suggests the better notation \(\mathcal{C}[\Phi]/(\Phi^2 = F)\)?) As a second example, we consider the hopping endomorphism \(\mathcal{H}\). This is a category with two objects \(\bullet\) and \(\circ\), with arrows \(\bullet \xrightarrow{\alpha} \circ\) and \(\circ \xrightarrow{\beta} \bullet\) and everything these arrows generate. Here too there is an endomorphism \(h\) of the identity functor completely determined by

\[
\begin{array}{c}
\bullet \\
\downarrow 1_{\alpha \beta} \\
\bullet
\end{array}
\]

\(\begin{array}{c}
\circ \\
\downarrow 1_{\beta \alpha} \\
\circ
\end{array}\)

\(\begin{array}{c}
\bullet \\
\downarrow 1_{\alpha \beta} \\
\bullet
\end{array}\)

And in this case \(\text{hom}((\mathcal{H}, h), (\mathcal{C}, F)) = m\mathcal{C}\), as introduced in Definition 8.1.2.

This makes certain observations easier. For instance, the reader can try to interpret some of the functors that we defined in Section 8.2 as coming from arrows between \((\mathcal{N}, s^1)\), \((\mathcal{N}, s^2)\) and \((\mathcal{H}, h)\).
2. One inconvenience with this notion arises as follows. The category \( t'\mathcal{C} \) is naturally endowed with an endomorphism \( \Phi \) of the identity functor, provided by the twisters. This tells the full story of \( t'\mathcal{C} \) in some sense and it would certainly be much better if we could write

\[
\text{hom}((\mathcal{N}, s^2), (\mathcal{C}, F)) = (t'\mathcal{C}, \Phi).
\]

But in the category \( m\mathcal{C} \), the situation is more complicated since we must incorporate information about the mixing maps into the picture. These produce a collection of maps \( \Phi_X : X \to \tau^*(X) \) which combine into natural transformations \( \Phi : \text{id}_{m\mathcal{C}} \to \tau^* \) and \( \tau^*\Phi : \tau^* \to \text{id}_{m\mathcal{C}} \). We see that somehow \( \tau^* \), \( \Phi \) and \( \tau^*\Phi \) must come from the endofunctor \( T : \mathcal{H} \to \mathcal{H} \) which is defined by \( T(\bullet) = \circ, T(\circ) = \bullet, T(\alpha) = \beta, T(\beta) = \alpha \), and the natural transformations \( u : \text{id}_\mathcal{C} \to T \) and \( v : T \to \text{id}_\mathcal{C} \) which satisfy \( uv = h \). Clearly, we need to incorporate this in the picture to tell the full story of \( m\mathcal{C} \).

3. Another intuition is that \( F \in \text{End}(\text{id}_\mathcal{C}) \) tears a hole in the category \( \mathcal{C} \). This became apparent already in Section 8.1 where we drew diagrams

\[
\bullet \overset{\alpha}{\leftarrow} \overset{\beta}{\rightarrow} \circ
\]

but had to warn the reader that this diagram does not commute but rather \( \alpha\beta = F_\bullet \) and \( \beta\alpha = F_\circ \), as if there was a hole in the middle of the diagram, preventing us from contracting paths. A related difficulty was encountered in Section 8.6, where we extended functors \( f^* \) and \( f_* \) to semi-linear maps. Somehow this keeps track of how many times a morphism has encircled such a hole, with the monoid \((\mathbb{N}, +)\) playing the role of a fundamental monoid underlying this phenomenon.

A similar situation occurs in semi-linear algebra when we are studying objects over a base object, say schemes \( X \) over a field \( K \), and suddenly become interested in morphisms \( X \to Y \) which are not linear over \( K \) but rather over a deeper lying object \( k \), say for a Galois extension \( K/k \). Every such morphism projects to an element of the Galois group \( \text{Gal}(K/k) \) which keeps track of the semi-linearity.
Although it is straightforward to generalize the notion of a category with endomorphism of the identity \((\mathcal{C}, F)\) to a category with a monoid \(M\) of endomorphisms of the identity functor, this cannot be the right approach for \(m\mathcal{C}\) or for the examples from Galois theory. (In fact \(\text{End}(\text{id}_\mathcal{C})\) is always a commutative monoid.) To include these cases, we will need a definition that is most elegantly stated in the language of (strict) 2-categories.

In general we will denote 2-categories with the Fraktur alphabet and in particular we will use \((\text{cat})\) to denote the 2-category of categories, functors and natural transformations.

**Definition B.1.1.** Let \(\mathfrak{m}\) be a 2-category with a single object \(\bullet\). An \(\mathfrak{m}\)-category is a category \(\mathcal{C}\) with a strict 2-functor \(\mathfrak{m} \rightarrow (\text{cat}) : \bullet \rightsquigarrow \mathcal{C}\).

We will not recall what a strict 2-functor is in general, but we will spell out in detail what it means to have a strict 2-functor \(\mathfrak{m} \rightarrow (\text{cat})\):

- The single object \(\bullet\) of \(\mathfrak{m}\) should be sent to a category \(\mathcal{C}\);
- every 1-cell \(u : \mathfrak{m} \rightarrow \mathfrak{m}\) should be sent to an endofunctor \(\mathcal{C} \rightarrow \mathcal{C}\);
- every 2-cell \(u \Rightarrow v\) between 1-cells should be sent to a natural transformation between the corresponding endofunctors;
- in such a way that these maps induce a morphism of 1-categories \(\text{End}(\mathfrak{m}) \rightarrow \text{End}(\mathcal{C})\).

Let us now explain the name \(\mathfrak{m}\)-category. If a monoid \(M\) acts on a set \(X\), we call \(X\) an \(M\)-set. There is then a morphism \(f : M \rightarrow \text{End}_{(\text{set})}(X)\) of monoids and thus a functor

\[\mathcal{M} \rightarrow (\text{set}) : \bullet \rightsquigarrow X,\]

where \(\mathcal{M}\) is the categorification of \(M\). So a \(\mathfrak{m}\)-category is just the 2-analogon of an \(M\)-set, with a ‘2-monoid’ acting on a ‘2-set’ (i.e. a category). We propose the following definition, which is a natural generalization of the idea of a crossed module.

**Definition B.1.2.** A crossed moduloid is a 4-tuple \((G, M, \partial, a)\) where

1. \(G\) and \(M\) are monoids;
2. \(M\) is a monoid;
3. \(\partial : M \rightarrow G\) is a homomorphism of monoids;
4. \( a : G \to \text{End}(M) \) is an action of \( G \) on \( M \), denoted \((m,g) \mapsto m^g\).

such that for all \( m,n \in M \) and \( g \in G \):

(M1) \( g\partial(m^g) = \partial(m)g \)

(M2) \( nm^\partial(n) = mn \)

(M3) \( G/\text{im} \partial \) is a group. (See further.)

If \( \partial \) and \( a \) are trivial, we simply write \((G,M)\).

The axioms (M1) and (M2) are the straightforward generalizations from the case of a crossed module; the axiom (M3) is new and, honestly, mainly a hunch. In fact, it has to be justified some more: it depends on the following proposition which observes that by axiom (M1), there is a monoid structure on \( G/\text{im} \partial \) and the last axiom demands that this is a group.

**Proposition B.1.3.** By axiom (M1), the relation on \( G \) given by

\[ g \sim h \iff (\exists u,v \in M)(g\partial u = h\partial v) \]

is a congruence relation.

**Proof.** Clearly the relation is symmetric and reflexive. For transitivity: if

\[ g\partial m = h\partial n \text{ and } h\partial n' = k\partial s, \text{ then} \]

\[ g\partial mn = h\partial n\partial n' = h\partial n'\partial(n^\partial n') = k\partial(sn^\partial n'), \]

Finally, we must show that \( \sim \) is a congruence relation. So suppose that

\[ g\partial m = h\partial n \text{ and } g'\partial m' = h'\partial n', \]

then we compute

\[ gg'\partial(m^g m') = g\partial(m)g'\partial(m') = h\partial(n)h'\partial(n') = hh'\partial(n'h'n'). \]
To a crossed moduloid \((G, M, \partial, a)\), we associate the following invariants:

- the group \(\pi_1 = G/\text{im} \partial\);
- the commutative monoid \(\pi_2 = \ker \partial\).

Special cases:

- For any monoid \(M\) and group \(G\) the trivial map \(\partial : M \to G : x \mapsto 1\) gives rise to a crossed moduloid with trivial action if (and only if) \(M\) is commutative. In that case \(\pi_1 = G\) and \(\pi_2 = M\).
- For any commutative monoid \(M\) and natural number \(n\) the map \(\partial : M \to M : x \mapsto nx\) gives rise to a crossed moduloid with \(G = M\), trivial action, \(\pi_1 = M/nM\) and \(\pi_2 = nM\).
- If \(G\) is an arbitrary group and \(N \trianglelefteq G\) a normal subgroup then \((G, N, \text{inc}, \text{Ad})\) is a crossed moduloid with action given by conjugation and \(\pi_1 = G/N\) and \(\pi_2 = 1\).

With a crossed moduloid we may form a 2-category as follows. First we construct the semi-direct product \(G \rtimes M\) in the usual way:

\[
(g, m)(h, n) = (gh, mh^n).
\]

We then define \(\mathfrak{c}(G, M, \partial, a)\): (i) there is a single object (or 0-arrow) \(\bullet\); (ii) the 1-arrows correspond to elements \(g : \bullet \to \bullet\) of \(G\) and (iii) the 2-arrows are given by \((g, m) : g \Rightarrow g\partial(m)\) for all \((g, m) \in G \rtimes M\).

