Chevalley Group Schemes As Varieties Over The Field Of One Element

by

Andrew Langworthy

A thesis submitted to

The University of Birmingham

for the degree of

Master Of Research

School of Mathematics

The University of Birmingham

September 2013

Abstract

The "field of one element" is a concept first suggested by J. Tits in 1957. It has been worked on and redefined many times over the past fifty years; in this paper we consider varieties over a quadratic extension of this field, introduced by C. Soulé and then refined by Connes and Consani. We follow the Connes and Consani paper "On The Notion Of Geometry Over $\mathbb{F}_1$" and present their results along with the necessary introduction to Chevalley groups and algebraic geometry.

Acknowledgements

I would like to thank my supervisor Dr Corneliu Hoffman for his continual support and guidance whenever it was needed during this project, and Dr Kay Magaard for his help and patience with all my questions. My thanks also extend to Prof. Chris Parker for his impromptu word counts and reassurance.

Contents

Chapter 1

Introduction

Here we will give a brief introduction to the structure of the paper, as well as to the field of one element. We begin with a familiar definition.

Definition 1.1. *A field $\mathbb{F}$ is a set with two binary operations, $+$ and $\cdot$, with identities 0 and 1, such that $\mathbb{F}$ and $\mathbb{F}\backslash\{0\}$ are abelian groups under $+$ and $\cdot$ respectively. We must also have that $\forall a, b, c \in \mathbb{F}$, $(a+b)\cdot c = a\cdot c + b\cdot c$, i.e. that multiplication distributes over addition. Lastly, we require that $0 \neq 1$.*

This final necessity, that $0 \neq 1$, seems very arbitrary, so we may ask what would happen if we let $0 = 1$. Well, consider a structure F such that this is the case, and pick some $a \in F$. Then we have that

$$(1+0)a = (1+1)a, \quad \text{and so}$$

$$a = a + a,$$

which means that a is the additive identity. This gives $a = 0 = 1$, and so this "field" has only one element.

Definition 1.2. *We define $\mathbb{F}_1$ to be the object described above; the unique one-element construction satisfying all the field axioms, but has $0 = 1$. We call this the field of one element.*

However, this is quite a naïve view of things since this is just the trivial ring, and so does not have some of the "nice" properties of fields. For this reason, constructions of $\mathbb{F}_1$ tend to be a little more abstract. However, this definition suffices to give a feel for $\mathbb{F}_1$. For the rest of the paper, we shall call $\mathbb{F}_1$ a field, even though classically there is obviously no such thing.

It is worth noting here that because of the non-concrete nature of $\mathbb{F}_1$, many people have looked at it in many different ways. In this introduction we draw from a few of these, but in the main body of the text we focus only on the viewpoint of Soulé in 1999, as used by Connes and Consani in [3].

We now give a few basic examples to give a feel of the mathematics of the field of one element, but before we dive into them it may be useful to see what q-theory is. Firstly, we have

$$\lim_{q\to 1}\frac{1-q^n}{1-q}=n,$$

which we can see by noting $1-q^n=(1-q)(1+q+...+q^{n-1})$. This then gives us a "$q$-like" version of n, known as the q-analog of n.

Definition 1.3. *The q-analog of n, written as $[n]_q$ is given by*

$$[n]_q=\frac{1-q^n}{1-q}.$$

We can redefine many familiar integer concepts in terms of q-analogs; for example, we define

$$\binom{n}{k}_q=\frac{[n]_q!}{[n-k]_q![k]_q!}$$

as a q-analog of the binomial coefficient, called the Gaussian coefficient. Given a vector space V of dimension n over the finite field $\mathbb{F}_q$, this calculates the number of k-dimensional vector subspaces of V. The obvious thing to do now, as we are considering a one-element field, is to see what happens when we let q approach 1 as a limit. Well, we have that

$$\lim_{q\to 1}\binom{n}{k}_q=\binom{n}{k}$$

by the basic properties of a limit. Note that as we are working over just one element, a vector space V over $\mathbb{F}_1$ has dimension equal to its order. The number of k-dimensional subspaces is $\binom{n}{k}$, but this is the same as the number of k-element subsets. Thus, a vector space over $\mathbb{F}_1$ is just a set, with subspaces being precisely subsets.

Now, what would a linear transformation do? Well, since there is no underlying structure on V, a linear transformation from V to another vector space over $\mathbb{F}_1$ is just a set map. A vector space isomorphism is more interesting though, since this is a $1-1$ V-map, and so is a permutation of V as

a set.

We have been talking in terms of linear algebra, looking at vector spaces and their morphisms, but it has turned out that over $\mathbb{F}_1$ these concepts are just those of sets and subsets. This is a general trend; linear algebra over $\mathbb{F}_1$ is just combinatorics of sets.

Another way of looking at $\mathbb{F}_1$ is to consider the general linear group on $\mathbb{F}_1$, which we denote by $GL_n(1)$. For $GL_n(q)$, the general linear group over the finite field of order q, we have a counting function $N(q)$ that counts its order. We know that

$$N(q) = \prod_{i=0}^{n-1} (q^n - q^i).$$

However, we can rewrite this in terms of our new q-analog notation. A little bit of algebra shows us that

$$N(q) = [n]_q!(q-1)^n q^{\binom{n}{2}}.$$

Now, we can divide this by $q-1$ to the power of the order of vanishing of $N(q)$ at $q=1$, i.e. divide by $(q-1)^n$. This gives us

$$\lim_{q \to 1} \frac{N(q)}{(q-1)^n} = n!,$$

which is the order of S_n. We use this to think of $GL_n(1) \cong S_n$.

We now consider a maximal torus (a group isomorphic to the Cartesian product of the ground field with itself j times) of GL_n, which is just its diagonal matrices. The normaliser of this group is the group of generalized permutation matrices, which are matrices like permutation matrices, but have any non-zero field element in place of their "1"s. Their quotient is called a Weyl group, which we discuss in a slightly different guise in section 3.1, and in this case is equal to S_n. As we have seen, $GL_n(1) = S_n$ and this suggests there is some sort of correspondence between the two, and this is correct. In fact, we may consider simple algebraic groups over $\mathbb{F}_1$ as Weyl groups.

Now we have seen this, we can branch out and perhaps consider field extensions of $\mathbb{F}_1$. To do this we think of them as pointed sets.

Definition 1.4. *A pointed set is a pair (A, a) for some set A and $a \in A$. A pointed set map between (A, a) and (B, b) is a map $f : A \to B$ such that $f(a) = b$.*

We can then look at the field extension of $\mathbb{F}_1$ of degree n for any $n \in \mathbb{N}$. We denote this object by $\mathbb{F}_{1^n}$.

Definition 1.5. *We define the field extension of* $\mathbb{F}_1$ *of degree* n *as*

$$\mathbb{F}_{1^n} = \mu_n \cup 0, \tag{1.1}$$

where $\mu_n \cup 0$ *is the group of* n^{th} *roots of unity, together with a point* 0*. We see this as a pointed set about* 0*.*

So what about vector spaces over this extension? Well, each vector v determines a set $\{\varepsilon_n^i v : 1 \leq i \leq n\}$, where ε_n is a primitive root. So, the vector space is a set upon which the group μ_n acts freely. In particular, a vector space of dimension d is just a set of size dn, together with "base" point. This can help us consider vector spaces over finite fields as vector spaces over some extension of $\mathbb{F}_1$. Indeed, if $q \equiv 1 \bmod n$, then we know that $\mu_n \leq \mathbb{F}_q^*$ as a group. Then we may consider $\mathbb{F}_q^* \cup 0 \cong \mathbb{F}_q$ as a pointed set and we have that $\mathbb{F}_q$ is an $\mathbb{F}_{1^n}$ vector space since μ_n acts freely upon it by multiplication. It will have dimension $\frac{q-1}{n}$. In particular, this gives us that $\mathbb{F}_q$ is an $\mathbb{F}_{1^n}$ algebra since we have a multiplication in $\mathbb{F}_q$.

What else do we study in this paper? Well, we give an introduction to algebraic geometry, giving it, to begin with, a primarily classical approach by motivating our constructions under the premise of finding zeros of polynomial equations. After this, we consider affine schemes, which are defined to be the spectrum of a unital ring, i.e. the set of prime ideals of a ring. We then give the definition of a group scheme, which are the structures of interest in chapter 4. Even a very brief overview of algebraic geometry will give much more information than we give here, as we have had to skip the theory of curves and projective spaces entirely, amongst other things.

We also consider Lie algebras, which are vector spaces with a binary Lie bracket, and subgroups of their automorphisms, specifically the Chevalley groups. Historically, the study of Lie algebras was a direct means to study Lie groups, which are of interest for many reasons but in particular the finite simple groups of Lie type are one of the four types of finite simple group in the well known classification of such groups. The Chevalley groups are groups with a very nice structure, and most of our discussion of them will be building up to investigate their Bruhat decomposition. This gives a canonical representation of every group element, which is vital in defining Chevalley group schemes as varieties over $\mathbb{F}_1$.

The language of [3] is written category theoretically, and so we also give a (very) short introduction in the form of appendix A, which is actually closer to a list of definitions that are required to understand

chapter 4.

We have tried to include as many proofs as possible, but due to space constraints, many have been left out. We have tried to make sure that those left out (that are not beyond the scope of this paper) are those that are unenlightening or unwieldy.

This paper will follow a few main sources; chapter 2 follows Ueno [15] for the first section, and then follows Gathmann [8] for section 2.2. Section 3.1 follows Humphreys [10], section 3.3 and appendix B follow Erdmann and Wilson [7] and 3.4 follows Carter [1]. Chapter 4 mirrors the Connes and Consani paper [3].

Chapter 2

Algebraic Geometry

Throughout this section, unless explicitly stated, we consider K to be an algebraically closed field.

2.1 Preliminaries

In this section we give a brief introduction to classical algebraic geometry. Consider a family of polynomials

$$\mathscr{F} = \{f_i(x_1, \ldots, x_n) : i \leq m\}$$

in n variables, with coefficients in K. We may then define an algebraic set.

Definition 2.1. *An algebraic set in K^n of a family of polynomials $\mathscr{F}$ is the set of solutions to all the polynomials in $\mathscr{F}$. We denote this set by $V(\mathscr{F})$.*

With this definition, we can now induce a topology onto K^n, by defining the algebraic sets of K^n to be the closed sets. We call this topology the Zariski topology, and for K^n with such a topology, we write $\mathbb{A}^n_K$. This is known as the n-dimensional affine space.

Definition 2.2. *We denote the ideal generated by a set X as $I(X)$. For our purposes, $X \subseteq \mathbb{A}^n$ and this set is given by*

$$I(X) = \{f \in K[x_1, \ldots, x_n] : f(x) = 0 \ \forall x \in X\}.$$

Note that this is indeed an ideal in ring-theoretic language since if $f \in I(X)$ and $g \in K[x_1, \ldots, x_n]$, then $fg(x) = f(x)g(x) = 0$ for all $x \in X$.

Definition 2.3. *An ideal I of a ring R is called a prime ideal if and only if for all $r, s \in R$, we have that if $rs \in I$ then either $r \in I$ or $s \in I$. Every ideal I has an associated radical $\sqrt{I} := \{r \in R : r^n \in I \text{ for some } n \in \mathbb{N}\}$. If $I = \sqrt{I}$ then I is called radical.*

Now, in classical algebraic geometry we also consider the polynomial ring $K[x_1, \ldots, x_n]$, which gives us more information about our families of equations defining algebraic sets. In particular, each family generates an ideal in this ring, which gives us the first result of the section.

Lemma 2.4. *Let I be the ideal generated by a family of equations $\mathscr{F}$ as in definition 2.1. Then in $\mathbb{A}^n_K$, we have that $V(I) = V(\mathscr{F})$.*

Proof. Firstly note that for any polynomial p, $V(\mathscr{F} \cup \{p\}) \subseteq V(\mathscr{F})$, and hence $V(I) \subseteq V(\mathscr{F})$. Now, take some $g \in I$. Then g is of the form

$$g = \sum_{i \in I} r_i f_i$$

for some $r_i \in K[x_1, \ldots, x_n]$ and $f_i \in \mathscr{F}$. Then if $a = (a_1, \ldots, a_n) \in V(\mathscr{F})$ then $f_i(a) = 0$ for all f_i, and so $g(a) = 0$, giving us that $V(I) \supseteq V(\mathscr{F})$, and so $V(I) = V(\mathscr{F})$ as required. □

We now give some basic properties of ideals of polynomial rings.

Proposition 2.5. *Let I, J be ideals of the polynomial ring $R \subseteq K[x_1, \ldots, x_n]$. Then*

1. $V(I) \cup V(J) = V(I \cap J)$,
2. $V(I) \subseteq V(J) \Rightarrow \sqrt{I} \supseteq \sqrt{J}$.

Proof. 1. Consider $a = (a_1, \ldots, a_n) \in V(I) \cup V(J)$. Then a is a solution to all polynomials in I or all polynomials in J. Without loss of generality, assume it is a solution to polynomials in I. Then since $I \cap J \subset I$, we have $a \in V(I \cap J)$. Now assume that $a \in V(I \cap J)$. Note a exists, since $0 \in V(I \cap J)$, so it must be the solution to some polynomial in $I \cup J$ and hence must be in $V(I \cup J)$.

2. It suffices to prove that $V(I) = V(\sqrt{I})$, since then we have

$$\begin{aligned} V(I) \subset V(J) &\Rightarrow V(\sqrt{I}) \subset V(\sqrt{J}), \\ &\Rightarrow \sqrt{I} \supset \sqrt{J}. \end{aligned} \tag{2.1}$$

Note that since $\sqrt{I} \supset I$, we have already that $V(\sqrt{I}) \subset V(I)$, so consider some $a \in V(I)$, and some $f \in \sqrt{I}$. Then $f^r \in I$ for some $r \in \mathbb{N}$, so $f^r(a) = 0$, so $f(a) = 0$ and $V(\sqrt{I}) \supset V(I)$ as required. □

Lemma 2.6. *If K is a finitely generated field over another field k, then K is algebraic over k.*

Proof. Omitted. □

Theorem 2.7. *Let I be an ideal of a polynomial ring R not containing the identity. Then $V(I) \neq 0$.*

Proof. Note that it suffices to show this for $I = \mathfrak{M}$ a maximal ideal, since any ideal is contained inside some maximal one, and if $\mathfrak{M}$ is maximal and $J \subset \mathfrak{M}$ is an ideal, then $V(\mathfrak{M}) \subseteq V(J)$.

Since $\mathfrak{M}$ is maximal, $K[x_1, \ldots, x_n]/\mathfrak{M}$ is a field. Obviously this contains K, and so by lemma 2.6, we see that $K[x_1, \ldots, x_n]/\mathfrak{M} \cong K$. The isomorphism will map x_i to some member of K, call it a_i. But then $x_i \equiv a_i \mod \mathfrak{M}$, and so $x_i - a_i \in \mathfrak{M}$, which gives $(a_1, \ldots, a_n) \in V(\mathfrak{M}) \neq \emptyset$ as required. □

This result is often called the Weak Hilbert Nullstellensatz.

Theorem 2.8 (Hilbert's Nullstellensatz)**.** *Let J be an ideal of a polynomial ring R. Then $I(V(J)) = \sqrt{J}$.*

Proof. We consider the first implication, that being that $\sqrt{J} \subseteq I(V(J))$. So take some $r \in \sqrt{J}$. Then $r^n \in J$ for some $n \in \mathbb{N}$. Now, consider some root x of r^n. Then $r(x)^n = 0$, which gives us that $r(x) = 0$. As $x \in V(J)$ and this is true for all such x, we have $r \in I(V(J))$ as required.

Now, we need that $\sqrt{J} \supseteq I(V(J))$. So, this means that for some $f \in I(V(J))$, we need that $f^m \in J$ for some $m \in \mathbb{N}$. We consider a new variable, call it x_0, and a new ideal, generated by $1 - x_0 f$ and J. Call this new ideal J'. Now, assume $V(J') \neq \emptyset$. Then some $a = (a_1, \ldots, a_n) \in V(J')$ and so $a \in J$ by definition of J'. Since $f \in J$, $f(a) = 0$. However, this is a contradiction since $1 - x_0 f(a) = 1 \neq 0$. Hence we must have $V(J') = \emptyset$. We can now use the weak Nullstellensatz and infer that $1 \in V(J')$ and so $V(J') = K[x_0, \ldots, x_n]$.

So, this means that we can write 1 as a sum of polynomials in J and $1 - x_0 f$ to give

$$1 = g_0(1 - x_0 f) + \sum_{i=1}^{d} g_i f_i$$

for $g_i \in K[x_0, \ldots, x_n]$, $f_i \in J$. Now, x_0 is a variable, so we may substitute it for $1/f$. Now we multiply by some power of f to remove all instances of f from the denominator of terms of the g_i. Then we have

$$f^m = \sum_{i=1}^{d} g_i' f_i \in K[x_1, \ldots, x_n],$$

as required. □

Corollary 2.9. *Let K be a field and consider the polynomial ring $K[x_1, \ldots, x_n]$. Then $\mathfrak{m}$ is a maximal ideal of the ring if and only if $\mathfrak{m}$ is of the form $(x_1 - a_1, x_2 - a_2, \ldots, x_n - a_n)$ for some $a_1 \ldots a_n \in K$.*

Proof. Omitted. □

We will now define a classical algebraic variety. For this, we need the concept of an irreducible set. This is just a set that cannot be written as the union of two closed sets in $\mathbb{A}^n_K$.