The horizontal and vertical composition laws of 2-morphisms are given by:

\[
\begin{align*}
(g, m)(h, n) &= (gh, mh^n), \\
(g, m)(h, n) &= (gh, mh^n).
\end{align*}
\]
This relates to our examples as follows.

- A category $\mathcal{C}$ with an endomorphism $F$ of the identity functor is an $m$-category for $m = c(\mathbb{N}, \mathbb{N}, \text{id}, \text{id})$ via the 2-functor $m \to \text{cat}$ defined by

\[
\begin{array}{ccc}
\bullet & \xrightarrow{(a,b)} & \bullet \\
\downarrow & & \downarrow \\
\quad a+b & & \quad F^b \end{array}
\Rightarrow
\begin{array}{ccc}
\bullet & \xrightarrow{1_{\mathcal{C}}} & \bullet \\
\downarrow & & \downarrow \\
\quad \mathcal{C} & & \quad \mathcal{C}.
\end{array}
\]

In particular, this also holds for the twisted category $(t\mathcal{C}, \Phi)$.

More generally, a category $\mathcal{C}$ endowed with a monoid $M$ of endomorphisms of the identity functor becomes an $m$-category with invariants $\pi_1 m = 1$ and $\pi_2 m = M$.

- The mixed category $m'\mathcal{C}$ becomes an $m$-category for

\[
m = c(\mathbb{N}, \mathbb{N}, x \mapsto 2x, \text{id})
\]

via the 2-functor $m \to \text{cat}$ given by

\[
\begin{array}{ccc}
\bullet & \xrightarrow{(0,1)} & \bullet \\
\downarrow & & \downarrow \\
\quad 0 & & \quad 1_{m'\mathcal{C}} \\
\quad Y & & \quad \Phi \quad m'\mathcal{C}
\end{array}
\Rightarrow
\begin{array}{ccc}
\bullet & \xrightarrow{1_{m'\mathcal{C}}} & \bullet \\
\downarrow & & \downarrow \\
\quad m'\mathcal{C} & & \quad m'\mathcal{C},
\end{array}
\]

where we leave the other assignments to the reader.

- Consider a category $\mathcal{D}$ with an object $K$ and a subgroup $G \leq \text{Aut}(K)$. The *fairy* $\mathcal{C}$ is just the slice category over $K$ with $G$-semilinear arrows—in detail: the objects are the arrows $q_X : X \to K$ in $\mathcal{D}$ and the arrows $X \to Y$ are the pairs of arrows $(f, f^\sharp)$ such that $f^\sharp \circ q_X = q_Y \circ f$. (We denote such an arrow succinctly as $X \xrightarrow{f,f^\sharp} Y$.)

Then $\mathcal{C}$ acquires the structure of a $c(G, G)$ category via the
following 2-functor

\[
\begin{array}{ccc}
  m & \rightarrow & (m,n) \\
  \circ & \downarrow & \circ \\
  mn & \rightarrow & \mathcal{C}
\end{array}
\xrightarrow{\alpha_{(m,n)}}
\begin{array}{ccc}
  \mathcal{C} & \rightarrow & \mathcal{C}
\end{array}
\]

The functors \( F_m \), one for every \( m \in G \), are given by

\[
F_m : \mathcal{C} \rightarrow \mathcal{C} : X \xrightarrow{f} Y \xrightarrow{q_X} K \xrightarrow{f^2} K \quad \xrightarrow{\alpha_{(m,n)}} \quad X \xrightarrow{f} Y \xrightarrow{q_Y} K \xrightarrow{(f^2)_m} K,
\]

denoted more succinctly by

\[
F_m : \mathcal{C} \rightarrow \mathcal{C} : (X \xrightarrow{f} Y) \mapsto (X^m \xrightarrow{f_{(f^2)^m}} Y^m).
\]

The natural transformations \( \alpha_{(m,n)} : F_m \rightarrow F_{mn} \), one for every pair \( m, n \in G \), are given by \( (\alpha_{(m,n)})_X = (\text{id}_X, n) \):

\[
\begin{array}{ccc}
  X^m & \xrightarrow{f_{(f^2)^m}} & Y^m \\
  \downarrow^{(\text{id}_X, n)} & & \downarrow^{(\text{id}_Y, n)} \\
  X^{mn} & \xrightarrow{f_{(f^2)^{mn}}} & Y^{mn}
\end{array}
\]

So we hope that a careful study of \( m \)-categories might be of higher explanatory value than a straightforward generalization of our construction.

### B.2 Fields

We will now prove two propositions on twisted and mixed fields. We have two reasons for doing so. The first reason is that we believe these propositions can be of direct interest for anyone willing to undertake the study of groups and geometries of types \( ^2G_2 \) and mixed \( F_4 \) from a Galois cohomology point of view, for instance see [CD15] for a
B.2. Fields

wild occurrence of a mixed field extension. A second reason is that it provides a peek behind the curtains of what to expect from a twisted or mixed Galois theory.

In this section, we will use exponential notations such as \( x^\theta = \theta(x) \) and \( x^{\theta-1} = \theta(x)/x \).

### B.2.1 Twisted fields and \( p = 3 \)

Let us first investigate blended fields \((k, \theta)\) also known as fields with Tits endomorphism. Surprisingly at first, the underlying field is never algebraically closed. For \( p = 2 \), it is shown in [DSW] that the equation \( x^2 + x + 1 \) has no solutions. For other characteristics, we have:

**Proposition B.2.1.** If \( p > 2 \) then equations \( x^{\theta-1} = -1 \) and \( x^{p-1} = -1 \) have no solutions.

*Proof.* For the first part apply \( \theta \) to \( x^\theta = -x \) to obtain \( x^p = -x^\theta = x \). This implies that \( x \in \mathbb{F}_p \), but \( \theta \) acts trivially on the prime field and therefore \( x^{\theta-1} = 1 \). For the second part, observe that that \( p - 1 = (\theta - 1)(\theta + 1) \).

This shows that \( \theta \) cannot be extended to the algebraic closure \( k_a \), and not even to the separable closure \( k_s \). On the other hand, \( \theta \) can always be extended to the *perfect closure* \( k_p \).

**Proposition B.2.2.** There exists an algebraic field extension \( k_p/k \) such that \( k_p \) is perfect and \( \theta \) can be extended to \( k_p \).

*Proof.* Clearly \( \theta \) can be extended to \( k^{p^-n} \subset k_a \) for all \( n \in \mathbb{N} \), by the isomorphism \( k^{p^-n} \rightarrow k : x \mapsto x^{p^n} \), so \( \theta \) can be extended to \( k_p = k^{p^-\infty} = \cup_n k^{p^{-n}} \).  

It turns out that when \( p = 2 \) resp. \( p = 3 \), the unsolvability of the equation \( x^2 + x + 1 \) resp. \( x^2 = -1 \) is essentially the *only* obstruction for extending \( \theta \) to a quadratic extension. For \( p = 2 \) this is implicit in [DSW], so from now on we will focus on \( p = 3 \).
More precisely, we will show that for \( p = 3 \), \( \theta \) can be extended to a field \( K \) where there are only two classes of squares: the class of 1 and the class of \(-1\).

**Lemma B.2.3.** Let \( p = 3 \) and assume \( \delta \in k^\times \). Then exactly one of the following occurs.

- \( \delta = x^2 \) for some \( x \);
- there exists a field extension \( \ell/k \) of degree 4 such that \( \delta \) has a square root in \( \ell \) and \( \theta \) can be extended to \( \ell \);
- \( \delta = -x^{\theta - 1} \) for some \( x \) and there exists no field extension \( \ell/k \) such that \( \delta \) has a square root in \( \ell \) and \( \theta \) can be extended to \( \ell \).

**Proof.** Assume that \( \delta \) and \( \delta^\theta \) belong to a different square class. Let \( \ell = k(\sqrt{\delta}, \sqrt{\delta^\theta}) \),

then clearly \([\ell : k] = 4\). Extend \( \theta \) to \( \ell \) by setting

\[
\sqrt{\delta^\theta} = \sqrt{\delta^\theta} = \delta \sqrt{\delta} \quad \sqrt{\delta^\theta + 1} = \sqrt{\delta^\theta + 3} = \delta \sqrt{\delta^\theta + 1},
\]

then we verify that this gives rise to an endomorphism of \( \ell \). It is sufficient to verify on the basis \( 1, \sqrt{\delta}, \sqrt{\delta^\theta}, \sqrt{\delta^\theta + 1} \) that \((uv)^\theta = u^\theta v^\theta \) and this is a quick exercise since if \( u = 1 \) or \( v = 1 \) this is trivial so we must only verify:

- \( \sqrt{\delta^\theta} \sqrt{\delta^\theta} = \sqrt{\delta^\theta} \sqrt{\delta^\theta} = \delta^\theta = (\sqrt{\delta} \sqrt{\delta})^\theta \)
- \( \sqrt{\delta^\theta} \sqrt{\delta^\theta + 1} \sqrt{\delta^\theta + 3} = \delta^3 = (\delta^\theta)^3 = (\sqrt{\delta^\theta} \sqrt{\delta^\theta})^\theta \)
- \( \sqrt{\delta^\theta + 1} \sqrt{\delta^\theta + 3} = \delta^3 + 3 = \delta^\theta + 3 = (\delta^\theta + 1)^3 = (\sqrt{\delta^\theta + 1} \sqrt{\delta^\theta + 1})^\theta \)
- \( \sqrt{\delta^\theta} \sqrt{\delta^\theta + 3} = \delta^3 + 3 = \delta^\theta + 3 = (\delta^\theta + 1)^3 = (\sqrt{\delta^\theta + 1} \sqrt{\delta^\theta + 1})^\theta \)
- \( \sqrt{\delta^\theta} \sqrt{\delta^\theta + 1} \sqrt{\delta^\theta + 3} = \delta^3 \sqrt{\delta^\theta + 1} = \delta^3 \sqrt{\delta^\theta} = (\delta^\theta \sqrt{\delta})^3 = (\sqrt{\delta^\theta} \sqrt{\delta^\theta + 1})^\theta \)
- \( \sqrt{\delta^\theta} \sqrt{\delta^\theta + 1} \sqrt{\delta^\theta + 3} = \delta^3 \sqrt{\delta^\theta + 1} \sqrt{\delta} = (\delta \sqrt{\delta^\theta})^3 = (\sqrt{\delta^\theta + 1} \sqrt{\delta})^\theta \)

Otherwise, \( \delta \delta^\theta = x^2 \) for some \( x \). Then \( \delta^\theta + 1 = (x^{\theta - 1})^{\theta + 1} \). Applying \( \theta - 1 \) we get \( \delta^2 = (x^{\theta - 1})^2 \). So either

- \( \delta = x^{\theta - 1} \). But now, either \( x^\theta \) and \( x \) belong to the same square class, then \( \delta \) is a square, or they belong to a different class and the field extension \( k(\sqrt{x}, \sqrt{x^\theta}) \) can be constructed by the first item and does the job, since \((\sqrt{x^\theta} / \sqrt{x})^2 = \delta \).
• $\delta = -x^{\theta-1}$, and then assume $\delta = y^2$ in the extension field $\ell$. Then applying $\theta + 1$ to $y^2 = (y^{\theta+1})(\theta-1) = -x^{\theta-1}$ gives $(y^{\theta+1})^2 = x^2$ so $x = \pm y^{\theta+1}$. But then $x^{\theta-1} = y^2$ and thus $\delta = -x^{\theta-1} = -y^2$, contradiction.

**Proposition B.2.4.** There exists an algebraic field extension $K/k$ such that $\theta$ can be extended to $K$ and every element of $K^\times$ is either of a square or minus a square.

**Proof.** Iteratively apply the previous lemma to obtain a field where every element $\delta$ is either a square or of the form $\delta = -x^{\theta-1}$.

If $\delta$ is a square, say $\delta = x^2$, and then it is of the form $\delta = y^{\theta-1}$ with $y = x^{\theta+1}$.

If $\delta = -x^{\theta-1}$, then either $x$ is a square, say $x = y^2$ and then $\delta = -(y^{\theta-1})^2$ is minus a square, or $x$ is of the form $x = -y^{\theta-1}$ and thus $\delta = -(-y^{\theta-1})^{\theta-1} = -y^{\theta^2-2\theta+1} = -y^{4-2\theta} = -(y^{2-\theta})^2$ and $\delta$ is minus a square again.

**B.2.2 Mixed fields and $p = 2$**

Recall from Example 9.4.1.2 that a mixed field $(k, \ell, \kappa, \lambda)$ always originates from a field extension $\ell/k$ such that $\ell^p \subseteq k$. Let $L/\ell$ be another field extension and assume $K$ is a subfield of $L$ such that $L^p \subseteq K$, $\ell \subseteq L$ and $k \subseteq K$ then $M = (K, L)$ is an extension of the mixed field $m = (k, \ell)$; moreover it is easily verified that every extension of mixed fields arised this way. In particular, taking $L = K = \ell^n$ the algebraic closure of $\ell$, we obtain something that could be thought of as an algebraic closure of the field $m$ (in contrast with Proposition B.2.1).

From now on, let $p = 2$ and let us study étale algebras over a field of degree 2, by which we mean extensions $(K, L)/(k, \ell)$ such that $K/k$ and $L/\ell$ are étale algebras of degree 2. In Grothendieck’s Galois theory, these should correspond to sets of order 2 with a continuous action of the absolute Galois group, although it remains to be seen what this means for mixed fields. Recall the following fact:
Proposition B.2.5. The étale extensions of $k$ of degree 2 are classified by the elements of $\text{coker } \varphi = k / \text{im } \varphi$, where

$$\varphi : k \rightarrow k : u \mapsto u^2 + u.$$ 

In this correspondence the element $u \in k$ correspond to the extension $k[X]/(X^2 + X + u)$ and in particular the trivial element of $\text{coker } \varphi$ corresponds to $k \oplus k$.

The mixed analog of Proposition B.2.5 is Proposition B.2.6:

Proposition B.2.6. The étale extensions of $m = (k, \ell, \kappa, \lambda)$ of degree 2 are classified by the elements of $\text{coker } \tilde{\varphi} = (k \oplus \ell) / \text{im } \tilde{\varphi}$ where

$$\tilde{\varphi} : k \oplus \ell \rightarrow k \oplus \ell : (x, y) \mapsto (x + y^\lambda, x^\kappa + y).$$

Proof. Clearly every étale extension can be realised as

$$\ell(\sqrt{e}) \xrightarrow{\kappa} k(\sqrt{d})$$

$$\ell \xrightarrow{\lambda} k$$

Where we have used the notation $\ell(\sqrt{e})$ for the extension $\ell[X]/(X^2 + X + e)$. If we denote the extensions of $\kappa$ and $\lambda$ by the same letters, we must have

$$\lambda(\sqrt{e}) = x + x' \sqrt{d}, x, x' \in k$$

$$\kappa(\sqrt{d}) = y + y' \sqrt{e}, y, y' \in \ell$$

We may now apply $\kappa$ to the first equation, substitute the second and express the result with respect to the $\ell$-basis $1, \sqrt{e}$, to obtain $(x')^\kappa y' = 1$ and $x^\kappa + (x')^\kappa y = e$. Mutatis mutandis, we also have $(y')^\lambda x' = 1$ and $y^\lambda + (y')^\lambda x = d$. This implies that $x' = y' = 1$ and thus

$$e = x^\kappa + y$$

$$d = x + y^\lambda$$
So every element \((x, y)\) of \(k \oplus \ell\) determines a mixed étale extension, since it determines both \(e, d\) and \(\kappa, \lambda\). Two such elements determine the same extensions if and only if their difference \((s, t)\) satisfies

\[
s^\kappa + t = a^2 + a \\
s + t^\lambda = b^2 + b,
\]

for some \(a \in \ell\) and \(b \in k\). Applying \(\kappa\) to the second equation and adding to the first, we obtain \(t + t^2 = (a + b^\kappa) + (a + b^\kappa)^2\). This implies that \(t = a + b^\kappa + 0\) or \(t = a + b^\kappa + 1\). Analogously \(s = a^\lambda + b + 0\) or \(s = a^\lambda + b + 1\). It is also clear that either both solutions are \(+0\) or both are \(+1\). So, after relabeling \(a + 1\) by \(a\) in the latter case, we see that \(t = a + b^\kappa\) and \(s = a^\lambda + b\), so \((s, t) \in \text{im} \tilde{\varphi}\).

**Corollary B.2.7.** There is a bijective correspondence between étale \(k\)-algebras of degree 2 and étale \(m\)-algebras of degree 2 provided by:

\[
c\ker \varphi \to c\ker \tilde{\varphi} : u \mapsto (u, 0) \\
c\ker \tilde{\varphi} \to c\ker \varphi : (u, v) \mapsto u + v^\lambda.
\]

**Proof.** It is immediately verified that the maps are well defined and inverses of each other, using the identities \((a^2 + a, 0) = \tilde{\varphi}(a, a^\kappa)\) and \((b^\lambda, b) = \tilde{\varphi}(0, b)\).

\[\square\]

### B.3 Conjectural taxonomy of \(F_4\)

The confusion surrounding mixed and twisted groups in the literature culminates around the particularly interesting case of groups of absolute type \(F_4\). In this appendix, we attempt to clear up the confusion and conjecturally postulate a taxonomy for \(F_4\).

Let us first focus on the mixed groups. Recall from [Spr09, 17.5.2] that the exceptional group \(F_4\) can admit the possible forms \(F_{4,4}\), \(F_{4,1}\) and \(F_{4,0}\) over an arbitrary field \(k\). Applying the mixing functor from Section 8.2, we get the same groups, interpreted as mixed algebraic groups. We assign a a Dynkin diagram simply by drawing parallel Dynkin diagrams for the two fibres, i.e. we double up the standard
diagrams. We emphasize that these groups are nothing special, they are just the standard groups of type $F_4$, but seen as a visible mixed group. Such a group is defined over a visible field $(k, k, fr_k, id_k)$ and then by base change over other fields as well.

![Diagram of $F_{4,0}$](image1)

![Diagram of $F_{4,1}$](image2)

![Diagram of $F_{4,4}$](image3)

By mixing them the other way, or applying the functor $\tau^*$, we also get a class of corresponding anti-visible groups, defined over anti-visible fields and then by base change over other fields too.

If $p = 2$ however, there are extra invisible options. Thanks to the very special isogeny $F_4 \to F_4$, we can mix $F_4$ with itself in a non-trivial manner over a visible field $k$. (See Proposition 10.2.1.) The most straightforward case is that of $F_{4,4}$ where we mix two split groups. The resulting group is the group associated to a mixed building of type $F_4$ as defined in [Tit74]. That this group actually corresponds to those mixed groups introduced by Tits is precisely the content of Theorem 10.4.1 in the case $F_4$. We associate to this group the following diagram, which indicates that the mixing maps align the long roots of one $F_4$ with the short roots of the other.