Definition 2.10. *Consider some irreducible set $V \subseteq \mathbb{A}^n_K$. If it is closed (i.e. of the form $V(I)$ for some ideal I), then we call it an affine algebraic variety.*

Note that every element of $\mathbb{A}^n_K$ is a solution of the zero polynomial, and so we may write $\mathbb{A}^n_K = V((0))$. It turns out that this itself is an affine algebraic variety, as shown by the following lemma.

Lemma 2.11. *An algebraic set X is irreducible if and only if the associated ideal $I(X)$ is a prime ideal.*

Proof. Omitted. □

Proposition 2.12. *We have*

1. $X = X_1 \cup X_2 \implies I(X) = I(X_1) \cap I(X_2)$,
2. $I(X) = I(X') \implies X = X'$.

Proof. 1. Consider some $X = X_1 \cup X_2$. Then if $f \in I(X)$ then $f(x) = 0$ for all $x \in X$. In particular, $f(x) = 0$ for $x \in X_1$, and so $f \in I(X_1)$. Similarly $f \in I(X_2)$, which gives one inclusion. Now let $f \in I(X_1) \cap I(X_2)$. Then f is zero on X_1 and X_2 and so is zero on their union, completing the proof.

2. Let $I(X) = I(X')$. Then consider the function

$$f(x) = \prod_{a \in X} (x - a). \tag{2.2}$$

Obviously this is in $I(X) = I(X')$, but since it is not zero on any points other than those in X, we must have $X' \subseteq X$. By symmetry then, we are done. □

We can now move onto more algebraic terminology, in particular that of morphisms and coordinate rings.

Definition 2.13. *The coordinate ring of an algebraic set Y is the ring $K[Y] := K[x_1, \ldots, x_n]/I(Y)$.*

Now, consider some set map ρ between two algebraic sets V and W lying inside $\mathbb{A}_K^n$ and $\mathbb{A}_K^m$ respectively. We say that ρ is a morphism if for $v = (v_1, \ldots, v_n) \in \mathbb{A}_K^n$ then $\rho(v) = (f_1(v_1, \ldots, v_n), \ldots, f_m(v_1, \ldots, v_n))$ where the f_i are polynomials for all $1 \leq i \leq m$.

One may guess that there is some correspondence between morphisms of algebraic sets and homomorphisms of the associated coordinate rings. This is indeed the case; in fact there is a natural bijection between them.

Let φ be a morphism $\varphi : V \to W$ with corresponding polynomials $f_1, \ldots, f_m$. Then we define a map $\overline{\varphi} : K[W] \to K[V]$ by $\overline{\varphi}(\theta) = \theta \circ \varphi + I(V)$. Note that if $a = (a_1, \ldots, a_n) \in V$, then $\varphi(a) \in W$, and so if $\theta \in I(W)$ then $\overline{\varphi}(\theta)(a) = 0$. This shows that the quotient map is indeed in $K[V]$. The fact that this map is a homomorphism follows from the distributive and associative laws of composition of functions, and is easily checked.

We claimed that this correspondence was bijective, and so now we must show that for every homomorphism of coordinate rings there is a corresponding morphism of algebraic sets. To see this, let $\Phi : K[W] \to K[V]$ be a homomorphism. We claim that Φ is of the form $\overline{\varphi}$ for some morphism φ. Take some representative f_i of the image under Φ of the equivalence class of x_i. Then we claim that we may take $\varphi = (f_1, \ldots, f_m)$ as our morphism.

Indeed, by construction φ is defined by polynomials, so all we now need show is that $\varphi(v) \in W$ $\forall v \in V$. So, let $g \in I(W)$, so $\Phi(g)(x) = 0$ in $K[V]$. But Φ is a homomorphism, so $g(\Phi)(x) = 0$, but $\Phi(x_i) = f_i$, so $g(f_1(x), \ldots, f_m(x)) = 0$, which gives us finally that $g(f_1(x), \ldots, f_m(x)) \in I(V)$, and so in particular $g\varphi$ vanishes on V, so for $v \in V$, $\varphi(v) \in V(I(W)) \subseteq W$, completing the claim. We can give this result as the following lemma.

Lemma 2.14. *Given two algebraic sets V and W, there is a bijective correspondence between homomorphisms of the form*

$$f : K[W] \to K[V]$$

and morphisms of the form

$$f' : V \to W.$$

2.2 Schemes

We now move on to the study of schemes. Affine schemes are a natural generalisation of affine varieties, and general schemes are constructed by "glueing" affine schemes together. In fact, up to equivalence

of categories, affine varieties are affine schemes. Schemes are more general than varieties in that while affine varieties correspond to certain types of ring (those of the form $K[V]$), affine schemes correspond to commutative rings, a much larger collection. For this reason, in this section we consider all rings to be commutative and unital unless stated otherwise.

Definition 2.15. *A mapping $f : \mathbb{A}_K^n \to \mathbb{A}_K^m$ is called regular at a point a if there exists an open set U_a containing a such that $f(u) = (f_1(u), \ldots, f_m(u))$ $\forall u \in U_a$, with f_i a polynomial for $1 \leq i \leq m$.*

We call a function regular if it is regular at all points.

Definition 2.16. *Let R be a commutative unitary ring. Then R is called local if there exists a unique maximal ideal $I \subseteq R$. We define*

$$\mathcal{O}_{v,V} = \{\frac{f}{g} : f, g \in K[V] \text{ and } f, g \text{ are regular at } v\}, \tag{2.3}$$

and call this the local ring of V at v. We define the ring of regular functions of V by

$$\mathcal{O}_V := \bigcap_{v \in V} \mathcal{O}_{v,V}.$$

Now, we consider the localisation of rings. This is a process analogous to creating the rational numbers from the integers. Take some arbitrary ring R and $D \subseteq R$ closed under multiplication. We call this a multiplicative subset of R. Then define a new set $S := \{r/d : r \in R, d \in D\}$. Now define an equivalence relation $\sim$ on S given by

$$\frac{r}{d} \sim \frac{s}{e} \Leftrightarrow \exists f \in D \text{ such that } (re - sd)f = 0. \tag{2.4}$$

It is easy to see that $\sim$ is an equivalence relation. We call the set equivalence classes of S under $\sim$ with multiplicative subset D the localisation of S at D, and we denote this by $D^{-1}S$. We define addition and multiplication in the same way as they are defined on $\mathbb{Q}$. Note that if D is multiplicative, then $1 \in D$, since if $u \in U$ then $uu^{-1} = 1 \in U$. So there is an additive identity $0/1$ and a multiplicative identity $1/1$.

As we have said, $\mathbb{Q}$ is a localisation, in particular the localisation of $\mathbb{Z}$ at $\mathbb{Z}\backslash\{0\}$. In this example, the constant f from equation 2.4 can always be taken to be 1. Note that R embeds naturally into $D^{-1}R$ by $r \mapsto r/1$. Also note that $\mathcal{O}_{v,V}$ can also be thought of as the localisation of the polynomial ring $K[V]$ at the maximal ideal $I(v)$.

We now look at sheaves and presheaves. These will be an integral part of the definition of a scheme later on.

Definition 2.17. *Let X be a space with topology $\mathcal{T}$. A presheaf $\mathcal{F}$ on X is a pair of collections*

$$\mathcal{F} := (\{\mathcal{F}(U) : U \in \mathcal{T}\}, \{\rho_{V,U} : U, V \in \mathcal{T} \text{ and } U \subseteq V\}),$$

such that $\mathcal{F}(U)$ are rings, whose elements we call sections, and $\rho_{V,U}$ is a ring homomorphism from $\mathcal{F}(U)$ to $\mathcal{F}(V)$ satisfying the following three conditions.

We need that $\mathcal{F}(\emptyset) = 0$, $\rho_{U,U}$ is identity for all U, and if we have three open sets $U \subset V \subset W$, then the restriction maps compose as one would expect; $\rho_{W,V} \circ \rho_{V,U} = \rho_{W,U}$.

Definition 2.18. *A sheaf is a presheaf with a "glueing" property. Let $\mathcal{F}$ be a presheaf on X and assume $U \subseteq X$ is open with a cover $\{U_i\}_{i\in\mathbb{N}}$ and for every ring $\mathcal{F}(U_i)$ take an arbitrary element f_i. Further assume that if for every pair U_i and U_j, we have that f_i restricted to $U_i \cap U_j$ is equal to f_j under the same restriction. We call $\mathcal{F}$ a sheaf if we get a unique $f \in \mathcal{F}(U)$ that for all i is equal to f_i restricted to U_i.*

Given a topological space X with two sheaves $\mathcal{F}$ and $\mathcal{G}$, we define a morphism $\psi : \mathcal{F} \to \mathcal{G}$ to be a set of ring homomorphisms $\psi(U) : \mathcal{F}(U) \to \mathcal{G}(U)$ for every open $U \subseteq X$ such that for every other open set $V \subseteq X$, the diagram

$$\begin{array}{ccc} \mathcal{F}(U) & \xrightarrow{\psi(U)} & \mathcal{G}(U) \\ \Big\downarrow{\scriptstyle \rho_{V,U}} & & \Big\downarrow{\scriptstyle \rho_{V,U}} \\ \mathcal{F}(V) & \xrightarrow{\phi(U)} & \mathcal{G}(V) \end{array}$$

commutes.

Example 2.19. *Take X an affine variety and let $U \subseteq X$ be open and consider the ring of regular functions on U, written as $\mathcal{O}_X(U)$. Consider such rings for all open U. We have the obvious restriction map between $\mathcal{O}_X(U)$ and $\mathcal{O}_X(V)$ where $V \subseteq U$. It can be seen that this gives us a presheaf, and in fact a sheaf, since the functions in the ring are nice enough due to regularity. We call this sheaf the structure sheaf on X, and write $\mathcal{O}_X$.*

Definition 2.20. *Let X and Y be topological spaces and $\theta : X \to Y$ a continuous function. Then given a sheaf $\mathcal{F}$ on X, we define the direct image sheaf $\theta_*\mathcal{F}$ to be the sheaf of rings $\mathcal{F}(\theta^{-1}U)$ such that*

$U \subseteq Y$ is open, with restriction maps $\rho_{\theta^{-1}(V),\theta^{-1}(U)}$ for every inclusion of open sets $V \subseteq U$.

Definition 2.21. *Let X be a topological space and $\mathcal{F}$ a sheaf on X. Then we call the pair $(X, \mathcal{F})$ a ringed space. Given another ringed space $(Y, \mathcal{G})$, the morphisms between them are pairs (θ, ψ), where $\theta : X \to Y$ is a continuous map and ψ is a morphism of sheaves from $\mathcal{G}$ to $\theta_* \mathcal{F}$.*

We can now give the definition of a stalk. This is basically the set of open sets with sections modulo an equivalence relation.

Definition 2.22. *Let X be a topological space, $x \in X$ a point of X and $\mathcal{F}$ a presheaf on X. Consider pairs of the form (U, ρ) where $U \subset X$ is open and $\rho \in \mathcal{F}(U)$. Then say that $(U, \rho) \sim (U', \rho')$ if there exists some open $V \subseteq U \cap U'$ such that $\rho\big|_V = \rho'\big|_V$. The stalk of $\mathcal{F}$ at x is the set of all these equivalence classes. Each individual class is called a germ.*

We give an example of a stalk, the stalk of the structure sheaf $\mathcal{O}_X$ at a point x. In fact, this gives us a familiar object; it is precisely $\mathcal{O}_{x,X}$, the local ring of X at x.

Proposition 2.23. *If X an affine variety and $x \in X$, then the stalk of $\mathcal{O}_X$ at x is the local ring $\mathcal{O}_{x,X}$.*

Proof. Omitted, see [13], p. 29. □

We now take a closer look at morphisms of varieties. This will allow us to generalise a little the concept of an affine variety, a generalisation which we will use later when looking at schemes.

Definition 2.24. *Consider two ringed spaces $(X, \mathcal{O}_X)$ and $(Y, \mathcal{O}_Y)$, whose structure sheaves are comprised of K-valued functions. Then a map $f : X \to Y$ is called a morphism if for any $\varphi : U \to K$ for $U \subseteq Y$ open, we have that $\varphi f : f^{-1}(U) \to K$ is a set theoretic function. We call this the pullback of φ. Also, we require f to be continuous, and pullback regular functions to regular functions.*

An isomorphism is a morphism with a two-sided inverse.

Now, to any finitely generated K-algebra $k[x_1, \ldots, x_n]/I$ we can associate the affine variety generating I, by lemma 2.14. However, there are no intrinsic zero sets to work with.

Definition 2.25. *A ringed space $(X, \mathcal{O}_X)$ is called an affine variety if*

1. *X is irreducible.*
2. *$\mathcal{O}_X$ is a sheaf and is made up of functions taking values on K.*
3. *The space is isomorphic, as a ringed space, to an affine variety in the sense of definition 2.10.*

Now we have covered affine varieties, we can extend our discussion to more general constructions, that of non-affine varieties. We use similar glueing requirements as in the definition of a sheaf.

Definition 2.26. *A prevariety is a ringed space $(X, \mathcal{O}_X)$ such that*

1. *X is irreducible.*
2. *$\mathcal{O}_X$ is a sheaf and is made up of functions taking values on K.*
3. *There is a finite open cover of X $\{U_i\}_{1 \leq i \leq m}$ such that $(U_i, \mathcal{O}_X\big|_{U_i})$ is an affine variety for all i.*

Example 2.27. *In this example we construct $\mathbb{P}^1$, the projective line, which is an example of a prevariety, by glueing two affine varieties together. In this case we will use two copies of the affine line $\mathbb{A}^1$. We "glue" them together using the map $x \mapsto \frac{1}{x}$. Every element has an inverse, other than zero, so this gives an extra point, "$\frac{1}{0}$", so naturally we label this ∞. The resulting space looks like $\mathbb{A}^1 \cup \{\infty\}$. We do not discuss projective spaces in this paper, but any introductory book on algebraic geometry will cover the topic.*

We will not rigorously show how this glueing works as that is not the topic of this project.

Definition 2.28. *Let $(X, \mathcal{O}_X)$ be a prevariety. We say it is a variety if for any prevariety $(Y, \mathcal{O}_Y)$ and morphisms f_1 and f_2 from Y to X, the set $\{x \in X : f_1(x) = f_2(x)\}$ is closed in Y.*

Now, this is not a particularly nice definition; having to deal with all prevarieties is not easy! Fortunately, the next proposition allows us to do away with this. For this we require the diagonal morphism. This is a morphism from X to $X \times X$ defined by $d : x \mapsto (x, x)$. The image $d(X)$ is called the diagonal of X.

Proposition 2.29. *X is a variety if and only if the diagonal is closed.*

Proof. Omitted. □

Corollary 2.30. *Any affine variety is a variety.*

Proof. Consider some $X \subseteq \mathbb{A}^n$ an affine variety. Then it defines an ideal $I(X) = (f_1, \ldots, f_k)$. Now, what equations define the diagonal? Well, it is defined by the same k equations as X, but it also needs $x_i = x_{n+i}$ where the coordinates of a point of the diagonal are given by $(x_1, \ldots, x_{2n})$. Now, if we remember in the Zariski topology, a closed set is one defined by the zeros of polynomial equations, and so this diagonal is obviously closed, and so by 2.29 X is a variety. □

We are now in a position to define the spectrum of a ring, the first building block in the definition of a scheme.

Definition 2.31. *Let R be a ring and $X = \{I \subseteq R : I \text{ is a prime ideal of } R\}$. Then we call X the prime spectrum of R and denote it by* $\operatorname{Spec} R$.

What we do now is to note that we can consider elements of R as functions on $\operatorname{Spec} R$. We do this in a way analogous to thinking of elements of the coordinate ring $K[V]$ as functions from V to K, by considering $K[V]$ as a polynomial ring modulo the ideal $I(V)$.

Definition 2.32. *Let $r \in R$ and define f_r to be the function from* $\operatorname{Spec} R$ *to the quotient $R/\mathfrak{p}$ where $f_r(\mathfrak{p})$ is the image of r in $R/\mathfrak{p}$.*

Definition 2.33. *For some ideal I of a ring R, define the set $V(I)$ to be the set of all ideals containing I. Now, we can define the Zariski topology on* $\operatorname{Spec} R$, *in much the same way we did in the case of affine varieties. Let the sets $V(I)$ be the closed sets of* $\operatorname{Spec} R$, *and call the topology defined by these the Zariski topology.*

Example 2.34. *Later in the paper, we will be considering the spectrum* $\operatorname{Spec} \mathbb{Z}$. *Consider some $x \in \mathbb{Z}$. Then x has a prime decomposition, $x = \prod_{i=1}^{n} p_i^{a_i}$, and so any ideal containing p_i also contains x, but not vice versa. Hence x cannot generate a prime ideal, unless $n = a_1 = 1$, or $x = 0$, which generates the trivial prime ideal (0). Obviously all prime numbers generate prime ideals, and so* $\operatorname{Spec} \mathbb{Z} = \{(p) : p \text{ a prime}\} \cup \{(0)\}$.

Note now that definition 2.32 actually has a very close relationship with our concept of $V(I)$. Indeed, we may rewrite our definition of $V(I)$ with respect to the elements of the ring, as in the following lemma.

Lemma 2.35. $V(I) = \{\mathfrak{p} \in \operatorname{Spec} R : f(\mathfrak{p}) = 0 \ \forall f \in I\}$.

Proof. If $f(\mathfrak{p}) = 0$ for $f \in I$, then the image of the ideal I in $R/\mathfrak{p}$ is zero, and so $\mathfrak{p} \supseteq I$. This argument works both ways. □

We will switch between these two ways of looking at $V(I)$, as each can be very useful. We also note that a direct analogy of proposition 2.5 applies to these closed sets, but we do not prove it here as the proof is very similar.

Now, we mention briefly the number of points in $\operatorname{Spec} R$. Note that a consequence of corollary 2.9 is that the maximal ideals of a coordinate ring $K[U]$ are in bijective correspondence with the points

of U. Now, $\operatorname{Spec} R$ has, in general, more points than this, because $\operatorname{Spec} R$ contains a point for every prime ideal, and maximal ideals are prime. We can ask what these extra points (those that are prime but not maximal) correspond to in the language of varieties. Well, by lemma 2.11, they correspond to irreducible sets, which are subvarieties. If a point $\mathfrak{p}$ corresponds to a subvariety X', then $\mathfrak{p}$ is known as the generic point of X'.

We also have a very useful lemma, which will come up again later in the paper.

Lemma 2.36. *Let R, S be rings and $\varphi : R \to S$ a ring homomorphism. Then φ induces a map*

$$\widetilde{\varphi} : \operatorname{Spec} S \to \operatorname{Spec} R$$
$$\mathfrak{p} \mapsto \varphi^{-1}(\mathfrak{p}),$$

such that $\widetilde{\varphi}$ is continuous with respect to the Zariski topology.

Proof. Note that for a ring homomorphism, only prime ideals map into prime ideals, and so this map does indeed map into $\operatorname{Spec} R$ as claimed. Now consider some I an ideal of R, and J the ideal generated by $\varphi(I)$. Then $\widetilde{\varphi}^{-1}(V(I)) = V(J)$, by definition of $\widetilde{\varphi}$.

So, for some $\mathfrak{q} \in \widetilde{\varphi}^{-1}(V(I))$, we see that $\varphi^{-1}(\mathfrak{q}) \supset I$. Now, this gives $\mathfrak{q} \supset \varphi(I) = J$. So, for a closed set $V(I)$, the inverse image under $\widetilde{\varphi}$ is $V(J)$, which is closed. This gives us continuity. □

It is worth noting that this lemma is a direct analog of lemma 2.14 for the language of schemes, and that as in that lemma, the converse is also true, i.e. a map between spectra give us an opposite map between the corresponding rings.

We are now in a position to define basic open sets in $\operatorname{Spec} R$.

Definition 2.37. *Take some $f \in R$ and let $\operatorname{Spec} R_f$ be the set of prime ideals in $\operatorname{Spec} R$ that do not contain f. Then this set is called a basic or distinguished open set in $\operatorname{Spec} R$.*

Note that $\operatorname{Spec} R_f$ is indeed open, since its compliment is equal to $V(f)$, which is by definition closed. The distinguished open sets are in fact the basis for the Zariski topology on $\operatorname{Spec} R$.

Definition 2.38. *Let U be open in $\operatorname{Spec} R$. We define $\mathcal{O}_{\operatorname{Spec} R}(U)$ to be the set of $f := (f_\mathfrak{p})_{\mathfrak{p} \in U}$ such that $f_\mathfrak{p} \in \mathfrak{p}^{-1}R$ and for every $\mathfrak{p} \in U$, there exists some open neighbourhood $V \subseteq U$ containing $\mathfrak{p}$ such that for any $\mathfrak{q} \in V$, $f_\mathfrak{q}$ is of the form $\frac{g}{h}$ for $f, g \in R$ and $g \notin \mathfrak{q}$.*

These are rings under the ordinary addition and multiplication of functions, and it can be shown that they define a sheaf on $\operatorname{Spec} R$, it is in fact the structure sheaf. We now see that the stalk of $\mathcal{O}_{\operatorname{Spec} R}$ at $\mathfrak{p}$ is in fact the local ring of R at $\mathfrak{p}$.

Proposition 2.39. *The stalk $\mathcal{O}_{\mathfrak{p},\operatorname{Spec} R}$ of the sheaf $\mathcal{O}_{\operatorname{Spec} R}$ is $\mathfrak{p}^{-1}R$.*

Proof. We consider the obvious homomorphism

$$\rho : \mathcal{O}_{\operatorname{Spec} R} \to \mathfrak{p}^{-1}R$$
$$(U, f) \mapsto f_{\mathfrak{p}},$$

and show that it is an isomorphism. To do this we first show that it is surjective. Indeed, consider some $f_{\mathfrak{p}} \in \mathfrak{p}^{-1}R$. Then $f_{\mathfrak{p}}$ is of the form g/h for some $h \notin \mathfrak{p}$. In particular, this means that it is defined on $\operatorname{Spec} R_h$, which is open. Hence we see the pair $(\operatorname{Spec} R_h, \frac{g}{h})$ maps onto it, proving surjectivity.

For injectivity, take the pairs (U, f_1) and (U, f_2) and remember by definition 2.38 we have that both functions can be thought of as tuples. Consider the element of the tuples corresponding to $\mathfrak{p}$ and assume that they are equal. Then it suffices to prove that the functions coincide on some neighbourhood of $\mathfrak{p}$. We may assume that both functions are representable as quotient functions $\frac{g_1}{h_1}$ and $\frac{g_2}{h_2}$ respectively. Now, since the functions have the same image, for some $d \notin \mathfrak{p}$ we have that $d(f_1 g_2 - f_2 g_1) = 0$. Hence this is true in all other local rings $\mathfrak{q}^{-1}R$, so $f_1/g_1 = f_2/g_2$ on those rings for $g_1, g_2, h \notin \mathfrak{q}$. But then the $\mathfrak{q}$ for which this is true are those in $\operatorname{Spec} R_{g_1} \cap \operatorname{Spec} R_{g_2} \cap \operatorname{Spec} R_h$, which contains $\mathfrak{p}$. Hence we have a neighbourhood, as required. □

We now have all the pieces of the puzzle necessary to rigorously define an affine scheme.

Definition 2.40. *An affine scheme is a pair $(\operatorname{Spec} R, \mathcal{O}_{\operatorname{Spec} R})$, where $\operatorname{Spec} R$ is defined with the Zariski topology.*

Before we give a definition of a general scheme, we need to discuss morphisms of schemes, again analogously to how we did things with varieties. However, we have a problem, in that pullbacks do not work nicely with schemes, since there is no ground field over which we are working. To solve this we actually make pullbacks part of the definition of a morphism between schemes.

Definition 2.41. *A morphism between affine schemes $(X, \mathcal{O}_X)$ and $(Y, \mathcal{O}_Y)$ is a function $f : X \to Y$ and for every $U \subseteq Y$, we have a pullback map $f_U : \mathcal{O}_Y(U) \to \mathcal{O}_X(f^{-1}(U))$, such that firstly, the following diagram*

$$\begin{array}{ccc} \mathcal{O}_Y(U) & \xrightarrow{\rho_{U,V}} & \mathcal{O}_Y(V) \\ {\scriptstyle f_U}\downarrow & & \downarrow{\scriptstyle f(V)} \\ \mathcal{O}_X(f^{-1}(U)) & \xrightarrow{\rho_{f^{-1}(U),f^{-1}(V)}} & \mathcal{O}_X(f^{-1}(V)) \end{array}$$

commutes. Secondly, we need some compatibility of f_U with the map f. Note that f_U induces a map between stalks of the form

$$\begin{aligned} f_{\mathfrak{p}} : \mathcal{O}_{f(\mathfrak{p}),Y} &\to \mathcal{O}_{\mathfrak{p},X} \\ (U,\rho) &\mapsto (f^{-1}(U), f\rho). \end{aligned} \tag{2.5}$$

These stalks are local rings by proposition 2.39, and so have unique maximal ideals $\mathfrak{m}_{f(\mathfrak{p},Y)}$ and $\mathfrak{m}_{f^{-1}(\mathfrak{p}),X)}$. We require that $f_{\mathfrak{p}}^{-1}(\mathfrak{m}_{\mathfrak{p},X}) = \mathfrak{m}_{f(\mathfrak{p}),Y}$.

Now we have the basic definition of a morphism, we may define a scheme proper.

Definition 2.42. *A scheme is a ringed space $(X, \mathcal{O}_X)$ with all its stalks local rings, and that admits an open covering $\{U_i\}$ with $U_i \subset X$ such that the ringed space $(U_i, \mathcal{O}_X\big|_{U_i})$ is isomorphic to an affine scheme* $\operatorname{Spec} R_i$.

We end this section by giving a brief definition which will crop up in the next section.

Definition 2.43. *Let X be a scheme. Then we say that a scheme over X is a scheme Y together with a morphism of schemes $f : Y \to X$.*

For ease of notation, if R is a ring, we say that an R-scheme is a scheme over $\operatorname{Spec} R$.

2.3 Group Schemes

We can now discuss group schemes. There are three ways to view group schemes, and we will talk about them all, if briefly. These are

- As representable functors from the category K-**Algebra** to the category **Group**,
- As Hopf algebras,
- As group objects in the category **Scheme**.

We give a table showing all the names of categories and their meanings in appendix A, but most should be obvious.

In chapter 4, we will be thinking of Chevalley group schemes in the first and third way, but the other is mentioned for completeness.

To understand the first definition, we need to look at the functor of points. Many algebraic structures in concrete categories can be thought of as Hom-sets.

Example 2.44. *If we have an abelian group G, then we may consider that $G \cong \mathrm{Hom}(\mathbb{Z}, G)$. To show this, consider the map*

$$\begin{aligned} F : G &\to \mathrm{Hom}(\mathbb{Z}, G) \\ g &\mapsto f, \end{aligned} \tag{2.6}$$

such that $f(1) = g$. This map is well defined since any homomorphism from $\mathbb{Z}$ is defined by its image of 1, since $f(a) = af(1)$. Now, we need to show that F is an isomorphism.

Consider $F(gh)$. This is the unique f_{gh} such that $f_{gh}(1) = gh$. But, $F(g)F(h) = f_g f_h$ and $f_g(1)f_h(1) = gh$. So $f_{gh} = f_g f_h$ and we have that F is a homomorphism. For injectivity, let $F(g) = F(h)$. Then $f_g(1) = f_h(1)$, but by the previous argument this means $f_g = f_h$, which gives injectivity. For surjectivity, consider some homomorphism f. Then $f(1) = g$ for some $g \in G$, so $f = f_g$.

In the example above, we showed that the abelian group was isomorphic to the group of homomorphisms between $\mathbb{Z}$ and itself, but we could generalize that by considering G to be a non-abelian group. Here we just get a bijection, which gives us the points of G as a Hom-set, without the group structure. This functor is obviously faithful. Any further generalization to non-faithful functors is of little use, as then we lose information about the points of our original structure.

So, what we are doing in general here is for some category $\mathscr{C}$, we take an element C and this will give us a functor

$$F : X \to \mathrm{Hom}(C, X).$$

However, as we stated before, if this functor is not faithful we will lose information in the mapping. In particular, in **Scheme**, such functors are not necessarily faithful, however carefully we choose our C. To get around this problem, we define a new functor and look at it over all such C.

Definition 2.45. *Let X be a scheme. We define the functor of points of X to be the functor*

$$\mathfrak{F}_X : \mathbf{Scheme}^V \to \mathbf{Set},$$

with $\mathfrak{F}_X(C) = \mathrm{Hom}(C, X)$. Given two schemes C and D, and a morphism $f : C \to D$, then the

induced morphism of sets $\mathfrak{F}_X(D) \to \mathfrak{F}_X(C)$ *is given by* $\mathfrak{F}(f)(g) = g \circ f$.

Fact 2.46. *This gives us an obvious further level of abstraction. We define a functor* $\mathfrak{F}$ *from* **Scheme** *to the category of functors between* $\mathbf{Scheme}^V$ *and* **Set**, *mapping* X *to* $\mathfrak{F}_X$. *For any morphism of schemes* $f : X \to Y$, *we let* $\mathfrak{F}(f)$ *map, for any scheme* Z, *some* $g \in \mathrm{Hom}(Z, X)$ *to* $f \circ g \in \mathrm{Hom}(Z, Y)$.

We now claim that this functor $\mathfrak{F}$ carries enough information in it to preserve the scheme in its image. In fact, we have a stronger statement; that the functor actually defines an equivalence of categories.

Theorem 2.47. *The functor* $\mathfrak{F}$ *defines an equivalence between the category* R**-Scheme** *and the category of functors between* R**-Algebra** *and* **Set**, *where we restrict* $\mathfrak{F}$ *to schemes over a ring* R.

Proof. This is an application of Yoneda's lemma, defining the equivalent category. Note that, using similar techniques to those in lemma 2.36, we have that contravariant functors on schemes are in correspondence with covariant functors on algebras, which is why we can exchange the categories as we have done. Hence all we need show is that for some morphism $\rho : \mathfrak{F}_X \to \mathfrak{F}_Y$, there is a unique associated $f : X \to Y$ such that f induces ρ in the way discussed in fact 2.46. So, consider the scheme X and an affine covering $\{U_\alpha\}$ of X. Let i_α be the inclusion maps of U_α into X and note that by applying ρ we get morphisms from U_α to Y. Now note that if U_α intersects with U_β we get that the associated inclusion maps are identical on the intersection. ρ obviously preserves this property. So we have that

$$f \circ i_\alpha = \rho(i_\alpha).$$

But since the i_α glue to make identity, we get $f = \rho(1)$, the required f. □

This discussion shows how schemes may be thought of as functors via Yoneda's lemma. We now define a representable functor.

Definition 2.48. *A functor is representable if it is isomorphic to* $\mathfrak{F}_X$ *for some object* X.

In particular, for a functor F from K**-Algebra** to **Set**, we have that F is isomorphic to some $\mathfrak{F}_A$ for a K-algebra A. Note that this representability clause was in fact tied up in our definition of $\mathfrak{F}$. This bring us to the first of our three definitions.

Definition 2.49. *A group scheme is a representable functor from the category* K**-Algebra** *to the category* **Group**.

This definition makes sense since the category of groups embeds naturally into the category of sets.

The second way of looking at group schemes requires knowledge of Hopf algebras. These are algebras over a field K which also have the structure of a coalgebra along with a K-linear map called an antipode, which in general has an inverse-like action.

Definition 2.50. *A Hopf algebra H over K is a set Γ along with*

- *multiplication $\Delta : H \otimes H \to H$,*
- *identity $\eta : \{e\} \to H$,*
- *inverse $Inv : H \to H$,*
- *comultiplication $\nabla : H \to H \otimes H$,*
- *augmentation $\varepsilon : H \to K$,*
- *antipode $S : H \to H$,*

such that these functions "work nicely" together. We leave the necessary commutative diagrams out here, but they can be found in [17] pp. 7-8.

This brings us to our second definition.

Definition 2.51. *A group scheme over K is a Hopf algebra over K.*

Finally, the third way of thinking of group schemes is the one implied by the name.

Definition 2.52. *Let $\mathscr{C}$ be a category that contains a zero object and all finite products. A group object in $\mathscr{C}$ is a triple (G, m, σ), where G is an object of $\mathscr{C}$, m is a morphism $m : G \times G \to G$, and σ is a morphism $\sigma : G \times G \to G$ such that the following diagrams commute.*

$$\begin{array}{ccc} G \times G \times G & \xrightarrow{m \times id_G} & G \times G \\ {\scriptstyle id_G \times m}\downarrow & & \downarrow{\scriptstyle m} \\ G \times G & \xrightarrow{m} & G \end{array} \qquad \begin{array}{ccccc} G \times G & \xleftarrow{\{id_G, 0_G\}} & G & \xrightarrow{\{0_G, id_G\}} & G \times G \\ & {\scriptstyle m}\searrow & \downarrow{\scriptstyle id_G} & \swarrow{\scriptstyle m} & \\ & & G & & \end{array}$$

$$\begin{array}{ccccc} G \times G & \xleftarrow{\{id_G, \sigma\}} & G & \xrightarrow{\{\sigma, id_G\}} & G \times G \\ & {\scriptstyle m}\searrow & \downarrow{\scriptstyle 0_G} & \swarrow{\scriptstyle m} & \\ & & G & & \end{array}$$

The first diagram gives us associativity, the second gives us the existence of a two-sided identity and the third gives us the existence of two-sided inverses.