![Diagram of $MF_{4,4}$](image4)

What if we try to mix non-split groups? There the situation gets more interesting. If the mixed base field is visible, this implies that one of the mixing maps must be linear. We suspect that this implies that one of the groups must be split as in Remark 10.5.2. This could give rise to the mixed Moufang sets of type $F_4$ from [CD15] and in principle there could also be a variant with an anisotropic $F_4$. So we get the following diagrams—where the $[*]$ means: hypothetical.
B.3. Conjectural taxonomy of $F_4$

If the field is non-visible, the condition that one of the groups must be split vanishes. A useful metaphor is that the non-perfect field $k^2 \subsetneq k$ generates a pool of non-splitness and by choosing an intermediary field $\ell$ strictly in between these extremes, both groups can tap into this pool which gives rise to extra possibilities where both components of the mixed group are non-split. This could give rise to the following diagrams, two of which are hypothetical and the latter of which we suspect is responsible for the mixed quadrangles of type $F_4$.

To figure out what twisted groups exist, we must ask ourselves which of the mixed groups admit twisted descent. Of course, the ground field over which they are defined must admit twisted descent but more importantly, both components must be isomorphic. This suggests three forms $^2F_{4,r}$, which arise by twisting $F_{4,r}$ for $r = 0, 1, 4$. The case $r = 4$ corresponds to the large Ree groups $^2F_4$ found by Ree [Ree61b], as we have shown in Theorem 10.3.1. The case $r = 1$ we conjecture to exist and correspond to the *doubly twisted Moufang sets of type $F_4$, introduced in [MM06] and studied thoroughly in [DSW]. Finally, the anisotropic $^2F_{4,0}$ is just hypothetical.
We will now provide a brief overview with some of the most important
definitions. This section contains nothing that is not widely known except for perhaps Section C.5.

As we mentioned in Section 7.4 there are certainly better references on
this material such as [Lei14] and [Mac71], but here we have collected
those facts which are most important for the rest of our work. Note
that we have chosen to ignore all foundational (set-theoretical) issues
throughout, although we might occasionally use words like *collection*
or *class* and (locally) *small* to minimize backlash from the more
informed readers, who should probably not be reading this Appendix.
To the other readers we will only issue a warning that in category
theory *small* does not mean finite, but it means *fits in a set*; in other
words *not small* means *too big to be a set* and how this should be
formalized depends on the underlying axiomatization of set theory.

C.1 Categories, functors, natural transformations

**Definition C.1.1.** A category \( \mathcal{C} \) consists of the following data: a
collection of objects \( \text{Ob}(\mathcal{C}) \); a collection of arrows \( \text{Ar}(\mathcal{C}) \); three maps
\( s \) (which stands for *source*), \( t \) (which stands for *target*) and \( \text{id} \) (which
stands for *identity*)

\[
\text{Ar}(\mathcal{C}) \xrightarrow{s} \text{id} \xleftarrow{t} \text{Ob}(\mathcal{C})
\]
Appendix C. Crash course in category theory

and a composition function
\[ \circ : \{(u, v) \in \text{Ar}(\mathcal{C}) \times \text{Ar}(\mathcal{C}) \mid s(u) = t(v)\} \rightarrow \text{Ar}(\mathcal{C}) \]

\[ (u, v) \mapsto u \circ v. \]

If a pair \((u, v)\) is in the domain of \(\circ\) we will say that \(u\) and \(v\) are composable and we will call \(u \circ v\) the composition of \(u\) and \(v\); more generally we call a tuple \((u_1, \ldots, u_n)\) of arrows composable if every two consecutive arrows are composable.

This data must satisfy the following conditions:

- \(s(\text{id}_X) = X\) and \(t(\text{id}_X) = X\) for every object \(X\);
- \(u \circ \text{id}_{s(u)} = u\) and \(\text{id}_{t(u)} \circ u = u\) for every arrow \(u\);
- \(s(u \circ v) = s(v)\) and \(t(u \circ v) = t(u)\) for every composable pair \(u, v\);
- \((u \circ v) \circ w = u \circ (v \circ w)\) for every composable triple \(u, v, w\).

Henceforth, if we say that \(f : X \rightarrow Y\) is an arrow, we mean that \(f\) is an arrow with \(s(f) = X\) and \(t(f) = Y\). The collection of all arrows between given objects \(X\) and \(Y\) is denoted by \(\text{hom}_{\mathcal{C}}(X, Y)\), although we will also write \(\text{hom}(X, Y)\) if no confusion is possible.

Example C.1.2. The main categories featuring in our work (apart from the newly defined ones) are:

- \((\text{set})\): the category of sets and functions;
- \((\text{grp})\): the category of groups and group homomorphisms;
- \((\text{ring})\): the category of unital commutative associative rings with ring homomorphisms;
- \((\text{sch})\): the category of schemes and morphisms of schemes.

Definition C.1.3. If \(\mathcal{C}\) is a category, then the opposite category \(\mathcal{C}^{\text{op}}\) denotes the category with the same classes of objects and arrows, but with the roles of \(s\) and \(t\) reversed; in particular the composition \(\circ^{\text{op}}\) is given by \(u \circ^{\text{op}} v = v \circ u\).

Definition C.1.4. An object \(X \in \text{Ob}(\mathcal{C})\) is called

- initial if for every other object \(Y\) there is a unique arrow \(X \rightarrow Y\);
- terminal if for every other object \(Y\) there is a unique arrow \(Y \rightarrow X\).
Lemma C.1.5. Let $\mathcal{C}$ be a category and let $E, E' \in \text{Ob}(\mathcal{C})$ both be initial resp. terminal objects, then there is a unique isomorphism $E \to E'$.

Proof. Consider the unique arrows $u : E \to E'$ and $v : E' \to E$ which exist because $E$ and $E'$ are initial or terminal. If we compose them in both orders, we find arrows $E \to E$ and $E' \to E'$. But there is a unique such arrow and the identity is such an arrow. Thus both compositions are the identity arrows on $E$ and $E'$.

In other words: initial and terminal objects are unique, up to a unique isomorphism. Therefore a common way to construct new concepts is by defining them to be the initial or terminal objects in a suitable category; this automatically makes them unique. A typical convention is to denote an initial object by $0$ and a terminal object by $1$. This notation is inspired by the category of sets, where the empty set is the unique initial object and every singleton is a terminal object.

Definition C.1.6. An arrow $f : X \to Y$ in a category $\mathcal{C}$ is called

- an **epimorphism** if for every pair of morphisms $g, g' : Y \to Z$ the equality $g \circ f = g' \circ f$ implies that $g = g'$.
- a **monomorphism** if for every pair of morphisms $g, g' : Z \to X$ the equality $f \circ g = f \circ g'$ implies that $g = g'$.
- an **isomorphism** if there is an arrow $g : Y \to X$ such that $f \circ g = \text{id}_Y$ and $g \circ f = \text{id}_X$.

Example C.1.7. In the category of sets, the epimorphisms are precisely the surjective maps and the monomorphisms are the injective maps. The same is true in the category $(\text{grp})$ of groups.

Example C.1.8. Every isomorphism is both an epimorphism and a monomorphism, but the converse is not true. For instance, the unique arrow $\mathbb{Z} \to \mathbb{Q}$ in the category $(\text{ring})$ is an epimorphism and a monomorphism, but not an isomorphism. This is also an example of a epimorphism which is not surjective on the underlying set.

We now come to the notion of a morphism between categories.
Definition C.1.9. Let $\mathcal{C}$ and $\mathcal{D}$ be categories. A functor $F : \mathcal{C} \to \mathcal{D}$ is a pair of maps $\text{Ob}(\mathcal{C}) \to \text{Ob}(\mathcal{D})$ and $\text{Ar}(\mathcal{C}) \to \text{Ar}(\mathcal{D})$, both denoted by $F$, such that

- $s(F(u)) = F(s(u))$ and $t(F(u)) = F(t(u))$ for every arrow $u$ in $\mathcal{C}$;
- $F(u \circ v) = F(u) \circ F(v)$ for every composable pair $u, v$ in $\mathcal{C}$;
- $F(\text{id}_X) = \text{id}_{F(X)}$ for every object $X$ in $\mathcal{C}$.

Sometimes a functor $\mathcal{C}^{\text{op}} \to \mathcal{D}$ is called a contravariant functor between $\mathcal{C}$ and $\mathcal{D}$, and the ordinary notion is then called a covariant functor.

Frequently we will denote a functor $F$ by an arrow $X \Rightarrow F(X)$ rather than an arrow “$\mapsto$”; the purpose of this change in notation is to inform the reader that although the functor has been defined on objects only and not on arrows, there is an obvious way to do so which we leave to the reader to figure out.

We stress that an assignment $X \Rightarrow F(X)$ should never be called a functor unless it can be extended to arrows.

Example C.1.10. Consider the following assignments from the category of groups to itself:

- $G \mapsto \text{Aut}(G)$ (the automorphism group);
- $G \mapsto \mathbb{Z}(G)$ (the center);
- $G \mapsto G/G'$ (the abelianization).

Only the third one gives rise to a functor.

Example C.1.11. If $\mathcal{C}$ denotes an arbitrary category then there always is a functor

$$\mathcal{C}^{\text{op}} \times \mathcal{C} \to (\text{set}) : (X, Y) \mapsto \text{hom}_\mathcal{C}(X, Y),$$

we leave as an exercise to the reader to work out the details. (Including the construction of the cartesian product of categories!)

Definition C.1.12. For a functor $F : \mathcal{C} \to \mathcal{D}$, we define the following properties:
• \( F \) is \textit{essentially surjective} if for every object \( Y \in \text{Ob}(\mathcal{D}) \), there is an object \( X \in \text{Ob}(\mathcal{C}) \) such that \( F(X) \cong Y \).

• \( F \) is \textit{full} resp. \textit{faithful} if for every pair of objects \( X, Y \in \text{Ob}(\mathcal{C}) \), the map

\[
\text{hom}_\mathcal{C}(X, Y) \to \text{hom}_\mathcal{D}(F(X), F(Y)) : u \mapsto F(u)
\]

is surjective resp. injective.

• \( F \) is \textit{constant} if \( F(X) = F(Y) \) for every \( X \in \text{Ob}(\mathcal{C}) \) and \( F(u) = \text{id}_{F(X)} \) for every arrow \( u : X \to Y \) in \( \mathcal{C} \).

A functor \( F \) is an \textit{equivalence} if it is full, faithful and essentially surjective. It is an \textit{embedding} if it is full and faithful.

\textbf{Example C.1.13.} The functor \((\text{ring})^{\text{op}} \to (\text{sch}) : R \mapsto \text{Spec}(R)\) is an embedding.

The collection of all functors between two categories can itself be endowed with the structure of a category. To make this precise we need a notion of morphisms between functors and these are called \textit{natural transformations}.

\textbf{Definition C.1.14.} Let \( F, G : \mathcal{C} \to \mathcal{D} \) be a pair of functors, then a \textit{natural transformation} \( \alpha : F \Rightarrow G \) is a collection of arrows in \( \mathcal{D} \) called the \textit{components} of \( \alpha \), one for every \( X \in \text{Ob}(\mathcal{C}) \) and denoted by \( \alpha_X : F(X) \to G(X) \), such that for every arrow \( f : U \to V \) in \( \mathcal{C} \) the following diagram commutes:

\[
\begin{array}{ccc}
F(U) & \xrightarrow{F(f)} & F(V) \\
\downArrow{\alpha_U} & & \downArrow{\alpha_V} \\
G(U) & \xrightarrow{G(f)} & G(V).
\end{array}
\]

The natural transformation is a \textit{natural isomorphism} if all components are isomorphisms.

Usually we will denote a natural transformation with just a single arrow \( \alpha : F \to G \); the underlying philosophy is that as we are making one construction on top of another, there is no point in keeping track of how high up we are. For instance, a morphism between group functors is technically a natural transformation but often it is irrelevant or
even confusing to point this out. When we do use the double arrow \( \alpha : F \Rightarrow G \) it is also an implicit warning that we specifically want to think of it as a morphism between functors.

**Definition C.1.15.** The hom-category \( \text{hom}(\mathcal{C}, \mathcal{D}) \) has as objects all functors \( F : \mathcal{C} \to \mathcal{D} \) and as morphisms between \( F \) and \( G \) all natural transformations \( \alpha : F \Rightarrow G \).

**Proposition C.1.16.** A functor \( F : \mathcal{C} \to \mathcal{D} \) is an equivalence if and only if there is a functor \( G : \mathcal{D} \to \mathcal{C} \) such that \( G \circ F \cong \text{id}_\mathcal{C} \) and \( F \circ G \cong \text{id}_\mathcal{D} \), where \( \cong \) denotes an isomorphism in the hom-category, i.e. a natural isomorphism.

### C.2 Diagrams, limits and continuity

**Definition C.2.1.** A functor \( D : \mathcal{D} \to \mathcal{C} \) is sometimes also called a diagram of shape \( \mathcal{D} \) in \( \mathcal{C} \); the diagram is small if \( \mathcal{D} \) is a small category. The diagram is said to commute if for every two objects in \( \mathcal{D} \), there is at most one arrow between them.

**Example C.2.2.** One should think of the category \( \mathcal{D} \) as a collection of placeholders, which must be filled with objects and arrows of \( \mathcal{C} \). Let us discuss in more detail the easy but extremely important case where \( \mathcal{D} \) is a category with three objects, the identity arrows and 2 more arrows, as depicted here:

![Diagram](https://via.placeholder.com/150)

A diagram of shape \( \mathcal{D} \) in a category \( \mathcal{C} \) is then just the information of three objects \( X, Y \) and \( S \) in \( \mathcal{C} \), together with two arrows:

![Diagram](https://via.placeholder.com/150)
Definition C.2.3. A cone over the diagram $\mathcal{D}$ is a constant functor $F : \mathcal{D} \to \mathcal{C}$ together with a natural transformation $\alpha^F : F \to \mathcal{D}$. A morphism between cones $F$ and $G$ is a natural transformation $\beta : F \to G$ such that $\alpha^G \circ \beta = \alpha^F$. A limit for the diagram $\mathcal{D}$ is a terminal object in the category of cones over $\mathcal{D}$, and it is denoted $\lim \mathcal{D}$ or also—by a common abuse of notation—$\lim_i X_i$ where the index $i$ runs over the objects of $\mathcal{D}$ and $X_i = \mathcal{D}(i)$.

A cocone over the diagram $\mathcal{D}$ is a constant functor $F : \mathcal{D} \to \mathcal{C}$ together with a natural transformation $\alpha^F : \mathcal{D} \to F$. A morphism between cocones $F$ and $G$ is a natural transformation $\beta : F \to G$ such that $\beta \circ \alpha^F = \alpha^G$. A colimit for the diagram $\mathcal{D}$ is an initial object in the category of cocones over $\mathcal{D}$, and it is denoted $\text{colim} \mathcal{D}$ or also—by a common abuse of notation—$\text{colim}_i X_i$ where the index $i$ runs over the objects of $\mathcal{D}$ and $X_i = \mathcal{D}(i)$.

Example C.2.4. (Continuation of Example C.2.2) A cone over this diagram is given by an object $C$ together with morphisms to each of the given objects, such that all the triangles which arise commute:

$$
\begin{array}{ccc}
C & \xrightarrow{\alpha_X} & X \\
\alpha_Y & & \alpha_S \\
\downarrow & & \downarrow \\
Y & \xrightarrow{\alpha_S} & S
\end{array}
$$

A limit $L = \lim \mathcal{D}$ for the diagram $\mathcal{D}$ is then a cone with the following property: whenever $C$ is a cone over the diagram $\mathcal{D}$, there is a unique arrow $u : C \to L$ such that that all the resulting triangles commute:

$$
\begin{array}{ccc}
C & \xrightarrow{\alpha_X^C} & L \\
\alpha_Y^C & & \alpha_S^C \\
\downarrow & & \downarrow \\
X & \xrightarrow{\alpha_S^L} & Y \\
\alpha_Y^L & & \alpha_S^L \\
\downarrow & & \downarrow \\
S & \xrightarrow{\alpha_S} & S
\end{array}
$$
For this specific diagram, the limit $L$ is called the fibered product $X \times_S Y$ of $X$ and $Y$ over $S$, which is an abuse of notation since it does not specify the arrows $X \to S$ and $Y \to S$. The maps $\alpha_X$ and $\alpha_Y$ are called the projections of the fibered product and the map $\alpha_S$ is usually not drawn, since it can be constructed from either projection by composing with the arrows to $S$.

**Definition C.2.5.** A category $\mathcal{C}$ is called (co)complete if it has all small (co)limits, i.e. every small diagram has a (co)limit. A functor $F : \mathcal{C} \to \mathcal{C}'$ between two categories is called (co)continuous if it preserves all (co)limits, i.e. whenever $D : D \to \mathcal{C}$ is a diagram in $\mathcal{C}$ and $F \circ D$ the corresponding diagram of the same shape in $\mathcal{C}'$ then

- (continuous) $F(\lim D) = \lim(F \circ D)$,
- (cocontinuous) $F(\colim D) = \colim(F \circ D)$.

In alternative notation, this becomes

- (continuous) $F(\lim_i X_i) = \lim_i F(X_i)$,
- (cocontinuous) $F(\lim_i X_i) = \lim_i F(X_i)$.

**Example C.2.6.** An important example of a continuous functor is the hom-functor. To illustrate this, let us continue our discussion from Example C.2.2 and Example C.2.4. Recall that we had a diagram $X \to S \leftarrow Y$ and a limit $L = X \times_S Y$, which means that for every choice of an object $C$ together with arrows $u$ and $v$ such that the following diagram (without the dashed arrow) commutes, there is a unique dashed arrow which makes the entire diagram commute.

![Diagram](image)

Conversely, given the dashed arrow we can reconstruct the arrows $u$ and $v$ by composing with $p_1$ and $p_2$. If we introduce the notation
\( \text{hom}_\mathcal{C}(C,U) = \mathbf{h}^C(U), \) where \( U \in \text{Ob}(\mathcal{C}) \) we have found that there is a bijection
\[
\{(u,v) \in \mathbf{h}^C(X) \times \mathbf{h}^C(Y) \mid q_x \circ u = q_y \circ v\} \to \mathbf{h}^C(X \times_S Y). \quad (\dag)
\]
\[(u,v) \mapsto u \times v.\]

We will call \( u \times v \) the pairing of \( u \) and \( v \). We may then interpret the left hand side of (\dag) as a fibered product in the category of sets, so we obtain a bijection
\[
\mathbf{h}^C(X) \times_{\mathbf{h}^C(S)} \mathbf{h}^C(Y) \to \mathbf{h}^C(X \times_S Y).
\]

But in fact \( \mathbf{h}^C \) is a functor
\[
\mathbf{h}^C : \mathcal{C} \to (\text{set}) : X \leadsto \text{hom}(C,X),
\]
and we have shown that this functor preserves the fibered product. The same argument will apply to diagrams of a different shape, which shows that the functor \( \mathbf{h}^C \) is in fact continuous.