It is worth noting that we have abused notation in the second and third commutative diagrams. We defined in appendix A the morphism 0_G to be the unique map $0_G : G \to Z$, where Z is the zero object of the category $\mathscr{C}$, whereas in the diagrams here we have defined 0_G to be map from G into itself. To explain this, note that since Z is a zero object, in particular it is initial, and so there is a unique embedding $e_G : Z \hookrightarrow G$. Hence when we write 0_G in the above diagrams, what we actually mean is $e_G \circ 0_G$. Since e is unique, this map is well-defined. We give two elementary examples.

Example 2.53. *Consider the set of permutations of three elements in the category* **Set**. *Then this admits the obvious group structure to give the group S_3. So this set is a group object in the category* **Set**.

Example 2.54. *An affine algebraic group is a group object in the category of varieties over a given field. Indeed, this is by definition since the group multiplication and inverse maps must be regular functions, which are morphisms of varieties.*

Group schemes then, are group objects in the category of schemes.

Definition 2.55. *A Chevalley group scheme is a scheme that admits the group structure of a Chevalley group.*

Now, we have three definitions of a group scheme, definitions 2.49, 2.51 and 2.55 and, unsurprisingly given that we have called all three a group scheme, these are equivalent.

Theorem 2.56. *Definition 2.49 is equivalent to definition 2.51.*

Proof. Omitted. See [16], pp. 55-56. □

The correspondence is fairly easy to see, and comes almost immmediately from the Yoneda lemma. By part 2 of the Yoneda lemma, we have that if F and G are functors represented by K-algebras A and B, then the morphisms between F and G correspond to homomorphisms between B and A. So, we now take $G = F \times F$, and hence $B = A \otimes A$.

The group map from $F(X) \times F(X)$ to $F(X)$ is the group multiplication, and Yoneda tells us this corresponds to another map, which is comultiplication $A \to A \otimes A$. Similarly we use Yoneda to give us augmentation from identity and antipode from the inverse.

Theorem 2.57. *Definition 2.51 and definition 2.55 are equivalent.*

Proof. Omitted. See [4] pp. 23-24. □

Chapter 3

Lie Algebras and Chevalley Groups

We start this section off with a quick discussion of root systems, which are used in the construction of a Chevalley basis for a Lie Algebra, which is turn is needed to define a Chevalley group.

3.1 Root Systems and their Weyl Groups

We start by noting that in a Euclidean space, that is, $\mathbb{R}^n$ for some $n \in \mathbb{N}$, every point a has a unique corresponding orthogonal reflection with respect to the standard Euclidean inner product $(\ ,\ )$. This is the linear transformation that fixes pointwise the hyperplane orthogonal to the point a and sends a to $-a$. This transformation w_a is given by

$$w_a(b) = b - \frac{2(a,b)}{(a,a)}a. \tag{3.1}$$

Note that if b is orthogonal to a, then $(a,b) = 0$ and so $w_a(b) = b$, and $w_a(a) = a - 2a = -a$, and so this is indeed the transformation we require. We now give a definition of a root system.

Definition 3.1. *A root system Φ is a set of points of E that satisfy four properties.*

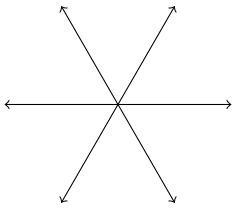
1. *There are finitely many points, and they span E,*
2. *$\forall a, b \in \Phi$, $2\frac{(a,b)}{(b,b)} \in \mathbb{Z}$,*
3. *If $a \in \Phi$ and $\gamma a \in \Phi$ for some $\gamma \in \mathbb{R}$, then $\gamma \in \{1, -1\}$,*
4. *If $a \in \Phi$, then the associated reflection w_a permutes the points of Φ.*

Each such Φ has a corresponding Weyl Group W that is the group generated by the reflections w_α for all $\alpha \in \Phi$.

Lemma 3.2. *The Weyl group W for some root system Φ is finite.*

Proof. By definition of a root system, w_a permutes Φ. But since W is generated by w_a, we have that any $w \in W$ permutes Φ. Hence W must be a subgroup of the symmetric group $S(\Phi)$, which has order $|\Phi|!$; in particular, this gives W is finite. □

Example 3.3. *Consider $E = \mathbb{R}^2$, and $\Phi = \{(1,0), (\frac{1}{2}, \frac{\sqrt{3}}{2}), (-\frac{1}{2}, \frac{\sqrt{3}}{2}), (-1,0), (-\frac{1}{2}, -\frac{\sqrt{3}}{2}), (\frac{1}{2}, -\frac{\sqrt{3}}{2})\}$. It can easily be seen that this is a root system, and now we can consider the Weyl group. It is generated by three involutions (since three reflections are repeats of the others), and we can see that any one of these conjugated by a second gives us the third. This means that the Weyl group must be isomorphic to S_3. Since we are in $\mathbb{R}^2$ we can draw the root system, as in the diagram below.*



Definition 3.4. *A subset Δ of a root system is a base if*

1. *Δ is a basis E,*
2. *For each root α in Φ, the representation with respect to the basis must have all non-positive or all non-negative integer coefficients.*

In the above example, the root system has base $\{(1,0), (-\frac{1}{2}, \frac{\sqrt{3}}{2})\}$.Note that W is in fact a Coxeter group; it has finite presentation, with generators Δ and relations of the form

$$(a_i a_j)^{m_{i,j}} \tag{3.2}$$

for $a_i, a_j \in \Delta$.

Now we have the definition of a base, we have a partition of Φ into Φ^+ and Φ^- in the obvious way. It is worth noting that these are both non-empty, for if $a \in \Phi^\pm$, then $-a \in \Phi^\mp$. This partition in turn gives us a partial order on E; if a a positive root, then $a \succ 0$, and for negative roots, $\alpha \prec 0$. Then for $v, w \in E$, we have that $v \prec w$ if and only if $w - v$ is a positive root.

Now, the subspace orthogonal to every root is a hyperplane; a subspace of dimension $n-1$ in E. These hyperplanes partition E into a number of connected parts. These are known as the Weyl chambers of the root system.

It is worth noting that since the inner product between any two roots must be an integer, the angles allowed between two roots are severely limited. This is because $(a,b) = |a||b|\cos\theta$, which gives us that

$$\frac{2(a,b)}{(b,b)} \cdot \frac{2(b,a)}{(a,a)} = 4\frac{|a|}{|b|}\cos\theta \cdot \frac{|b|}{|a|}\cos\theta = 4\cos^2\theta. \tag{3.3}$$

Since $\cos^2\theta \leq 1$, this gives us seven possibilities for θ. The full table can be seen in [10], page 45 . This can now lead us to state some small, but useful, lemmas.

Lemma 3.5. *Let $r \in \Delta = \{r_i\}_{i=1}^n$. Then w_r transforms all positive roots into positive roots, other than r, which is transformed to $-r$.*

Proof. We have already seen that $w_r(r) = -r$, so let us take some $r' \in \Phi^+ \backslash \{r\}$. Then

$$r' = \sum_{i=1}^{n} \lambda_i r_i, \tag{3.4}$$

with $\lambda_i \geq 0$ for all $i \leq n$. Now, we have that $n > 1$, so there exists some j such that $r_j \neq r$, and $\lambda_j > 0$. So, in $w_r(r')$, $\lambda_j > 0$ and hence $w_r(r') \in \Phi^+$ as required. □

Lemma 3.6. *Let a, b be non-proportional roots. If $(a,b) < 0$, then $a+b$ is a root.*

Proof. Note that if $(a,b) < 0$ then $\frac{2(a,b)}{(b,b)} < 0$, since the inner product is positive definite. By the table mentioned above in Humphreys [10], we may just read off the entries, and we notice that if $(a,b) < 0$, then one of $\frac{2(a,b)}{(b,b)}$ or $\frac{2(b,a)}{(a,a)}$ is equal to -1. If it is the former, then $w_b(a) = a - (-1)b = b + a$. If the latter, then $w_a(b) = a + b$. Since the Weyl group leaves the root system invariant, we have that $a+b$ is a root, as required. □

Corollary 3.7. *Let a, b be non-proportional roots. Then if $(a,b) > 0$, then $a-b$ is a root.*

Proof. This follows from the previous lemma by substituting b for $-b$, since the form is bilinear. □

Corollary 3.8. *Let a, b be non-equal roots in a base Δ. Then if $(a,b) \leq 0$, $a-b$ is not a root.*

Proof. If not, then $(a,b) > 0$, so by lemma 3.6, $a-b$ is a root, but this is a contradiction from the definition of a base, since one coefficient is positive, and the other negative. □

The base of a root system always exists, but it is not unique. In fact, there are as many bases as there are Weyl chambers of E, and there is a natural bijection between them. To show this, take some $c \in X$, where $X \subseteq E$ is a Weyl chamber, and we consider all the roots in $a \in \Phi$ such that $(a, c) > 0$. We call this set $\Phi^+(c)$. Now we take some $b \in \Phi^+(c)$ and call it indecomposable if it can't be written as the sum of two other roots b_1 and b_2 in $\Phi^+(c)$ (and decomposable otherwise). Then we claim that the set of indecomposable root in $\Phi^+(c)$ defines a base for Φ. The proof of this can be found in [10].

The concept of a base gives us another natural ordering on a root system. We give its definition now, and state a short lemma, which will be used later when considering subgroups of Chevalley groups.

Definition 3.9. *Consider a root $a \in \Phi$ and a base Δ of Φ. Then a can be uniquely written as*

$$a = \sum_{r \in \Delta} \lambda_r r \tag{3.5}$$

for $\lambda_r \in \mathbb{R}$. We define the height of a to be $\sum \lambda_r$.

Lemma 3.10. *We may choose an ordering $\prec$ such that if $a \prec b$, then $h(a) \leq h(b)$.*

Proof. Let Φ be our root system, and V the vector space over $\mathbb{R}$ spanned by it, of dimension d, with basis $\{r_1, ..., r_d\}$. Then we may think of the height function as a linear map from Φ to $\mathbb{R}$. Note that the vectors $\{r_1 - r_2, r_2 - r_3, ..., r_{d-1} - r_d\}$ are all in the kernel of h and are linearly independent by construction, hence $\dim(\ker(h)) \geq d-1$, but since h is obviously not the zero map, $\dim(\ker(h)) = d-1$.

Now consider some v a vector of height 1. So we now have a basis of V, given by $\{v, r_1 - r_2, r_2 - r_3, ..., r_{d-1} - r_d\}$ and so any vector can be written as the sum

$$\lambda v + \sum_{i=1}^{d-1} \lambda_i (r_i - r_{i+1}). \tag{3.6}$$

The height then is obviously λ, since the height map is linear. We can now choose Φ^+ to be the intersection of Φ with vectors whose first non-zero coefficient under the above basis is positive. Looking back at the definition of $\prec$, we can see this gives us a working Φ^+, ending the proof. □

Now we have looked at the height of roots, let us consider their length.

Definition 3.11. *The length of some $w \in W$ with respect to a base $\Delta = \{r_1, ..., r_k\}$ of a root system is the least t such that $w = w_{r_{i_1}} ... w_{r_{i_t}}$ for some $i_1 \dots i_t$. We write this as $l(w)$.*

Example 3.12. *This obviously gives $l(1) = 0$ (by convention) and $l(w_r) = 1$ for all $r \in \Delta$.*

We can also consider the function $n : W \to \mathbb{N}$ defined by $n(w) = |\{r \in \Phi^+ : w(r) \in \Phi^-\}|$. We can now give a lemma, with some basic properties of n.

Lemma 3.13. *Let $r \in \Delta$ and $w \in W$. Then*

1. *$n(w_r w) = n(w) + 1$ if $w_r^{-1} \in \Phi^+$,*
2. *$n(w_r w) = n(w) - 1$ if $w_r^{-1} \in \Phi^-$,*
3. *$n(w w_r) = n(w) + 1$ if $w_r \in \Phi^+$,*
4. *$n(w w_r) = n(w) - 1$ if $w_r \in \Phi^-$.*

Proof. This is relatively straightforward, using lemma 3.5. From this, we know that $w(r)$ only moves one root from Φ^+ to Φ^-; r itself. Hence $n(w_r w) = n(w) \pm 1$. If r is in the image of $\Phi^\pm$ under w, then w_r will put it into $\Phi^\mp$, increasing (resp. decreasing) $n(w)$ by one, proving 1 and 2. We make a near identical argument for 3 and 4. □

It turns out that this n function is actually equal to the length of a Weyl group element, something that will be useful later when discussing finite Chevalley groups.

Theorem 3.14. *$n(w) = l(w)$ for all $w \in W$.*

Proof. Consider some $w \in W$ of length k. Then

$$w = w_{r_1} ... w_{r_k}. \tag{3.7}$$

Now, $n(w) \leq n(w_{r_1} w) + 1$, by lemma 3.13. Inducting, we obtain $n(w) \leq n(w_{r_k} w_{r_{k-1}} ... w_{r_1} w) + k$. But since all of these w_{r_i} are involutions, $w_{r_k} w_{r_{k-1}} ... w_{r_1} w = 1$, and so we get $n(w) \leq k$.

Now, suppose $n(w) < k$. Then by the previous lemma there must be some $j \leq k - 1$ such that

$$w_{r_1} ... w_{r_j}(r_{j+1}) \in \Phi^-. \tag{3.8}$$

Hence, we can take off the first $i - 1$ terms (for some i) to give $w_{r_i} ... w_{r_j}(r_{j+1}) \in \Phi^-$ and removing the w_{r_i} gives us an expression in Φ^+. Hence that expression must equal r_i by 3.5. From this argument, we can see that

$$w_{r_i} = w_{r_{i+1}} ... w_{r_j} w_{r_{j+1}} w_{r_j} ... w_{r_{i+1}}. \tag{3.9}$$

Now, taking inverses on our expression gives us

$$w_{r_i}...w_{r_j} = w_{r_{i+1}}...w_{r_{j+1}}.$$

By assumption we have $w = w_{r_1}...w_{r_k}$, and we may now trade out $w_{r_{i+1}}...w_{r_{j+1}}$ for $w_{r_i}...w_{r_j}$. This removes the $j+1$ term and leaves us with two r_i terms adjacent, which cancel to leave us with

$$w = w_{r_1}...w_{r_{i-1}}w_{r_{i+1}}...w_{r_j}w_{r_{j+2}}...w_{r_k},$$

giving us a shorter expression for w, a contradiction. This completes the proof. □

We now briefly discuss a theorem about relations of bases to each other, a corollary of which will come up later in the paper.

Theorem 3.15. *If Δ_1 and Δ_2 are bases of Φ, then there exists a unique $w \in W$ such that $w(\Delta_1) = \Delta_2$.*

Proof. Let $\Phi^{+,1}$ and $\Phi^{+,2}$ be the positive root systems containing Δ_1 and Δ_2 respectively. We proceed by induction on $|\Phi^{+,1} \cap \Phi^{-,2}| := n$. For the base case, if $n = 0$, then $\Phi^{+,1} = \Phi^{+,2}$, and so $w = 1$ does the job.

Now , assume $n > 0$. Thus we have that $\Delta_1 \cap \Phi^{+,2}$ is non-empty, since Δ_1 generates Φ. So consider some $r \in \Delta_1 \cap \Phi^{+,2}$, and note that $|w_r(\Phi^{+,1}) \cap \Phi^{+,2}| = n - 1$ since $w_r(r) = -r \notin \Phi^{+,1}$.

So, we have $w_r(\Delta_1)$ being the base for $w_r(\Phi^{+,1})$ and so by our inductive hypothesis, we have that there exists some w' such that

$$w'w_r(\Phi^{+,1}) = \Phi^{+,2},$$

which gives us our required $w = w'w_r$.

For uniqueness, assume $w_1(\Phi^{+,1}) = \Phi^{+,2}$ and $w_2(\Phi^{+,1}) = \Phi^{+,2}$. Then $w_2^{-1}w_1(\Phi^{+,1}) = \Phi^{+,1}$ and so by theorem 3.14, we get that $n(w_2^{-1}w_1) = 0$ and so $l(w_2^{-1}w_1) = 0$ and hence $w_1 = w_2$ as required. □

Corollary 3.16. *Consider a root system Φ. Then there is a unique $w_0 \in W$ such that $w_0(\Phi^+) = \Phi^-$.*

Proof. Note that Φ^+ defines a base Δ, and since Φ^- is just the opposite of a positive root system, it also defines a base, Δ'. Then by theorem 3.15, we have a unique w_0 as required. □

We now move to briefly explain how root systems can be classified, and for that we require the definition of irreducible root systems, which are to be the building blocks for all root systems.

Definition 3.17. *Let Φ be a root system. We say that Φ is reducible if there exist Φ_1, Φ_2 non-empty subsets of Φ such that $\Phi_1 \cup \Phi_2 = \Phi$, and for all $a \in \Phi_1$ $b \in \Phi_2$, we have that $(a, b) = 0$. If two subsets Φ_1 and Φ_2 have this second property, we say they are orthogonal.*

We say that Φ is irreducible if and only if it is not reducible. We claim that a root system has a unique decomposition in terms of irreducible root systems, but this will have to wait until we give some elementary properties of these irreducible systems.