**Example C.2.7.** If we want to form the product \( G \times H \) of two groups \( G \) and \( H \) in the category of groups, the underlying set is the product of these sets in the category of sets. This follows from the fact that the forgetful functor
\[
\text{forget} : (\text{grp}) \to (\text{set}) : (G, \cdot) \leadsto G
\]
is continuous. This in turn follows from the observation that this functor is naturally isomorphic to the functor
\[
\mathbf{h}^\mathbb{Z} : (\text{grp}) \to (\text{set}) : G \leadsto \text{hom}_{(\text{grp})}(\mathbb{Z},G),
\]
which is continuous by Example C.2.6.

So the fact that the group \( \mathbb{Z} \) gives rise to the forgetful functor is a somewhat special situation, and deserves a separate definition.

**Definition C.2.8.** An object \( X \in \text{Ob}(\mathcal{C}) \) is a generator for the category \( \mathcal{C} \) if the functor
\[
\mathbf{h}^X : \mathcal{C} \to (\text{set}) : Y \leadsto \text{hom}_\mathcal{C}(X,Y)
\]
is faithful.
C.3 Presheaves and the Yoneda embedding

**Definition C.3.1.** If $\mathcal{C}$ is an arbitrary category, then we call $\mathcal{C} = \text{hom}(\mathcal{C}^{\text{op}}, \text{(set)})$ the category of presheaves on $\mathcal{C}$. There is a functor

$$h : \mathcal{C} \to \mathcal{C} : X \mapsto h_X,$$

defined by setting

$$h_X : \mathcal{C}^{\text{op}} \to \text{(set)} : Y \mapsto h_X(Y) = \text{hom}(Y, X).$$

A presheaf $F \in \text{Ob}(\mathcal{C})$ is called *representable* if it is isomorphic to a presheaf of the form $h_X$ for some $X \in \text{Ob}(\mathcal{C})$; one says that $F$ is in the *essential image* of the functor $h$.

**Lemma C.3.2 (Yoneda.).** There is an isomorphism

$$\text{hom}_{\mathcal{C}}(h_X, F) \cong F(X), \text{ natural in } F \text{ and } X.$$

The *natural in $F$ and $X$* part is just a different way of saying that if we treat both sides as functors

$$\mathcal{C}^{\text{op}} \times \mathcal{C} \to \text{(set)} : (X, F) \mapsto F(X), \text{ and}$$

$$\mathcal{C}^{\text{op}} \times \mathcal{C} \to \text{(set)} : (X, F) \mapsto \text{hom}_{\mathcal{C}}(h_X, F),$$

then these are naturally isomorphic. In other words the collection of bijections should be interpreted as a collection of components of a natural transformation. Since a morphism in a product category

$$(d, e) : (D, E) \to (D', E') \in \text{Ar}(\mathcal{D} \times E)$$

can always be factorized into a pair of morphisms which are constant in either component

$$(D, E) \xrightarrow{d, \text{id}_E} (D', E) \xrightarrow{\text{id}_{D'}, e} (D', E'),$$

it is sufficient to verify that the bijection is natural in $F$ and $X$ separately.
**Proof.** We will only provide the bijection and leave the verification to the reader to complete or look up in any of the cited references such as [Lei14].

Consider a natural transformation $\alpha : h_X \Rightarrow F$. Its component at the specific object $X$ is the arrow $h_X(X) \rightarrow F(X)$. But the set $h_X(X)$ contains a special element, namely the identity arrow $X \rightarrow X$, and its image will be an element of $F(X)$. The map $\alpha \mapsto \alpha_X(id_X)$ provides one part of the bijection.

For the other part, assume that we have been given a prefered element $u \in F(X)$. Inspired by our construction in the previous paragraph, we want construct a natural transformation $\alpha$ for which $u$ becomes equal to $\alpha_X(id_X)$. To construct the entire natural transformation $\alpha$ amounts to constructing all its components $\alpha_Y$, using only the element $u \in F(X)$. To do this, let us consider the following diagram:

$$
\begin{array}{ccc}
  h_X(X) & \xrightarrow{\alpha_X} & F(X) \\
  \downarrow^{f} & & \downarrow^{F(f)} \\
  h_X(Y) & \xrightarrow{\alpha_Y} & F(Y)
\end{array}
$$

If we investigate specifically what happens with $id_X$ along both paths, we obtain the identity

$$\alpha_Y(f) = F(f)(\alpha_X(id_X)).$$

This inspires us in the context where we know only $u$ to make the definition

$$\alpha_Y : h_X(Y) \rightarrow F(Y) : f \mapsto F(f)(u).$$

The rest of the proof is then a matter of verifying that this indeed defines a functor, that both maps are inverses to each other, and that the resulting bijection is natural in both arguments. \(\square\)

**Corollary C.3.3** (Yoneda embedding). *The functor $h : \mathcal{C} \rightarrow \widehat{\mathcal{C}}$ is an embedding.*

(Recall that an embedding is a functor which is full and faithfull.)
Proof. If we apply Lemma C.3.2 to the case $F = h_Y$ we find a natural bijection

$$\hom_{\mathcal{C}}(X, Y) \cong \hom_{\mathcal{C}}(h_X, h_Y),$$

and if we go through the construction we see that the bijection is indeed the map $h$ on arrows.

By applying this to the dual category, we find a contravariant version of the Yoneda embedding.

**Proposition C.3.4** (Co-Yoneda lemma and Yoda embedding). In the hom-category $\mathcal{C}^{\text{op}} = \hom(\mathcal{C}, (\text{set}))$, we have the bijection

$$\hom_{\mathcal{C}^{\text{op}}}(h^X, F) \cong F(X), \text{ natural in } F \text{ and } X,$$

where

$$h^X : \mathcal{C} \to (\text{set}) : Y \leadsto h^X(Y) = \hom_{\mathcal{C}}(X, Y).$$

As a consequence, the functor $h^- : \mathcal{C}^{\text{op}} \to \hom(\mathcal{C}, (\text{set})) : X \leadsto h^X$ is an embedding.

The name Yoda embedding comes from the joke: contravariant, the Yoda embedding is, which is actually useful to remember.

## C.4 Adjunctions

**Definition C.4.1.** Let $F : \mathcal{C} \to \mathcal{D}$ and $G : \mathcal{D} \to \mathcal{C}$ be functors. Then there are two ways to construct a functor $\mathcal{C}^{\text{op}} \times \mathcal{D} \to (\text{set})$: one given by $(X, Y) \leadsto \hom_{\mathcal{D}}(F(X), Y)$ and one given by $(X, Y) \leadsto \hom_{\mathcal{C}}(X, G(Y))$. We will call $(F, G)$ an adjoint pair if there is a natural isomorphism $\alpha$ between these two functors:
One frequently phrases this as follows: there is a bijection

\[ \text{hom}_\mathcal{C}(X, G(Y)) \cong \text{hom}_\mathcal{D}(F(X), Y), \text{natural in } X \text{ and } Y. \]

The fact that \( F \) and \( G \) form an adjoint pair or adjunction is frequently denoted by the notation \( F \dashv G \), which suppresses \( \alpha \). We then call \( F \) the left adjoint and \( G \) the right adjoint of the pair. The induced maps are sometimes denoted by \( \flat \) and \( \sharp \), which further suppresses \( \alpha \) from the notation.

\[ \text{hom}_\mathcal{C}(X, G(Y)) \to \text{hom}_\mathcal{D}(F(X), Y) : x \mapsto x^\sharp \]
\[ \text{hom}_\mathcal{D}(F(X), Y) \to \text{hom}_\mathcal{C}(X, G(Y)) : y \mapsto y^\flat. \]

Let us state in full detail what it means that \( \alpha \) is a natural transformation:

Consider an arbitrary pair of arrows \( f : X' \to X \) in \( \mathcal{C} \) and \( g : Y \to Y' \) in \( \mathcal{D} \). Then for every arrow \( u : X \to G(Y) \) in \( \mathcal{C} \) and \( v : F(X) \to Y \), the following equations hold:

\[ [G(g) \circ u \circ f]^\sharp = g \circ u^\sharp \circ F(f) \]
\[ G(g) \circ v^\flat \circ f = [g \circ v \circ F(f)]^\flat \]

(\( \bullet \))

The reader is encouraged to verify this in detail from the definition of a natural transformation. The best way to understand these equations is probably by imagining a huge commuting diagram which consists of two parts, as depicted below. The rules for forming such a diagram are as follows:

- the entire diagram must commute;
- the right box is a commuting diagram in \( \mathcal{D} \);
- the left box is an arbitrary commuting diagram in \( \mathcal{C} \) on which \( F \) was applied;
- all other arrows must run from the left box to the right box.
Let us transform this diagram as follows:

- remove $F$ from the left box;
- apply $G$ to the right box;
- apply $\flat$ to all the connecting arrows.

On the given diagram, the result would look like this:

The given equations (●) then tell us that the resulting diagram commutes as well:

**Proposition C.4.2.** The former diagram is a commuting diagram, with in the left box $F$ applied to a commuting diagram if and only if the latter diagram is a commuting diagram, with in the right box $G$ applied to a commuting diagram.

**Proof.** Consider two paths in the latter diagram with the same first
vertex and the same last vertex. If both these vertices are contained in the left box, then both paths are the same because we assumed that the left box was a commuting diagram with $F$ applied to it; and if they are both contained in the right box then the resulting paths commute because $G$ is a functor.

Otherwise both paths start in the left box and end in the right box. But then they are both of the form

$G(g_1) \circ \ldots \circ G(g_i) \circ u \circ f_1 \circ \ldots \circ f_j = p$

and then we know that applying the given transformation to such a path will always result in $p^\flat$ precisely because ($\bullet$) holds.

The other implication holds *mutatis mutandis*.

We obtain interesting information by applying this formalism in non-obvious manners. For instance, consider the following diagram.

There is an obvious, but not very interesting way to apply the formalism by drawing a box around the two left nodes. But we can also draw the left box just around the top vertex! If we then apply the transformation—remove $F$ from one box, add $G$ to the other box, and apply $\flat$ to the connections—we obtain the following diagram:

This is a bit remarkable, because it means that if for every object $X \in \text{Ob}(\mathcal{C})$, we keep track of the important arrow

$\eta_X := \text{id}^\flat_{F(X)} : X \rightarrow G(F(X))$, 

then we can reconstruct all the maps \( u \mapsto u^{♭} \) by defining them as the composition depicted in the diagram. Similarly we may derive the diagram

\[
\begin{array}{ccc}
F(u) & \to & F(G(Y)) \\
\downarrow_{\id_G^{♭}} & & \downarrow_{\id_G^{♭}} \\
Y & \to & Y \\
\end{array}
\]

and this tells us that for every object \( y \in \text{Ob}(\mathcal{D}) \), there is an important arrow

\[
\varepsilon_Y := \id_G^{♭} : F(G(Y)) \to Y
\]

which permits to reconstruct the maps \( u \mapsto u^{♭} \).

**Proposition C.4.3.** The arrows \( \eta_X \) form the components of a natural transformation \( \eta : \id \to G \circ F \); similarly the arrows \( \varepsilon_Y \) form the components of a natural transformation \( \varepsilon : F \circ G \to \id \).

We leave the proof to the reader; it is an (easy) application of this adjunction formalism.

**Definition C.4.4.** The natural transformation \( \eta : \id \to G \circ F \) is called the *unit* of the adjunction \( F \dashv G \); the natural transformation \( \varepsilon : F \circ G \to \id \) is called the *co-unit* of the adjunction \( F \dashv G \).

It is possible to characterize an adjunction \( F \dashv G \) entirely in terms of the natural transformations \( \varepsilon \) and \( \eta \), this is the *unit-co-unit definition* of an adjunction.

**Proposition C.4.5.** Let \( F \dashv G \) be an adjoint pair of functors. Then \( F \) is cocontinuous and \( G \) is continuous.

**Proof.** Using continuity of the hom-functor in its second argument, we compute for every object \( X \) and diagram \((Y_i)_i\):

\[
\text{hom}(X, \lim_i G(Y_i)) \cong \lim_i \text{hom}(X, G(Y_i)) \\
\cong \lim_i \text{hom}(F(X), Y_i) \\
\cong \text{hom}(F(X), \lim_i Y_i)
\]
\[ \cong \text{hom}(X, G(\lim_{i} Y_{i})). \]

Therefore if we denote \( U = \lim_{i} G(Y_{i}) \) and \( V = G(\lim_{i} Y_{i}) \), then the presheaves \( h_{U} \) and \( h_{V} \) are isomorphic and thus, because the Yoneda embedding is an embedding, \( U \cong V \). The other claim can be proven similarly, \textit{mutatis mutandis}.

\[ \square \]

**Example C.4.6.** In Example C.2.7 we saw that the forgetful functor

\[ \text{forget} : (\text{grp}) \to (\text{set}) : (G, \cdot) \rightsquigarrow G \]

which maps a group to its underlying set, is continuous. Let

\[ \text{free} : (\text{set}) \to (\text{grp}) : X \rightsquigarrow \text{free}(X) \]

be the functor which assigns to every set the free group \( \text{free}(X) \) (also often denoted by \( F(X) \)) constructed in the usual manner as reduced words in the alphabet \( X \cup X^{-} \) with concatenation as operation. Note that this construction comes equipped with a monomorphism of sets

\[ \eta_{X} : X \to \text{forget}(\text{free}(X)). \]

The universal property of a free group over a set \( X \) states precisely that for every group \( G \) there is a bijection

\[ \text{hom}_{(\text{set})}(X, \text{forget}(G)) \cong \text{hom}_{(\text{grp})}(\text{free}(X), G) \]

\[ u \mapsto u^{\sharp} \]

\[ v^{\flat} \leftarrow v \]

with the property that \( \text{forget}(u^{\sharp}) \circ \eta_{X} = u \):

\[
\begin{array}{ccc}
X & \xrightarrow{u} & G \\
\downarrow{\eta_X} & & \downarrow{\text{forget}(u^{\sharp})} \\
& \text{forget(} \text{free}(X) \text{)} & \\
\end{array}
\]

One can then verify that this bijection is natural in \( X \) and \( G \), which means that there is an adjunction

\[ \text{free} \dashv \text{forget.} \]
Thus by Proposition C.4.5, the forgetful functor is continuous. In fact, we learn a little more: since the functor free is also cocontinuous, we see that the coproduct of free groups can be obtained as follows

$$\text{free}(X) \sqcup \text{free}(Y) = \text{free}(X \sqcup Y),$$

i.e. we recover the fact that the coproduct of free groups is their free product. Note that the co-unit of the adjunction is a map

$$\text{free} \circ \text{forget} : G \to \text{free}$$

which is surjective, so this says that every group has a presentation.

**Example C.4.7.** The Spec-functor and the global section functor

$$\text{Spec} : (\text{ring})^{\text{op}} \to (\text{sch}) : R \mapsto \text{Spec}(R)$$

$$\Gamma : (\text{sch}) \to (\text{ring})^{\text{op}} : X \mapsto \mathcal{O}_X(X)$$

determine an adjunction $\Gamma \dashv \text{Spec}.$

### C.5 Endomorphisms of the identity functor

**Definition C.5.1.** Consider categories $\mathscr{C}, \mathscr{D}, \mathscr{E}$ with functors $F, G, H$ and a natural transformation $\alpha$ as in on of the following diagrams:

\[
\begin{array}{ccc}
\mathscr{C} & \xrightarrow{F} & \mathscr{D} \\
\downarrow G & & \downarrow H \\
\mathscr{E} & & \text{resp.}
\end{array}
\]

Then the left resp. right whiskering of $\alpha$ and $F$ are the respective natural transformations, given by providing their components:

$$F \circ \alpha : G \circ F \implies H \circ F$$

$$\alpha \circ F : F \circ G \implies F \circ H$$

$$(F \circ \alpha)_X = \alpha_{F(X)}$$

$$(\alpha \circ F)_X = F(\alpha_X).$$

Since we will be dealing with endomorphisms of the identity functor in our work, let us provide the following proposition; the main purpose of it is to bound the collection of endomorphisms of the identity functor in some way.
If $\alpha : \text{id}_C \Rightarrow \text{id}_C$ is an endomorphism of the identity functor, then it is clear that for every object $X$, there is a map $\alpha_X \in \text{End}_C(X)$ and thus we obtain a map (in fact a morphism of monoids) 

$$\text{End}(\text{id}_C) \to \text{End}(X) : \alpha \mapsto \alpha_X.$$ 

So one way to construct endomorphisms of $\text{id}_C$ would be to see how an endomorphism of $X$ can be extended. The trouble with that is that a given endomorphism $X \to X$ could correspond to 0, 1 or many such natural transformations $\alpha$, so it seems hopeless to reconstruct $\alpha$ from $\alpha_X$. But we will see in Corollary C.5.3 that if $X$ is a generator for the category, the map is injective and the given endomorphism $X \to X$ extends to at most one such $\alpha$.

**Proposition C.5.2.** Consider the categories $\mathcal{B}$, $\mathcal{C}$, $\mathcal{D}$ and $\mathcal{E}$ with functors $E$, $F$, $G$ and $H$:

$$\mathcal{B} \xrightarrow{E} \mathcal{C} \xleftarrow{G} \mathcal{D} \xrightarrow{F} \mathcal{E}.$$ 

If $E$ is essentially surjective and $F$ is faithful, then the map 

$$\text{hom}(G, H) \to \text{hom}(F \circ G \circ E, F \circ H \circ E) : \alpha \mapsto E \circ \alpha \circ F$$ 

between sets of natural transformations is injective.

**Proof.** Assume that $E \circ \alpha \circ F = E \circ \beta \circ F$ for $\alpha, \beta \in \text{hom}(G, H)$. Then $F(\alpha_{E(X)}) = F(\beta_{E(X)})$ for all objects $X \in \text{Ob}(B)$ and thus, since $F$ is faithful, $\alpha_{E(X)} = \beta_{E(X)}$ for all objects $X$; but then $\alpha = \beta$ since $E$ is essentially surjective. \(\square\)

**Corollary C.5.3.** If $X \in \text{Ob}(\mathcal{C})$ is a generator for the category $\mathcal{C}$, then there is an injective map 

$$\text{End}(\text{id}_C) \to \text{End}_C(X).$$

**Proof.** If $X$ is a generator, then $h^X$ is faithful. So set $\mathcal{B} = \mathcal{C} = \mathcal{D}$, $\mathcal{E} = (\text{set})$, $E = G = H = \text{id}_C$ and $F = h^X$ in Proposition C.5.2 to obtain an embedding 

$$\text{End}(\text{id}_C) \to \text{End}(h^X),$$
and note that by the Yoneda lemma $\text{End}(h^X) = \text{End}_\mathcal{C}(X)$. □

**Example C.5.4.** In the category of $R$-algebra’s, where $R$ is an arbitrary ring of characteristic $p$, the object $R[x]$ is a generator. An endomorphism of $R[x]$ is completely determined by the image of $x$, which is a polynomial $f$. This tells us that the only candidates for endomorphisms of the identity functor are given by

$$\alpha_X : A \to A : u \mapsto f(u).$$

This polynomial must then be additive, multiplicative and map $1$ to itself, and with some effort one can conclude that the only possibilities are the powers of the Frobenius $f(u) = u^p$. 
Één van de hoekstenen van de algebraïsche groepentheorie is de structuurtheorie van de half-enkelvoudige groepen over een algebraïsch gesloten veld, die we grotendeels aan Chevalley danken. De theorie is daarna snel uitgebreid naar een theorie over (samenhangende) reductieve groepen over willekeurige velden, en zelfs over een willekeurig basisschema door vele anderen, zoals Borel, Tits, en de auteurs van [SGA3]. Maar in de loop van de tweede helft van de 20ste eeuw werden bij een aantal gelegenheden groepen ontmoet die nauw verwant lijken aan reductieve groepen, maar op een vreemde, exotische manier.