Lemma 3.18. *If the root system Φ (with base Δ) is irreducible, then there exists a unique maximal (with respect to the ordering $\prec$) $a \in \Phi$.*

Proof. Note firstly that Φ is finite, and so a maximal element has to exist. Let this element be $a := \sum_{v \in \Phi} c_v v$. Since a is maximal, $a \in \Phi^+$. We now partition Δ into two sets, by looking at the coefficients of a in the above sum. Let $\Delta_1 = \{v \in \Delta : c_v > 0\}$, and $\Delta_2 = \{v \in \Delta : c_v = 0\}$.

We now show that $\Delta_2 = \emptyset$. Assume it is not. Then take some $u \in \Delta_2$. Then $(u, v) \leq 0$, by corollary 3.8, and there is some $u' \in \Delta_1$ such that u is not orthogonal to u'. This follows from the fact that Φ is irreducible; if there was no such u', Φ could be split into two orthogonal subsets. These two observations combined tell us that $(u, u') < 0$, and so again by corollary 3.8, we have $u + u'$ is a root. But then $a \prec u + u'$, a contradiction. Hence Δ_2 is void.

Now, this shows that for all $b \in \Phi$, $(a, b) \geq 0$ and there is at least one b such that $(a, b) > 0$. If a' is another maximal element, then the same applies to it, and so if follows that $(a, a') > 0$. But then by definition of $\prec$, we have either $a \prec a'$ or $a' \succ a$, a contradiction, which finishes the proof. □

Proposition 3.19. *Let Φ be an irreducible root system. Then the Weyl group W acts irreducibly on the underlying Euclidean space E.*

Proof. Omitted. □

3.2 Root System Constructs

To get any more information about Chevalley groups, we now consider root systems in a slightly different light, and consider some other groups related to the Weyl group, along with some functoral constructions that will help shed light on Chevalley group schemes.

We start by noting that the Weyl group is in fact a Coxeter group; i.e. a group that is defined by generators $\{w_1, ..., w_k\}$ and relations of the form $(w_i w_j)^{m_{ij}} = 1$. We do not prove this fact here, but

refer the reader to [1] page 23.

Definition 3.20. *The Braid group B for a root system Φ is defined with generators in the finite set Λ of Φ and the relations*

$$\{\underbrace{q_iq_j....q_iq_j}_{m_{i,j}} : q_i \in \Lambda \ \forall i,j\},$$

where $m_{i,j}$ are the same as in the definition of the Weyl group of Φ (as in equation 3.2), and $|\Lambda| = |\Delta|$.

We now consider the canonical homomorphism between the Braid group and the Weyl group,

$$\begin{aligned} p_1 : B &\to W \\ q_i &\mapsto r_i. \end{aligned} \tag{3.10}$$

This is surjective. Consider $\ker(p_1)$ and let $X = [\ker(p_1), \ker(p_1)]$, the commutator subgroup.

Definition 3.21. *The extended Coxeter group V is defined to be the quotient B/X.*

This group can also be realised by its presentation. It has the generators and relations of the Braid group, but it has extra generators

$$\{g(s) : s \in S\},$$

where we define the set S to be the set of reflections of the root system Φ; all of those elements in W conjugate to some w_a for $a \in \Phi$, and extra relations

1. $q_i^2 = g(w_{a_i})\ \forall a_i \in \Phi$,
2. $g(s)^{q_i} = g(w_{a_i}(s))$,
3. $g(s)$, $g(s')$ commute with each other for all $s, s' \in S$,

where $w_{a_i}(s)$ is s conjugated by w_{a_i}.

Note now that there is another homomorphism $p_2 : V \to W$ that acts in the same way as p_1. The kernel of this map is the group generated by $g(s)$ for $s \in S$. We call this group U, and it is normal since it is the kernel of a homomorphism, and also abelian by point 3 above.

Definition 3.22. *A root lattice L for a root system Φ is the $\mathbb{Z}$-span of Φ.*

In 1966, Tits introduced a functor, for a given root lattice L and root system Φ, called $\mathfrak{N}$ that maps (D, ε) in $\mathbf{AbGroup}^{(2)}$ to the data $(N, N_s, p)_{s\in S}$ for some group N, homomorphism $p : N \to W$ and subgroups $N_s \subseteq N$. The data must satisfy three conditions:

1. $\ker(p)$ is abelian,

2. If $p(n) = w$, then $nN_sn = N_{w(s)}$,

3. $p(N_s) = \{1, s\}$.

Example 3.23. *The data given by $(V, p_2, V_s)_{s\in S}$ where V_s is the subset of V generated by the set $\{v \in V : v^2 = g(s)\}$. We omit the proof, but it is relatively straightforward using the given definitions.*

We can now construct the data canonically given an abelian group D with involution ε. This can be found in [3], we give the result here. To do this, we look at the group $\mathrm{Hom}(L, D) =: T$. This has an obvious left action of W, where $w(t(\alpha)) = t(w(\alpha))$. We can consider a particular set of maps in T, those of the form

$$h_s(l) = d^{n_\alpha(l)}, \tag{3.11}$$

where n_α is the co-root associated to α, and is thought of as a function via the Killing form of definition 3.29.

We now define T_s (for $s = w_r$ for some root r) to be the subgroup of T defined by homomorphisms mapping some $x \in L$ to something of the form $d^{\nu(x)}$ where $d \in D$ and $\nu(x)$ is linear from L to $\mathbb{Z}$ and proportional to n_r, the co-root associated to r.

We now have the the data

$$\{T, T_s, h_s\}_{s\in S} \tag{3.12}$$

satisfies the three conditions of Tits. Using these two, we can get our canonical data for (D, ε).

Definition 3.24. *The group N for a lattice L and root system Φ is defined to be the quotient*

$$(V \ltimes T)/\{u, f(u)\}_{u\in U},$$

where f is the map mapping $g(s)$ to h_s^{-1}.

It is not immediately obvious that the set $\{u, f(u)\}_{u\in U}$ is a subgroup, however once you realise f is a homomorphism, this quickly becomes apparent, since $(u, f(u))(v(f(v)) = (uf(u)^{-1}vf(u), f(u)f(v))$ and since U is abelian, we can cancel the conjugating $f(u)$ terms to give us $(uv, f(uv))$. This subgroup as a general construction is known as the graph of the homomorphism f.

We define the homomorphism p by having it act trivially on T, and defining $p = p_2$ on V. The subgroups N_s are those generated by the image of $T_s \times V_s$, under the quotient projection. We can then

claim that the data $\{N, N_s, p\}_{s\in S}$ satisfies Tits' conditions.

Theorem 3.25. *Let L be a lattice and Φ a root system. The functor $\mathfrak{N}$ corresponds a pair (D, ε) to the extension*

$$1 \to T \to N \xrightarrow{p} W \to 1.$$

We do not give a proof here, but we note that since the quotient defining N is the graph that quotients out U of V to leave W, this makes intuitive sense.

3.3 Lie Algebras

We assume that the reader is familiar with the basic theory of Lie algebras. If not, the theorems used, along with some of the proofs, can be found in appendix B.

Consider a complex semisimple Lie algebra $\mathfrak{L}$ with Cartan subalgebra $\mathfrak{C}$. We can begin to look at the Cartan decomposition of the Lie algebra. For this we define a weight of a Lie algebra. This notion generalises eigenvectors for a group of linear transformations.

Definition 3.26. *Consider a complex vector space V and a subalgebra $\mathfrak{L}$ of $\mathfrak{gl}(V)$. Given a linear map $\alpha \in \mathfrak{L}^*$ ($\mathfrak{L}^*$ being the dual of $\mathfrak{L}$), we call α a weight for $\mathfrak{L}$ if the set*

$$\mathfrak{L}_\alpha := \{v \in V : [x, v] = \alpha(x)v \ \forall x \in \mathfrak{L}\} \tag{3.13}$$

is non-empty.

It is easy to check that in fact we have that $\mathfrak{L}_\alpha$ is actually a vector subspace of V, and indeed a Lie algebra, since it is one dimensional and therefore trivially closed under Lie multiplication. Note also that by definition, these spaces are invariant under Lie multiplication by elements of $\mathfrak{C}$.

Definition 3.27. *A weight of a Lie algebra is a linear map α such that $\mathfrak{L}_\alpha$ is non-zero. We call this $\mathfrak{L}_\alpha$ the weight space of α.*

Now, we consider the Cartan subalgebra $\mathfrak{C}$ and $\mathfrak{L}$ its containing semisimple Lie algebra. Then since all $c \in \mathfrak{C}$ are semisimple, it follows that so are the elements $\mathrm{ad}(c)$ in $\mathrm{End}(\mathfrak{L})$. Since these maps commute (again, because their associated vectors commute) we may simultaneously diagonalise $\mathfrak{C}$. So, $\mathfrak{L}$ may be decomposed into a sum of weight spaces of the Cartan subalgebra.

So, the Cartan decomposition of a simple Lie algebra $\mathfrak{L}$ is given by $\mathfrak{L} = \mathfrak{L}_0 \oplus \mathfrak{L}_{\alpha_1} \oplus ... \oplus \mathfrak{L}_{\alpha_k}$. Note that since $\mathfrak{C}$ is abelian, $[c, c'] = 0 \ \forall c, c' \in \mathfrak{C}$, and so the Cartan subalgebra lies inside $\mathfrak{L}_0$. In fact, the maximality of $\mathfrak{C}$ implies that $\mathfrak{L}_0 = \mathfrak{C}$, but we omit the proof here, it can be found in [7].

Lemma 3.28. $[\mathfrak{L}_{r_1}, \mathfrak{L}_{r_2}] \subseteq \mathfrak{L}_{r_1+r_2}$.

Proof. Let $x \in \mathfrak{L}_{r_1}$ and $y \in \mathfrak{L}_{r_2}$. Then $[h, [x, y]] = -[x, [y, h]] - [y, [h, x]] = [x, [h, y]] + [[h, x], y]$ by Jacobi identity. But we know how the Lie bracket works with h; we get that this is equal to $[x, r_2(h)y] + [r_1(h)x, y] = (r_1 + r_2)(h)[x, y]$ by bilinearity of the Lie bracket. This gives us what we claimed. □

It turns out that the weights from the Cartan decomposition form a root system, but due to the need for brevity we do not prove this here. However, we do give a few of the necessary tools. Firstly, one requires an isomorphism between the Cartan subalgebra $\mathfrak{C}$ and its dual $\mathfrak{C}^*$. For $c \in \mathfrak{C}$, we associate the map $k(c, \_)$, where $k(c, \_)$ is the Killing form.

Definition 3.29. *The Killing form on a complex Lie algebra $\mathfrak{L}$ is a symmetric, bilinear form given by $k(x, y) = \mathrm{tr}(\mathrm{ad}(x), \mathrm{ad}(y))$.*

Lemma 3.30. *The Killing form restricted to some $\mathfrak{L}_\alpha$ is an inner product.*

Proof. Omitted. □

For ease of use, we shall now just use parentheses for the Killing form, since we will only use it now in the context of the inner product. The proofs that weights define a root system can be found in [7].

Now it turns out, rather nicely, that simple Lie algebras are in bijection with the root systems as given by their Cartan decomposition. It is worth noting here that as the Cartan subalgebra is not necessarily unique for a given Lie algebra $\mathfrak{L}$, the roots $\alpha_1, \ldots, \alpha_k$ are not unique for given $\mathfrak{L}$, but they define isomorphic root systems. So, the roots are not unique, but the root system is (up to isomorphism).

Now, consider for some semisimple complex Lie algebra $\mathfrak{L}$ the Cartan decomposition $\mathfrak{L} = \mathfrak{C} \oplus \bigoplus_{r \in \Phi} \mathfrak{L}_r$, and for every root r consider $h_r = \frac{2r}{(r,r)}$. Also for every r pick some $e_r \in \mathfrak{L}_r$, noting that if we have picked $e_r \in \mathfrak{L}_r$, we can pick a corresponding $e_{-r} \in \mathfrak{L}_{-r}$ such that $[e_r, e_{-r}] = h_r$. Choose the e_r in this way.

Now consider $\{h_r : r \in \Delta\} \cup \{e_r : r \in \Phi\} =: B$, where Δ is a base for Φ. This is a basis for $\mathfrak{L}$.

Example 3.31. *Consider the Lie algebra* $\mathfrak{L} := \mathfrak{sl}_3(\mathbb{C})$. *Then the Cartan subalgebra of* $\mathfrak{L}$ *is* 2*-dimensional and consists of the diagonal matrices of trace zero. This is obvious, since there are no other semisimple elements in* $\mathfrak{L}$. *It is also easy to see that each* $\mathfrak{L}_r$ *is spanned by some* e_{ij} *for* $1 \leq i, j \leq 3$, $i \neq j$ *where* e_{ij} *is the empty matrix, with a* 1 *in the* i, j^{th} *position. The root system of this Lie algebra is* A_2.

It will be useful later to see how these basis elements multiply together in $\mathfrak{L}$, and so give the Lie brackets of these elements here

Fact 3.32. *We have*

1. $[h_r, h_s] = 0$,
2. $[h_r, e_s] = A_{rs} e_s$,
3. $[e_r, e_{-r}] = h_r$,
4. $[e_r, e_s] = 0$ for $r + s \notin \Phi$,
5. $[e_r, e_s] = N_{r,s} e_{r+s}$.

Note that for 1), this is by definition of the Cartan subalgebra; since it is abelian we get the Lie bracket is zero. For 2), we have that the weight spaces are 1-dimensional and invariant under $\operatorname{ad}\mathfrak{C}$, and so we get a multiple of e_s back. 3) is by our choice of e_{-r}. 4) and 5) come from lemma 3.28.

We call the field elements $N_{r,s}$ structure constants for $\mathfrak{L}$. We have various relationships between them, the proofs of which are relatively unenlightening; they consist mainly of case analysis and liberal use of the Jacobi identity, so we will omit the proofs and just state the results when and where we require them.

The first of these we shall use is the fact that $N_{r,s} N_{-r,-s} = -(p+1)^2$, where p is the smallest coefficient of s in the s string through r for the root system Φ. In fact it turns out that we may choose our vectors e_r in such a way that $N_{r,s} = \pm(p+1)$. This is the last piece of information we need to define a Chevalley basis.

Definition 3.33. *Consider a Lie algebra* $\mathfrak{L}$ *with Cartan decomposition* $\mathfrak{C} \oplus \bigoplus_{r \in \Phi} \mathfrak{L}_r$. *Let* h_r *be as defined above, and* e_r *vectors in* $\mathfrak{L}_r$ *such that* $[e_r, e_{-r}] = h_r$ *and* $[e_r, e_s] = \pm(p+1)$. *Then we have that* $\{h_r : r \in \Delta\} \cup \{e_r : r \in \Phi\} =: B$ *is a basis for* $\mathfrak{L}$, *and we call this a Chevalley basis.*

The next step in defining the Chevalley group is to consider the exponential map for nilpotent derivations for a Lie algebra $\mathfrak{L}$. Consider some nilpotent derivation D such that $D^m = 0$ and m is the smallest such integer.

Definition 3.34. *We define the exponential map* exp (D) *to be*

$$\exp(D) := 1 + \sum_{r=1}^{m-1} \frac{D^r}{r}.$$

We note that this is an automorphism of $\mathfrak{L}$. We now consider the derivation $\mathrm{ad}(e_r)$, a derivation by example B.6. We also claim this is nilpotent. Indeed, by fact 3.32, we have $\mathrm{ad}(e_r)(\mathfrak{C}) = \mathfrak{L}_r$, so in this case $\mathrm{ad}(e_r)^2 = 0$. Similarly $\mathrm{ad}(e_r)(\mathfrak{L}_{-r}) \subseteq \mathfrak{C}$ so here $\mathrm{ad}(e_r)^3 = 0$. Obviously $\mathrm{ad}(e_r)(\mathfrak{L}_r) = 0$ and then lastly $\mathrm{ad}(e_r)(\mathfrak{L}_{r'}) \subseteq \mathfrak{L}_{r+r'}$. But since root strings are of finite length (i.e. $r + nr'$ is not a root for $n \geq m$ for some finite m), we can only continue doing this finitely many times before it terminates. Hence $\mathrm{ad}(e_r)$ is nilpotent.

We now do the obvious, and consider exp $\mathrm{ad}(e_r)$, in particular exp $\mathrm{ad}(\zeta e_r)$ for some $\zeta \in \mathbb{C}$. We denote this automorphism by $x_r(\zeta)$.

It is a simple exercise to see that these automorphisms send elements of a Chevalley basis to other $\mathbb{Z}$-linear combinations of Chevalley basis elements.

Now, we expand on the previous ideas by allowing us to recreate these structures over arbitrary fields by tensoring. Take a Lie algebra $\mathfrak{L}$ with Chevalley basis B and take the $\mathbb{Z}$-linear span of B: call it $B_{\mathbb{Z}}$. Then we may tensor this construction with an arbitrary field K to give $\mathfrak{L}_K := K \otimes B_{\mathbb{Z}}$.

Now, consider some element of $\mathfrak{L}_K$

$$k \otimes \left(\sum_{r \in \Delta} n_{h,r} h_r \oplus \sum_{r \in \Phi} n_{e,r} e_r \right),$$

where n's are all in $\mathbb{Z}$. But, by the distributive properties of tensors and direct sums, and the bilinearity of the tensor product, we can rewrite this as

$$\sum_{r \in \Delta} a_r (1 \otimes h_r) + \sum_{r \in \Phi} b_r (1 \otimes e_r)$$

for a_r and b_r in K.

It is obvious that this gives us a vector space over K, and in fact we may use our old Lie bracket to define a new one. Define a Lie bracket on $\mathfrak{L}_K$ by $[a \otimes x, b \otimes y] = ab \otimes [x, y]$. We identify h_r and e_r

from $\mathfrak{L}$ into $\mathfrak{L}_K$ as $1 \otimes h_r$ and $1 \otimes e_r$ respectively, and so we may talk of these elements as belonging to either Lie algebra.

Finally, we make an analogous connection with the automorphisms of $\mathfrak{L}$, $x_r(\zeta)$. Take the matrix representing $x_r(\zeta)$ with respect to the Chevalley basis of $\mathfrak{L}$ and in their coefficients change all instances of ζ to some $t \in K$. As before, we do not change notation, but now consider $x_r(t)$ to be an automorphism of $\mathfrak{L}_K$.

3.4 Chevalley Groups

We are now in command of the necessary tools to define a Chevalley group for a Lie algebra.

Definition 3.35. *Let $\mathfrak{L}$ be a Lie algebra and K a field. Then we denote by $\mathfrak{L}(K)$ the group generated by $x_r(t)$ for $t \in K$, $r \in \Phi$. This is a subgroup of the automorphisms of $\mathfrak{L}_K$ and we call it the Chevalley group of type $\mathfrak{L}$ over the field K.*

We claim that $\mathfrak{L}(K)$ is independent of the choice of basis. The proof can be found in [1]. Let us now look and see how these elements combine together.

Lemma 3.36. *Let $x_r(t) \in \mathfrak{L}(K)$ and $e_r \in \mathfrak{L}_r$ for $r \in \Phi$. Then $x_r(t)e_r = e_r$.*

Proof. We have that

$$x_r(t) = \exp(t.\mathrm{ad}(e_r)) = 1 + t.\mathrm{ad}(e_r) + \cdots + \frac{(t.\mathrm{ad}(e_r))^{n-1}}{(n-1)!}. \tag{3.14}$$

Now, by fact 3.32, we know that $[e_r, e_r] = 0$ since $2r \notin \Phi$, by definition of a root system. Hence all the terms of $x_r(t)$ cancel other than the first, and we have our proof. □

We can make similar arguments using the identities from fact 3.32 to obtain the following formulae:

Fact 3.37. *We have*

1. $x_r(t)e_r = e_r$,
2. $x_r(t)e_{-r} = e_{-r} + t.h_r - t^2.e_r$,
3. $x_r(t)h_s = h_s - A_{sr}t.e_r$,
4. $x_r(t)e_s = \sum_{i=0}^{q} M_{rsi}t^i.e_{ir+s}$,

where $M_{rsi} = \pm\binom{p+i}{p}$*, and* p *is defined by* r *and* s *via* N_{rs}*.*

The proofs are very similar to that of the previous lemma, and so we omit them.

We can now look at some subgroups of our Chevalley group to get a feel for its structure. Note that by definition $x_r(s)x_r(t) = x_r(s+t)$, and so we can consider the group generated by $x_r(t)$ for $t \in K$. Call it X_r. As we can see, there is an obvious isomorphism from X_r to K as an additive group. We call these X_r the root subgroups. From these we can create two more subgroups, one that is generated by X_r for $r \in \Phi^+$ and another generated by X_r for $r \in \Phi^-$.

We now give a small lemma detailing how elements $x_r(t)$ are normalised by other automorphisms of the Lie algebra.

Lemma 3.38. *Consider an element* $x_r(t)$ *as above. Then for some* ρ *an automorphism of our Lie algebra* $\mathfrak{L}$*, we have*

$$\rho x_r(t)\rho^{-1} = \exp(t.\mathrm{ad}(\rho e_r)).$$

Proof. This follows directly from the definition of $x_r(t)$, and properties of the Lie bracket. □

We now consider the Chevalley commutator formula, which gives us a value for the commutator of $x_r(u)$ and $x_s(v)$ for some $u, v \in K$, and linearly independent $r, s \in \Phi$. This is

$$[x_r(u), x_s(v)] = \prod_{i,j>0,\ ir+js\in\Phi} x_{ir+js}(C_{ijrs}(-v)^i u^j), \tag{3.15}$$

where the constant $C_{ijrs} \in \{\pm 1, \pm 2, \pm 3\}$, and the terms x_{ir+js} occur in increasing order of $i+j$. The proof of this, as with quite a few of these identities, is long but not particularly difficult. It uses case analysis of different root combinations, combined with a few small case-specific lemmas.

We can use this formula to aid us in discovering the structure of the subgroup generated by X_r for positive roots r. Call this subgroup U. The structure of U is nice, as we have a unique representation of every $u \in U$.

Theorem 3.39. U *is nilpotent, and has central series of the form*

$$U = U_1 \supset U_2 \supset \cdots \supset U_{h+1} = 1, \tag{3.16}$$

where U_i *is the subgroup of* U *generated by* X_r *with* $h(r) \geq i$*, and* h *is the highest height of a root in*

Φ. *Also, the elements of* U *are uniquely expressible in the form*

$$\prod_{r\in\Phi^+} x_r(t_r). \tag{3.17}$$

To prove this, we need one lemma, which we state here without proof.

Lemma 3.40. *Let* $r \in \Phi$. *Then* h_r *in* $\mathfrak{L}_K$ *is non-zero.*

We are now in a position to prove our claim.

Sketch proof of theorem 3.39. To prove the first part, we note that $U_i \trianglelefteq U$ by using the commutator formula and considering the heights of the roots. To show that it has such a central series, we consider the quotient map

$$\psi : U \to U/U_{m+n},$$

and look at the image of the commutator formula under such ψ, to give us that $x_s(t)$ and $x_r(t')$ commute for $h(r) \geq m$ and $h(s) \geq n$.

The second part uses the commutator formula to give us a nice representation in terms of increasing root size, which gives existence.

Finally, uniqueness follows from a decreasing induction on root size, and the action of u (with two such representations) on the basis elements e_{-r}. This is where the previous lemma comes in, since we consider

$$ue_{-r} = e_{-r}t_r h_r + x$$

for $x \in \sum_{r\in\Phi^+} \mathfrak{L}_r$. The lemma then tells us $h_r \neq 0$ and using this we can show that the representations are equal, using the decomposition of the root system. The full proof can be found in [1]. □

Corollary 3.41. *We may define an isomorphism between* U *and the free* K*-module taking* Φ^+ *as its basis; define for some* $u = \prod_{r\in\Phi^+} x_r(t_r)$ *in* U

$$\prod_{r\in\Phi^+} x_r(t_r) = \phi(t_r)_{r\in\Phi^+}. \tag{3.18}$$

Proof. The previous theorem gives the uniqueness of the representation in U, and so ϕ is an isomorphism. □

We now state a few facts about the subgroup $\langle X_r, X_{-r}\rangle$, which will help us in defining the Bruhat decomposition of the Chevalley group, which will turn up again in chapter 4. It turns out that this subgroup is the image of $SL_2(K)$ under a certain homomorphism, and this is what we now explore.

The first thing to note is that $SL_2(K)$ is generated by the elements

$$\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \text{ and } \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix}.$$

There seems to be an obvious choice for the images of these matrices under our homomorphism, that being

$$\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \mapsto x_r(t) \text{ and } \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix} \mapsto x_{-r}(t).$$

Fact 3.42. *It turns out that this indeed defines a homomorphism, but we omit the proof, which can be found in [1].*

Now, under this homomorphism, we consider two images; the image of the matrices

$$\begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix}$$

for $\lambda \in K$, and the image of the matrix

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

We call the images of these matrices under the homomorphism mentioned $h_r(\lambda)$ and n_r respectively. To see how these act upon the Chevalley basis, we consider the action of $SL_2(K)$ on polynomial spaces of polynomials of homogeneous degree.

Definition 3.43. *Let $\mathbb{C}[x,y]_q$ be the ring of complex polynomials in two variables, x and y, homogeneous of degree q, and let $v_i = x^i y^{q-i}$. Then polynomials in $\mathbb{C}[x,y]_q$ will be of the form*

$$\sum_{i=0}^{q} z_i v_i \tag{3.19}$$

for $z_i \in \mathbb{C}$.

The operation we consider by an element

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

is $x \mapsto ax + by$ and $y \mapsto cx + dy$, then extend linearly to $\mathbb{C}[x, y]_q$.

Example 3.44. *Consider $\mathbb{C}[x, y]_2$ with basis $-x^2, 2xy, y^2$. Then we have the following actions;*

$$\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \text{ sends } \begin{array}{rcl} -x^2 & \mapsto & -x^2 \\ 2xy & \mapsto & 2xy - 2t(-x^2) \\ y^2 & \mapsto & y^2 + t(2xy) - t^2(-x^2). \end{array}$$

$$\begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix} \text{ sends } \begin{array}{rcl} -x^2 & \mapsto & -x^2 - t(2xy) - t^2y^2 \\ 2xy & \mapsto & 2xy + 2ty^2 \\ y^2 & \mapsto & y^2. \end{array}$$

Note that these actions are the same as those of $x_r(t)$ and $x_{-r}(t)$ on the elements e_r, h_r, e_{-r}, which form a basis of the space $\mathfrak{L}_r \oplus \mathfrak{C} \oplus \mathfrak{L}_{-r}$. This is in fact true in general, a fact which is used to prove the existence of the earlier homomorphism.

We can now study $h_r(\lambda)$ and n_r more closely by looking at the way they act on elements of the Chevalley basis. This is completely determined by lemmas 3.46 and 3.47. First we need one more proposition, which determines the actions of $x_r(t)$ on vectors e_{ir+s} where $ir + s$ is a root, and so a member of an r-chain. We omit the proof, but it relies on relations of structure constants as briefly talked about in fact 3.32.

Proposition 3.45. *Consider r and s independent roots, so that there is an r-chain through s of the form $s, r + s, \ldots, qr + s$. Then we have*

$$x_r(t)e_{ir+s} = \sum_{j=0}^{q-i} \delta_i \delta_{i+1} \ldots \delta_{i+j-1} \binom{i+j}{j} t^j e_{(i+j)r+s},$$

and

$$x_{-r}(t)e_{ir+s} = \sum_{j=0}^{q-i} \delta_{i-1} \delta_{i-2} \ldots \delta_{i-j} q - \binom{i+j}{j} t^j e_{(i-j)r+s},$$

where δ_i is defined by $N_{r,ir+s} = \delta_i(i + 1)$.

Lemma 3.46. *We have that $h_r(\lambda)$ acts on the Chevalley basis in the following manner*

$$h_r(\lambda)h_s = h_s \ \forall s \in \Delta,$$

$$h_r(\lambda)e_s = \lambda^{A_{rs}}e_s \ \forall s \in \Phi,$$

where A_{rs} is given by fact 3.32.

Proof. Omitted □

Lemma 3.47. *We have that n_r acts on the Chevalley basis in the following manner*

$$n_r h_s = h_{w_r(s)} \ \forall s \in \Delta,$$

$$n_r e_s = \varepsilon_{rs} e_{w_r(s)} \ \forall s \in \Phi \text{ and } \varepsilon_{rs} = \pm 1.$$

Proof. Let s and r be linearly independent roots. Then the action of n_r on e_s is the same as the action of the associated matrix on $u_i = \delta_0 \dots \delta_{i-1}\binom{q}{i}v_i$. Hence, we see that

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}(v_i) = (-1)^i v_{q-i},$$

by definition of the operation of $SL_2(K)$ on the polynomial ring.

Then we may instead look at the action on u_i, and it is easy to see that we have

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}(u_i) = (-1)^i \frac{\delta_0 \dots \delta_{i-1}}{\delta_0 \dots \delta_{q-i-1}} u_{q-i}.$$

Since the operations are the same, we may now see that $n_r e_s = \varepsilon e_{(q-i)r+s} = c e_{w_r(s)}$ by definition of w_r, the reflection, where $\varepsilon = (-1)^i \frac{\delta_0 \dots \delta_{i-1}}{\delta_0 \dots \delta_{q-i-1}} = \pm 1$.

Also, we have that n_r operates on e_r, h_r, e_{-r} in the same way as the associated matrix acts on the basis $-x^2, 2xy, y^2$. So plugging in the values, we can see that $n_r e_r = -e_{-r}$, $n_r e_{-r} = -e_r$ and $n_r h_r = -h_r$. Noting that $\langle X_r, X_{-r} \rangle$ acts as identity on $h \in \mathfrak{C}$ such that $(h_r, h) = 0$, we see that we have $n_r h_s = h_{w_r(s)}$ for $s \in \Phi$. □

Now, we can define a new subgroup of $\mathfrak{L}(K)$. We let H be the subgroup generated by $h_r(\lambda)$ for

$r \in \Phi$ and $\lambda \in K \setminus \{0\}$. Then by lemma 3.46, this subgroup operates trivially on $\mathfrak{C}$ and preserves $\mathfrak{L}_r$.

Now, we can take the $\mathbb{Z}$-span of the roots Φ to give us an additive abelian group P (the root lattice), with basis Δ. This gives us a definition of a K-character.

Definition 3.48. *A K-character of P is a homomorphism from P onto the multiplicative group K^*. Since it is a homomorphism, it is completely determined by its values on the basis Δ. For ease of use, we shall just call K-characters characters.*

So, we will look in particular at maps of the form $\chi_{\lambda,s} : r \mapsto \lambda^{A_{sr}}$. This map can be extended linearly to give a character. Also, characters can give rise to some automorphisms of $\mathfrak{L}_K$ by fixing pointwise $\mathfrak{C}$ and mapping every e_s to $\chi(s)e_s$. Note that if $\chi = \chi_{\lambda,s}$, then this automorphism is precisely $h_s(\lambda)$. Hence we may say that our group H is a subgroup of the group of automorphisms defined by characters in the way shown above. Call this group $\hat{H}$.

So the obvious question is which χ give rise to elements of H? To answer this, we require the concept of fundamental weights $\{q_1, \ldots, q_n\}$, which are a basis of the Cartan subalgebra that are in some respect dual to the base Δ of the root system Φ. Indeed, if $\Delta = \{r_1, \ldots, r_n\}$, then the corresponding elements of the Cartan subalgebra $h_{r_1}, \ldots, h_{r_n}$ are related via the Killing form by $(h_{r_i}, q_j) = \delta_{i,j}$, the Kronecker delta function. We also have that, since $\{q_1, \ldots, q_n\}$ is a base,

$$r_i = \sum_{j=1}^{n} \alpha_j q_j$$

and, in fact, every α_j is integral. The proof of this is found in [1].

Now, let Q be the $\mathbb{Z}$-span of these fundamental weights. Obviously now we see that $P \leq Q$ by the above claim. We may consider characters of Q and P. Obviously characters of Q are also characters of P, but not necessarily vice versa. In fact, the characters that are are precisely the ones that are associate to members of H.

Theorem 3.49. *The characters χ of P that extend to characters of Q are precisely those characters that give rise to automorphisms in H.*

Proof. Firstly, consider some element $h_r(\lambda) \in H$. Then by the observation above, we see that $h_r(\lambda)$ is the automorphism generated by the character $\chi_{\lambda,r}$, which gives us the first implication direction, since $\chi_{\lambda,r}$ can be extended to a character of Q in the obvious way.

Now, consider some character χ of Q. Take the values on the basis; $\chi(q_i) = \phi_i$ for $i \leq l$. We can write this as the product of P characters, that is to say consider χ_{ϕ_i,p_i} and note that it takes the values

we require;

$$\chi_{\phi_i,p_i}(q_j) = \begin{cases} \phi_i & \text{if} \quad i = j \\ 1 & \text{if} \quad i \neq j \end{cases}$$

Hence, $\chi = \prod_i^l \chi_{\phi_i,p_i}$ and hence the automorphism given by χ is the product of automorphisms in H, and hence is in H as required. □

We are now in a position to define our next subgroup of the $\mathfrak{L}(K)$. Consider the group generated by H and the elements n_r for $r \in \Phi$. We call this group N. We immediately dive into our main theorem regarding N.

Theorem 3.50. *There exists a homomorphism $\Psi : N \to W$ such that $n_r \mapsto w_r$ with kernel H.*

Proof. Omitted □

So, what does this tell us about the structure of N? Well, it tells us that $N/H \cong W$ and so $H \trianglelefteq N$ is normal in N.

Corollary 3.51. *$U \cap N = 1$.*

Proof. Consider some $\phi \in U \cap N$. Then we consider the action of ϕ on some e_s. Now, since $\phi \in N$ we have that $\phi e_s = \varepsilon_{r,s} e_{w_r(s)}$ for some $r \in \Phi$. Now, since $\phi \in U$, we have that $\phi e_s = e_s + x$ for some $x \in \mathfrak{C} \oplus \sum_{r \prec s} \mathfrak{L}_r$ by lemma 3.37.

Combining these, we see $\varepsilon_{r,s} = 1$, $x = 0$ and $w_r(s) = 1$. But this is true for all s, hence $w_r = 1$. This gives us that ϕ is not of the form n_r by lemma 3.47 and so it must be that $\phi \in H$. But, $U \cap H = 1$ and we get the required result. □

We can now come onto the Bruhat decomposition of Chevalley groups, culminating in a formula that gives us the structure and size of the finite Chevalley groups. Firstly, we need two definitions. The first is that of the Borel subgroup.

Definition 3.52. *A Borel subgroup of an algebraic group is a maximally connected solvable subgroup. A theorem by Borel tells us that any two of these are conjugate.*

The second is that of a (B, N)-pair.

Definition 3.53. *A group G has a (B, N)-pair if*

- *G is generated by B and N,*

- $B \cap N \trianglelefteq N$,
- $W = N/B \cap N$ *is generated by involutions* $w_1, \ldots, w_l$, *and is in fact a Coxeter group,*
- *for* $n_i \in N$ *the preimage of* $w_i \in W$ *under the homomorphism described in fact 3.50 and* $n \in N$ *then we have that*

$$Bn_iB.BnB \subseteq Bn_inB \cup BnB,$$

- n_i *as above does not normalise* B.

As the reader may have guessed, Chevalley groups have a (B, N)-pair. N is as described earlier, and we may take B to be the subgroup UH. As the notation suggests, this is a Borel subgroup, although we do not prove this here.

Theorem 3.54. *Chevalley groups have a* (B, N)*-pair.*

Proof. To prove the first point, we show that $X_r \subseteq \langle B, N\rangle$ and since $\langle X_r\rangle_{r\in\Phi}$ generate $\mathfrak{L}(K)$, this is sufficient. Now, since Φ is a root system, we have that every root is the image of a root in the basis under the action of some member of the Weyl group W. Hence we have $X_r = X_{w(r')}$ for some $r' \in \Delta$. Now, note that

$$\begin{aligned} n_r x_s(t) n_r^{-1} &= n_r.\exp(\mathrm{ad}(te_s))n_r^{-1} \\ &= \exp(\mathrm{ad}(n_r te_s)) \text{ by lemma 3.38} \\ &= \exp(\mathrm{ad}(\varepsilon_{rs} te_{w_r(s)})) \text{ by lemma 3.47} \\ &= x_{w_r(s)}(\varepsilon_{rs} t). \end{aligned}$$

This shows us that, in particular, $n_r X_s n_r^{-1} = X_{w_r(s)}$. Hence, $X_{w(r')} = n_r X_{r'} n_r^{-1}$. But r' is a base root and so belongs to Φ^+ and hence $X_{r'} \subseteq U \subseteq B$ and so is contained in $n_r B n_r^{-1}$ and so in $\mathfrak{L}(K)$ as required.

The other parts are fairly easy to prove; note that $B \cap N = UH \cap N = H$ by corollary 3.51 and is normal in N by the preceding theorem, giving us the second axiom.

The third comes from the same theorem, which gives us that the W mentioned in definition 3.53 is precisely the Weyl group, which is generated by involutions.

If $r \in \Delta$, then $X_r \subseteq U \subseteq B$, but $n_r X_r n_r = X_{-r} \nsubseteq B$, giving part five.

Part four follows directly from the next lemma, whose proof we omit. □

Lemma 3.55. *Take $r \in \Delta$, $n \in N$ and w the pre-image of n under our homomorphism. Then*

$$BnB.Bn_rB \subseteq Bnn_rB \cup BnB.$$

We now look at the parabolic subgroups of a group with a (B, N)-pair, which will gives us some more insight into the structure of our Chevalley group with respect to the subgroups B and N.

Definition 3.56. *A parabolic subgroup of a group with a (B, N)-pair is a subgroup that contains some conjugate of the Borel subgroup; $g^{-1}Bg$.*

We now look at specific parabolic subgroups, those of the form P_J. Consider the Weyl group W and its generators $\{w_1, \ldots, w_l\}$. take some $J = \{j_1, \ldots, j_{l'}\} \subseteq \{1, \ldots, l\}$ and consider the pre-image of $w_{j_1}, \ldots, w_{j_{l'}}$ under the homomorphism mentioned in theorem 3.50. Then this is a subgroup of N, we call N_J. Define P_J to be the group BN_JB.

Lemma 3.57. *Let G be a group with a (B, N)-pair. Then $P_J \leq G$.*

Proof. Omitted. □

Corollary 3.58. $G = BNB$.

Proof. Take $J = \{1, \ldots, l\}$, so $N_J = N$. Then the above holds and $BNB \leq G$. But $B, N \subseteq BNB$ and $\langle B, N \rangle = G$, so we have $G = BNB$. □

Now, it is worth noting that this decomposition is not, in itself, unique for a Chevalley group. However, there is a canonical form, which we will discuss further, after a theorem about the nature of the double cosets BnB of G.

Theorem 3.59. *Consider G a group with a (B, N)-pair. If $n, n' \in N$ with the same image in W under the homomorphism of 3.50, then we have that $BnB = Bn'B$. Hence we have a bijective map from double cosets of B and elements of the Weyl group.*

Proof. Omitted. □

Definition 3.60. *A set of roots Ψ of a root system Φ is called closed if for every $r, s \in \Psi$, $ir + js \in \Psi$ for all i, j positive integers such that $ir + js \in \Phi$.*

It turns out that every root system can decompose Φ^+ into two disjoint closed subsets, and this is achieved by considering the actions of the Weyl group on roots. Given an element w of the Weyl group of a root system Φ with the set of positive roots Φ^+ and negative roots Φ^-, we may define

$$\Psi_{1,w} = \{r \in \Phi^+ : w(r) \in \Phi^+\},$$

$$\Psi_{2,w} = \{r \in \Phi^+ : w(r) \in \Phi^-\}.$$

These are disjoint closed subsets whose union is Φ^+. We may now look at this with respect to the subgroup U, and we may take a similar disjunction. Define

$$U_w^+ = \prod_{r \in \Psi_{1,w}} X_r,$$

$$U_w^- = \prod_{r \in \Psi_{2,w}} X_r.$$

We claim that $U = U_w^+ U_w^-$ and $U_w^+ \cap U_w^- = 1$. The proof of this can be found in [1]. We can state our main theorem, giving us a canonical form for the decomposition of our Chevalley group, after a technical lemma.

Lemma 3.61. $n_w U_w^+ n_w^{-1} \subseteq U$.

Proof. By the beginning of the proof to theorem 3.53, we know that $n_r X_s n_r^{-1} = X_{w_r(s)}$. Now consider w as a product of involutions $w = w_{r_1} \dots w_{r_k}$ and then n_w and $n_{r_1} \dots n_{r_k}$ have the same image in W and so we may multiply one with an element of H (which is the kernel of the homomorphism) to give equality; $n_w = h n_{r_1} \dots n_{r_k}$.

Hence, we get

$$\begin{aligned} n_w X_s n_w^{-1} &= h n_{r_1} \dots n_{r_k} X_s n_{r_k}^{-1} \dots n_{r_1}^{-1} h^{-1} \\ &= h X_{w_{r_1 \dots w_{r_k}}(s)} h^{-1} \\ &= h X_{w(s)} h^{-1} \\ &= X_{w(s)}. \end{aligned}$$

Now, this means that $n_w U_w^+ n_w^{-1} = n_w \prod_{s \in \Psi_1} X_s n_w^{-1} \subseteq U$. □

Note that this proof holds analogously to give us $n_w U_w^- n_w^{-1} \subseteq V$.

Theorem 3.62. *Given some $w \in W$, take some n_w that maps into it under the homomorphism mentioned in theorem 3.50. Then each element $g \in G$ has the unique form $g = uhn_w u_w^-$ for $u \in U$, $h \in H$, $u_w^- \in U_w^-$.*

Proof. We firstly show existence. Consider the double coset Bn_wB. We may say

$$\begin{aligned} Bn_wB &= Bn_wHU \\ &= Bn_wHU_w^+U_w^- \\ &= BHn_wU_w^+U_w^- \text{ since } n_w^{-1}Hn_w = H \\ &\subseteq BUn_wU_w^- \text{ by lemma 3.61} \\ &= Bn_wU_w^-. \end{aligned}$$

However, now we see that $U_w^- \subseteq B$ and so we have equality here. So every element of G that is a member of a double coset can be written in the form we want, but that is precisely every element of G and we have the required form; bn_wu_w, which gives us $uhn_wu_w^-$.

Now we show uniqueness. Suppose that $u_1h_1n_wu_w = u_2h_2n_{w'}u_{w'}$. Then by theorem 3.59, we see $w = w'$ and $n_w = n_{w'}$. Hence $(u_2h_2)^{-1}u_1h_1 = n_{w'}u_{w'}u_w^{-1}n_w^{-1}$.

But now, note that $n_wU_w^-n_w^{-1} \subseteq V$ by lemma 3.61 and we know $B \cap V = 1$, and so we get the required result. □

The finite Chevalley groups are worth a brief mention here, since now we have such a precise decomposition for all the elements of $\mathfrak{L}(K)$, we can fairly easily discover the size of the finite Chevalley groups. So, we consider the case where K is a finite field $\mathbb{F}_q$.

Now, we know that every element of $\mathfrak{L}(\mathbb{F}_q)$ is contained in some double coset Bn_wB, and these are disjoint. Hence we get

$$\begin{aligned} |\mathfrak{L}(\mathbb{F}_q)| &= |\sum_{w\in W} Bn_wB| \\ &= |\sum_{w\in W} UHn_wU_w^-| \text{ by theorem 3.62} \\ &= |U||H|\sum_{w\in W} |U_w^-|. \end{aligned}$$

Let us examine each term individually. We know that U is, by definition, generated by the elements

$x_r(t)$ for $r \in \Phi^+$ and $t \in \mathbb{F}_q$ and, from equation 3.17, we can see that

$$|U| = q^{|\Phi^+|},$$

since elements of U are uniquely expressible in the form given. Obviously, the size of U_w^{-1} is closely related; looking at the definition in equation 3.4, we see that the order is q to the power of $|\Phi^+ \cap \{r \in \Phi^+ : w(r) \in \Phi^-\}|$. But as we have already seen, this is precisely the length of a Weyl group element, $l(w)$. Hence we have

$$|U_w^-| = q^{l(w)}.$$

So lastly, we need the order of H. As we have seen, H is generated by the automorphisms of $\mathfrak{L}(q)$ associated to characters of P that can be extended to Q as in theorem 3.49. So for a character, each generator of P, that is Δ, can be mapped to $q-1$ possibilities, as χ maps into $\mathbb{F}_q$. Since we have l generators, this would give us an order of $(q-1)^l$. However, not all of these characters can be extended to Q; this is actually the order of the group $\hat{H}$. So consider the surjective map $\hat{H} \to H$ given by restricting the characters down to P. It can be shown that the kernel of this map is isomorphic to the characters of the group Q/P. Say this has order d. Then we have

$$|H| = \frac{1}{d}(q-1)^l.$$

Combining all this gives us the last theorem of this section.

Theorem 3.63. *Consider the finite field $\mathbb{F}_q$. Then the Chevalley group of type $\mathfrak{L}$ over this field has order given by*

$$|\mathfrak{L}(\mathbb{F}_q)| = \frac{1}{d} q^{|\Phi^+|}(q-1)^l \sum_{w \in W} q^{l(w)}.$$

Chapter 4

Geometry over $\mathbb{F}_1$

This section of the paper will review the Connes and Consani paper, which culminates in proving that Chevalley group schemes determine varieties over an extension of the field of one element. To do this, we give a quick theorem relating to the structure of a Chevalley group scheme $\mathfrak{G}$ of definition 2.55, and then define a variety over $\mathbb{F}_1$.

Definition 4.1. *If A is commutative ring, we can define the rational points over A of $\mathfrak{G}$ by $\mathfrak{G}(A)$, the set*

$$\mathrm{Hom}(\mathrm{Spec}\, A, \mathfrak{G}).$$

We may define identical sets for rational points of A over $\mathscr{T}$ and $\mathscr{N}$, which correspond to the subgroups H and N of section 3 respectively.

These groups associate to those in the Chevalley group, and from this we get that $\mathscr{N}(A)/\mathscr{T}(A) \cong W$, the Weyl group of the root system.

This allows us to give one more theorem, due to Tits, that gives us insights into the structure of $\mathscr{N}(A)$ and $\mathscr{T}(A)$, without considering the structure of $\mathfrak{G}$.

Theorem 4.2. *The group extension*

$$1 \to \mathscr{T}(A) \to \mathscr{N}(A) \xrightarrow{p} W \to 1$$

is isomorphic to the extension

$$1 \to \mathrm{Hom}(L, A^*) \to N_{A^*,-1} \xrightarrow{p} W \to 1,$$

where $N_{A^*,1}$ is the group defined in definition 3.24 for $(D, \varepsilon) = (A^*, -1)$, and p is the quotient map $p : \mathcal{N}(A) \to W$ with quotient $\mathcal{T}(A)$.

4.1 Varieties over $\mathbb{F}_1$

Definition 4.3. *Consider a triple $G = (\underline{G}, G_{\mathbb{C}}, e_G)$ where*

1. *$\underline{G}$ is a covariant functor from* **AbGroup** *(the category of finite abelian groups) to* **Set**,
2. *$G_{\mathbb{C}}$ is a variety over $\mathbb{C}$,*
3. *e_G is a natural transformation from the functor $\underline{G}$ to the functor $\mathrm{Hom}(\mathrm{Spec}\,\mathbb{C}[\_], G_{\mathbb{C}})$.*

We call this a gadget over $\mathbb{F}_1$. If $\underline{G}(D)$ is finite for all abelian groups D, we say that the gadget G is finite.

We may refine this definition a little to consider graded gadgets.

Definition 4.4. *A gadget G over $\mathbb{F}_1$ is graded when the associated covariant functor is graded and maps to $\mathbb{N}$-graded sets. We write*

$$\underline{G} = \coprod_{k \geq 0} \underline{G}^k : \mathbf{AbGroup} \to \mathbf{Set}.$$

As an example of this that we will come back to a few times, we consider the gadget defined by an affine variety over $\mathbb{Z}$.

Example 4.5. *Let V be an affine variety over $\mathbb{Z}$. Then we have that V defines a gadget $\mathcal{G}(V)$ as follows.*

1. *$\underline{V}$ is the covariant functor from* **AbGroup** *to* **Set** *given by $\underline{V}(G) = \mathrm{Hom}(\mathcal{O}, \mathbb{Z}[D])$,*
2. *$V_{\mathbb{C}}$ is the variety $V \otimes \mathbb{C}$,*
3. *e_G is the natural transformation from the functor $\underline{V}$ to the functor $\mathrm{Hom}(\mathrm{Spec}\,\mathbb{C}[\_], V_{\mathbb{C}})$ defined by applying the functor $\_ \otimes_{\mathbb{Z}} \mathbb{C}$ giving a natural inclusion.*

For part 3, note that by lemma 2.36, we have an injection $\mathrm{Hom}(\mathrm{Spec}\,\mathbb{C}[\_], V_{\mathbb{C}})$ into $\mathrm{Hom}(\mathcal{O}_{\mathbb{C}}, \mathbb{C}[\_])$, and we then cite [9] corollary 1.7.4 for the remaining inclusion.

Note here that given a variety V, we have defined it's associated ring by $\mathcal{O}$, its ring of regular functions. We will use this notation for all rings, as the context will show which variety we are taking a ring of.

This definition of a gadget is actually a refinement of one by Soulé in 1999, the most important difference being that Soulé defined his functor from abelian group rings, whereas these are defined more generally, from abelian groups.

We can now consider morphisms of gadgets. To do this we need two maps, one to map between functors and one to map between varieties over $\mathbb{C}$. A natural construction would be to require the functor map to be a natural transformation, and the map between the varieties to be a morphism. This would mean that that we would get a nice commutative diagram. Let the gadgets be given by $X = (\underline{X}, X_{\mathbb{C}}, e_X)$ and $Y = (\underline{Y}, Y_{\mathbb{C}}, e_Y)$, and the gadget morphism be $\phi = (\underline{\phi}, \phi_{\mathbb{C}})$. We get the commutative diagram

$$\begin{array}{ccc} \underline{X}(A) & \xrightarrow{\phi_A} & \underline{Y}(A) \\ \Big\downarrow{\scriptstyle e_X(A)} & & \Big\downarrow{\scriptstyle e_Y(A)} \\ \mathrm{Hom}(\mathrm{Spec}\mathbb{C}[A], X_{\mathbb{C}}) & \xrightarrow{\phi'_{\mathbb{C}}(A)} & \mathrm{Hom}(\mathrm{Spec}\mathbb{C}[A], Y_{\mathbb{C}}) \end{array}$$

where $\phi'_{\mathbb{C}}(A)(f) = f \circ \phi_{\mathbb{C}}$. We now need one more definition before defining a variety over $\mathbb{F}_1$. We define an immersion, which is a refinement of the notion of a gadget morphism.

Definition 4.6. *An immersion is a gadget morphism $\phi = (\underline{\phi}, \phi_{\mathbb{C}})$ as above such that $\underline{\phi}(A)$ is injective for all $A \in$ **AbGroup**, and $\phi_{\mathbb{C}}$ is an embedding.*

We now give the definition of an affine variety over $\mathbb{F}_1$ as given by [3].

Definition 4.7. *An affine variety over $\mathbb{F}_1$ is a triple $(X, X_{\mathbb{Z}}, i)$ for a finite graded gadget X, an affine variety over $\mathbb{Z}$, $X_{\mathbb{Z}}$, and an immersion of gadgets $i : X \to \mathcal{G}(X_{\mathbb{Z}})$ such that for any affine variety V, and gadget morphism $\rho : X \to \mathcal{G}(V)$, there exists a unique morphism of gadgets $\rho_{\mathbb{Z}} : \mathcal{G}(X_{\mathbb{Z}}) \to \mathcal{G}(V)$ such that*

$$\rho = \rho_{\mathbb{Z}} \circ i.$$

We now give two examples of varieties over $\mathbb{F}_1$. The first is an affine gadget, used later in the paper.

Example 4.8. ***The Affine Case***

Let F be a finite set. We consider the graded functor

$$\underline{\mathbb{A}}^F : \mathbf{AbGroup} \to \mathbf{Set}, \tag{4.1}$$

graded by

$$\underline{\mathbb{A}}^F(D)^{(k)} = \coprod_{\substack{Y \subset F, \\ |Y|=k}} D^Y, \tag{4.2}$$

along with a variety over $\mathbb{C}$; $\mathbb{C}^F$, and the natural transformation

$$e_F : \underline{\mathbb{A}}^F \to \mathrm{Hom}(\mathrm{Spec}\mathbb{C}[\_], \mathbb{C}^F). \tag{4.3}$$

Note that $\mathrm{Hom}(\mathrm{Spec}\mathbb{C}[\_], \mathbb{C}^F) \cong \mathrm{Hom}(\mathbb{C}[t_j]_{j\in F}, \mathbb{C}[\_])$*, and so we can define e_F as follows*

$$e_F(D)(d_j)_{j\in Y}(t_i)_{i\in F} = (\alpha_i)_{i\in F} \text{ where } \alpha_i = \begin{cases} \chi(d_i) & \text{if} \quad i \in Y \\ 0 & \text{if} \quad i \notin Y \end{cases} \tag{4.4}$$

for any $\chi \in \mathrm{Spec}\,\mathbb{C}[D]$ a character.

The gadget defined by the three conditions above defines a variety over $\mathbb{F}_1$.

For this to be true, we require a variety over $\mathbb{Z}$ and an immersion i onto it as per definition 4.7. In this example we take our $\mathbb{Z}$-variety to be $X_\mathbb{Z} = \mathrm{Spec}(\mathbb{Z}[t_j]_{j\in F})$, which defines a gadget in the way described in example 4.5, and our immersion i to be the pair $(\underline{i}, i_\mathbb{C})$ as follows

$$\begin{aligned} \underline{i} &: \underline{\mathbb{A}}^F \to \mathrm{Hom}(\mathbb{Z}[t_j]_{j\in F}, \mathbb{Z}[\_]) \\ i_\mathbb{C} &: \mathbb{C}^F \to \mathbb{C}^F, \end{aligned} \tag{4.5}$$

where $i_\mathbb{C}$ is identity. This then is obviously an embedding. It is also easy to see that $\underline{i}$ is injective, for it is obvious that we have "enough" t_j unknowns to completely determine D with all homomorphisms since $|Y| < |F|$.

Lemma 4.9. *The gadget $\mathbb{A}^F := (\underline{\mathbb{A}}^F, \mathbb{C}^F, e_F)$ with the variety $X_\mathbb{Z}$ over $\mathbb{Z}$ and immersion i define a variety over $\mathbb{F}_1$.*

Proof. Consider some affine variety V, and let $\mathcal{G}(V) = (\underline{V}, V_\mathbb{C}, e_V)$ be the associated gadget. Assume

that there exists some morphism of gadgets $\phi : \mathbb{A}^F \to \mathcal{G}(V)$, defined by a pair as above, where

$$\begin{aligned} \underline{\phi} &: \underline{\mathbb{A}}^F \to \ \mathrm{Hom}(\mathcal{O}, \mathbb{Z}[\_]), \\ \phi_{\mathbb{C}} &: \mathbb{C}^F \to V_{\mathbb{C}}. \end{aligned} \tag{4.6}$$

Now, by lemma 2.36, and due to the fact that $\mathbb{C}^F \cong \mathrm{Spec}\mathbb{C}[t_j]_{j\in F}$ and $V_{\mathbb{C}} \cong \mathrm{Spec}\mathcal{O} \otimes \mathbb{C}$, we can see that $\phi_{\mathbb{C}}$ induces a new map,

$$\widetilde{\phi} : \mathcal{O} \otimes \mathbb{C} \to \mathbb{C}[t_j]_{j\in F}. \tag{4.7}$$

Consider some $f \in \mathrm{Hom}(\mathbb{C}[t_j]_{j\in F}, \mathbb{C}[D])$ and define $\psi(f) := f \circ \widetilde{\phi}$. Since the e functors are natural transformations, by the definition of a morphism of gadgets, we have the following commutative diagram

$$\begin{array}{ccc} \underline{\mathbb{A}}^F(D) & \xrightarrow{\quad\underline{\phi}\quad} & \mathrm{Hom}(\mathcal{O}, \mathbb{Z}[D]) \\ \Big\downarrow e_F(D) & & \Big\downarrow e_V(D) \\ \mathrm{Hom}(\mathbb{C}[t_j]_{j\in F}, \mathbb{C}[D]) & \xrightarrow{\quad\psi\quad} & \mathrm{Hom}(\mathcal{O} \otimes \mathbb{C}, \mathbb{C}[D]). \end{array}$$

Now, consider

$$e_V(D) \circ \underline{\phi}(d)(a \otimes 1) \tag{4.8}$$

for some $a \in \mathcal{O}$ and $d \in D$. Note that we may think of $a \otimes 1$ as being in $\mathcal{O}$ by the inverse of the inclusion map

$$\begin{aligned} g : \mathcal{O} &\to \mathcal{O} \otimes \mathbb{C} \\ a &\mapsto a \otimes 1. \end{aligned} \tag{4.9}$$

But note that $e_V(D)(f)(a \otimes c) = f(a) \otimes c$ and so we have equation 4.8 becomes $\underline{\phi}(d)(a) \otimes 1 \in \mathbb{Z}[D]$ by the obvious inclusion.

Now, given the commutativity of the diagram we have that

$$e_V(D) \circ \underline{\phi}(d) = \psi \circ e_F(D)(d), \tag{4.10}$$

and so we consider the right hand side. Then by definition of ψ we have

$$\begin{aligned}\psi \circ e_F(D)(a) &= e_F(D)(d_i)_{i\in Y} \circ \widetilde{\phi}(a) \\ &= e_F(D)(d_i)_{i\in Y}(P),\end{aligned} \tag{4.11}$$

for some complex polynomial $P \in \mathbb{C}[t_j]_{j\in F}$. Now, by equation 4.4, we have $e_F(D)(d_i)_{i\in Y}(P) = P(\alpha_j)_{j\in F}$. Since the left hand side of equation 4.10 is in $\mathbb{Z}[D]$, we have that $P(\alpha_j)_{j\in F} \in \mathbb{Z}[D]$.

What we now need is for all the coefficients of P to be in $\mathbb{Z}$, and to show this we do a double induction, firstly on $|F|$ and then on a polynomial degree. So, assume that for all sets of size less than $|F|$, the coefficients of P are in $\mathbb{Z}$. Now, note that since $Y \subsetneq F$, we have that, by definition of the α_i, at least one α_i is zero for some i. Hence we have the associated polynomial $P_i \in \mathbb{C}[t_{j_1}, \ldots, t_{j_{|F|-1}}]$, which is the polynomial for a set of size $|F| - 1$, and so by induction hypothesis all coefficients of P_i are in $\mathbb{Z}$.

Now, we have a decomposition of the polynomial P to the form

$$P = P_i + g. \tag{4.12}$$

We now induct on the highest degree of t_i in g. If the highest degree of t_i is zero, then g is the zero polynomial and we are done. So assume that for all h where the degree of t_i in h is less than that of g, we have h has integer coefficients. But note that by its construction $g = t_i g'$, and so we can use our induction hypothesis on g' to get that all coefficients are in $\mathbb{Z}$, which completes the claim; $P \in \mathbb{Z}[t_j]_{j\in F}$.

This shows that $\widetilde{\phi}\Big|_{\mathcal{O}} : \mathcal{O} \to \mathbb{Z}[t_j]_{j\in F}$. Now, what does this mean? Well, for the gadget morphism ϕ we need to show, by definition 4.7, that there is another morphism $\phi_{\mathbb{Z}}$ such that $\phi = \phi_{\mathbb{Z}} \circ i$. So, $\phi_{\mathbb{Z}}$ is, as before, a pair, and from the above argument, we can see that

$$\phi_{\mathbb{Z}} = (\widetilde{\phi}\Big|_{\mathcal{O}}, \phi_{\mathbb{C}}). \tag{4.13}$$

This is the required gadget morphism, which completes the proof, as uniqueness is given by construction. □

The next example is that of the variety $\operatorname{Spec} D$ for some abelian group D.

Example 4.10. *For* $\operatorname{Spec} D$ *to be a variety, we need three things; the gadget, the variety over* $\mathbb{Z}$ *and*

an immersion.

For the gadget, we take the functor $\operatorname{Spec} D$ *given by* $\underline{SpecD}(D') = Hom(D, D')$ *and the variety over* $\mathbb{C}$ *to be* $\operatorname{Spec} \mathbb{C}[D]$*. The natural transformation then takes the functor* $Hom(D, D')$*, as per the definition, to* $\operatorname{Hom}(\operatorname{Spec} \mathbb{C}[D'], \operatorname{Spec} \mathbb{C}[D])$*.*

Now, by lemma 2.36, we have that

$$\operatorname{Hom}(\operatorname{Spec} \mathbb{C}[D'], \operatorname{Spec} \mathbb{C}[D]) = \operatorname{Hom}(\mathbb{C}[D], \mathbb{C}[D']),$$

and so we can make the natural transformation do the obvious thing; take some $f \in \operatorname{Hom}(D, D')$ *and extend in the standard way to a map* $f' \in \operatorname{Hom}(\mathbb{C}[D], \mathbb{C}[D'])$*.*

The variety over $\mathbb{Z}$*, written as* $\operatorname{Spec}(D)_{\mathbb{Z}}$ *is given by the functor*

$$\underline{\operatorname{Spec}(D)}_{\mathbb{Z}}(D') = \operatorname{Hom}(\mathbb{Z}[D], \mathbb{Z}[D']), \tag{4.14}$$

along with the variety over $\mathbb{C}$*, which is* $\mathbb{C}[D]$*, and then the natural transformation which we define analogously to that of a gadget* $\operatorname{Spec} D$*:*

$$e_{\operatorname{Spec}(D),\mathbb{Z}}(D') : \operatorname{Hom}(\mathbb{Z}[D], \mathbb{Z}[D']) \to \operatorname{Hom}(\mathbb{C}[D], \mathbb{C}[D']). \tag{4.15}$$

We also require an immersion. The functor map is the obvious inclusion, and the map of varieties over $\mathbb{C}$ *is identity. These two maps trivially define an immersion.*

Lemma 4.11. *The gadget* $\underline{\operatorname{Spec}} D$ *can be realised as a variety over* $\mathbb{F}_1$*.*

Proof. We can use exactly the same proof technique as in lemma 4.9; drawing a commutative diagram and then showing that the restriction of the map from $\mathcal{O} \otimes \mathbb{C}$ to $\mathbb{C}[D']$ down to $\mathcal{O}$ maps into $\mathbb{Z}[D']$. □

Unfortunately, Chevalley groups schemes are not varieties over $\mathbb{F}_1$. However, it turns out that we can define them as varieties over the quadratic extension of $\mathbb{F}_1$. To do this, our gadget functor must map not from **AbGroup**, but from $\mathbf{AbGroup}^{(2)}$, the category of finite pointed abelian groups, where the distinguished points have order exactly 2.

Morphisms in $\mathbf{AbGroup}^{(2)}$ are obviously group homomorphisms than map distinguished points to distinguished points.

Consider now the ring $R_2 = \mathbb{Z}[T]/(T^2 - 1)$. Note that we have two homomorphisms into $\mathbb{Z}$; the one that maps T to 1, and the other mapping T to -1. This gives us that $\operatorname{Spec} R_2$ is actually two copies

of $\operatorname{Spec}\mathbb{Z}$. We only consider the copy given by the non-trivial homomorphism. We define

$$\beta[D,\varepsilon] = \mathbb{Z}[D] \otimes_{\mathbb{Z}[\mathbb{Z}/2\mathbb{Z}]} \mathbb{Z},$$

where we have $\varepsilon = -1$. We define $\mathbb{C}[D,\varepsilon]$ as the extension of scalars of $\beta[D,\varepsilon]$ into $\mathbb{C}$.

So what does this mean for characters of $\mathbb{C}[D,\varepsilon]$? Well, its characters are the characters of $\mathbb{C}[D]$ with the value of ε under them taking -1. Note that these separate the elements of D, so if there are $d_1 \neq d_2$ in D, then there is some character χ such that $\chi(d_1) \neq \chi(d_2)$.

We define a variety over $\mathbb{F}_{1^2}$ in a similar way to example 4.5, but making the functor map D to $\operatorname{Hom}(\mathcal{O}, \beta[D,\varepsilon])$, with corresponding inclusion natural transformation, and for the variety $V_{\mathbb{C}}$ to sit over the non-trivial copy of $\operatorname{Spec}\mathbb{Z}$ in $\operatorname{Spec} R_2$.

4.2 Chevalley Group Schemes as Varieties over $\mathbb{F}_{1^2}$

We now give the functor, complex variety and natural transformation required to define a gadget for our Chevalley group scheme. For the functor we take

$$\underline{G} : \mathbf{AbGroup}^{(2)} \to \mathbf{Set}$$

$$\underline{G}(D,\varepsilon) = \underline{\mathbb{A}}^{\Phi^+}(D) \times \coprod_{w \in W} (p^{-1}(w) \times \underline{\mathbb{A}}^{\Phi_w}(D)),$$

where p is the projection $p : N_{D,\varepsilon} \to W$ as in the definition of Tits' data $\{N, N_s, p\}_{s \in S}$. Note that this looks very similar to the Bruhat decomposition of the Chevalley group we gave in theorem 3.62. This will be of vital importance, as it allows us to break down our functors into nice affine ones.

For the complex variety, we take $G_{\mathbb{C}} = \mathfrak{G}(\mathbb{C})$, the set of rational points over $\mathbb{C}$ of $\mathfrak{G}$.

For our natural transformation, we require some

$$\begin{aligned} e_G : \underline{G} &\to \operatorname{Hom}(\operatorname{Spec}\mathbb{C}[\_], G_{\mathbb{C}}) \\ (D,\varepsilon) &\mapsto \operatorname{Hom}(\operatorname{Spec}\mathbb{C}[D,\varepsilon], G_{\mathbb{C}}). \end{aligned} \tag{4.16}$$

To define this, we reuse the affine natural transformation of example 4.8 to give three separate functors

$$e_{\Phi^+} : \underline{\mathbb{A}}^{\Phi^+}(D) \to \mathbb{C}^{\Phi^+},$$

$$e_{\Psi_{2,w}} : \underline{\mathbb{A}}^{\Psi_{2,w}}(D) \to \mathbb{C}^{\Psi_{2,w}},$$

$$e_N : N_{D,\varepsilon} \to N(\mathbb{C}),$$

where $\Psi_{2,w}$ is as defined in the discussion of definition 3.60.

Now, we can define $\mathscr{U}(A)$ generated by $x_r(t)$ for $t \in A$, the direct analogue of the group U in section 3.4, and similarly define $\mathscr{U}_w^-(A)$. From these we can get a decomposition of $\mathfrak{G}(K)$ into disjoint cells of the form

$$C_w = \mathscr{U}(K)\mathscr{T}(K)n_w\mathscr{U}_w^-(K).$$

The proof of this fact is identical to that of theorem 3.62.

Now, by corollary 3.41, we have that $\mathscr{U}(A)$ elements are in bijection with the group A^{Φ^+}, and by direct analogy that $\mathscr{U}_w^-(A)$ is in bijection with $A^{\Phi_w^-}$, where A is a field (or more generally, a commutative ring, but we do not need this generalisation here). So, since the two functors e_{Φ^+} and $e_{\