De eerste keer dat zo’n ontmoeding plaatsvond was omstreeks 1960. Terwijl Suzuki een klasse eindige enkelvoudige groepen classificeerde, ontdekte hij een nieuwe klasse, nu bekend als de Suzukigroepen. Zijn ontdekking was de prelude op de ontdekking van een meer algemene constructie door Ree later dat jaar die ook andere, gelijkaardige klassen groepen produceerde: de getwiste Chevalley groepen. Wat later—vermoedelijk omstreeks 1970—bestudeerde Tits reductieve groepen met zijn gebouwetheorie. Toen hij een bepaalde klasse van gebouwen classificeerde, ontdekte hij dat hoewel de meeste van die gebouwen afkomstig waren van reductieve groepen, er ook gebouwen waren die weliswaar verwant waren met reductieve groepen, maar niet op een rechtstreekse manier: dit waren de gemixte groepen en gebouwen. In 1997 vervolledigde Weiss de classificatie van een andere klasse van gebouwen en ontdekte groepen die nog vreemder zijn, maar nog steeds herkenbaar als verre neven van de reductieve groepen. Tot slot, omstreeks 2010 ontdekten Conrad, Gabber and Prasad, als deel van hun classificatie en structuurtheorie voor pseudo-reductieve groepen dat—de naam suggereert het al!—de meeste van die groepen nauw verwant zijn met reductieve groepen. Maar ook hier zijn er een paar
bizarre snuiters die eerder van ver verwant zijn.

In al deze gevallen onderscheiden we twee gemeenschappelijke fenome-
nen. De eerste observatie is dat de combinatoriek van wortelsystemen
met wortels met twee verschillende lengtes nodig is om de constructie
te laten slagen; met andere woorden een Dynkindiagram van type $B_n,
C_n, F_4$ of $G_2$ speelt een belangrijke rol. De tweede observatie is dat
de constructies bepaalde ingrediënten vereisen die zeer typisch zijn
voor ‘wiskunde in positieve karakteristiek’, zoals het Frobeniusendo-
morfisme van een veld of algebraïsche groep, en soms hangen ze ook
af van de aanwezigheid van inseparable velduitbreidingen.

Deze twee fenomenen zijn uiteraard nauw verwant. Bijvoorbeeld:
de verhouding van de wortels in het wortelsysteem houdt in dat de
karakteristiek van het veld 2 of 3 moet zijn. Desondanks was het ons
uitgangspunt dat om het inzicht in deze groepen te verdiepen, we
dezelfde fenomenen zo goed als mogelijk moeten ontwarren. Doorheen
het grootste deel van ons werk hebben we ons daarom geconcentreerd
op het tweede aspect, de ‘wiskunde in positieve karakteristiek’. Het is
daarna vrij eenvoudig om de combinatoriek van wortelsystemen via de
achterdeur binnen te smokkelen, door het bestaan van de zogenaamde
zeer speciale isogenieën te veronderstellen. Maar in onze aanpak is
dit eerder een nevengedachte.

Laten we nu de hoofdideeën achter onze aanpak samenvatten. Tra-
ditioneel denkt men over een algebraïsche groep als een groepsobject
in de categorie van de schema’s. Maar als we ons beperken tot de
categorie van schema’s in een vaste positieve karakteristiek dan is er
een bijzonderheid in de vorm van de absolute Frobenius. We kunnen
deze gebruiken om nieuwe, nauw verwante, categorieën te definiëren:
de categorieën van de getwiste en gemixte schema’s. Er zijn dan
ook groepsobjecten in deze categorieën en dit is waar die exotische
abstract en algebraïsche groepen vandaan komen.

Om meer in detail te kunnen treden, moeten we in herinnering brengen
dat een gewone algebraïsche groep over een veld $k$ een geassocieerde
puntefunctor $K \rightarrow G(K)$ bezit die een abstracte groep $G(K)$ produ-
ceert voor elke $k$-algebra $K$. Deze procedure, maar uitgevoerd in
de context van getwiste en gemixte groepsschema’s, produceert een
groot aantal abstracte groepen uit getwiste en gemixte ringen en dit zijn de groepen waar we naar op zoek waren.

De categorie van getwiste schema’s die we geconstrueerd hebben ligt een beetje dieper dan de categorie van de schema’s, en kan worden gezien als de categorie van schema’s over $\text{Spec } F_{\sqrt{p}}$. (Het veld met $\sqrt{p}$ elementen!) Er bestaat een inbedding$^1 \ m : X \rightarrow mX$ die een gewoon schema$^2$ omzet in een getwist schema. Met een typisch misbruik van notatie noteren we voor de eenvoud vaak $mX$ gewoon met $X$; in het bijzonder heeft de notatie “$\text{Spec } F_p$” nu ook de betekenis van een getwist schema. Als we de functor $m$ toepassen op een schema $X$ samen met het unieke structuur morfisme $X \rightarrow \text{Spec } F_p$, bekomen we een getwist schema dat gedefinieerd is over $F_p$.

Om dit wat minder abstract te maken beschouwen we een gewone algebraïsche groep $G$ over $F_p$ samen met een homogene ruimte $\mathcal{H}$ en een afbeelding $\pi : G \rightarrow \mathcal{H}$, allen gedefinieerd over $F_p$. Dit wordt gewoonlijk voorgesteld met het volgende diagram van schema’s.

![Diagram](image)

Maar als we dit inbedden in de categorie van de getwiste schema’s, zien we dat er vele andere dingen gebeuren waar we ons eerder niet van bewust waren:

![Diagram](image)

$^1$D.i. een getrouwen (maar niet noodzakelijk volle) functor.

$^2$Deze functor wordt eigenlijk genoteerd met $\delta \circ m$ in Sectie 8.1 maar we boeten hier wat in aan precisie om de helderheid van de uiteenzetting te bevorderen.
Voor een aantal buitengewoon merkwaardige objecten die gedefinieerd zijn over $\mathbf{F}_{\sqrt{p}}$ maar niet over $\mathbf{F}_p$. Bijvoorbeeld hebben we een object getekend dat suggestief met $2^2G$ genoteerd werd, en een object $\text{Spec} \mathbf{F}_{\sqrt{p}}^3$. Maar er is nog een tweede verrassing: er zijn getwiste schema’s die gedefinieerd zijn over $\mathbf{F}_p$, maar niet afkomstig zijn van een gewoon schema! Deze denkbeeldige of onzichtbare objecten zijn voorgesteld door de $\mathfrak{S}$ op het diagram en zij zijn verantwoordelijk voor de gemixte groepen van Tits. Dus als we de getwiste schema’s over $\mathbf{F}_p$ voortaan gemixte schema’s noemen, hebben we net geobserveerd dat er gemixte schema’s zijn die geen gewone schema’s zijn.

Er zijn een aantal belangrijke subtiliteiten om voor uit te kijken. Bijvoorbeeld, als $X \to \text{Spec} \mathbf{F}_p$ een gemixt schema is, dan is het structuurmorfisme uniek als een morfisme van gemixte schema’s maar niet als een morfisme van getwiste schema’s. Bijvoorbeeld binnen de gemixte schema’s is er precies één morfisme van $\text{Spec} \mathbf{F}_p$ naar zichzelf; maar als een getwist schema zijn er twee! Anders gezegd heeft de automorfismegroep binnen de categorie van getwiste schema’s orde 2; met andere woorden “$\text{Gal}(\mathbf{F}_p/\mathbf{F}_{\sqrt{p}}) \cong C_2$”.

Onze hoofdstellingen zijn drie toepassingen die tonen hoe de vreemde groepen die we aan het begin van deze uiteenzetting geïntroduceerd hebben in deze context meer natuurlijk opduiken; we vatten dit nu gauw samen. Stelling 10.3.1 zegt dat de getwiste groepen van Suzuki en Ree ontstaan als groepsobjecten in de categorie van getwiste schema’s. Stelling 10.4.1 zegt dat de gemixte groepen van Tits ontstaan als groepsobjecten in de categorie van de gemixte schema’s. Bovendien zijn de getwiste en gemixte groepen nauw met elkaar verwant: door het veld uit te breiden langs de uitbreiding $\mathbf{F}_p/\mathbf{F}_{\sqrt{p}}$ wordt een getwiste groep omgezet in een gemixte groep; en er is een afdalingscriterium Propositie 8.3.6 dat gebruikt kan worden om uit te zoeken welke getwiste groepen op deze manier ontstaan uit een gemixte groep. De exotische groepen van Conrad, Gabber and Prasad zijn ook nauw verwant met deze onzichtbare gemixte reductieve groepen: onze stelling Stelling 10.5.1 stelt dat ze ontstaan als Weilbeperkingen van gemixte reductieve groepen. Het voordeel hier is dubbel: enerzijds wordt de standaardconstructie voor pseudo-reductieve groepen hierdoor nog meer alomtegenwoordig, anderzijds wordt de classificatie
van sferische gebouwen hierdoor meer uniform.
Bibliography


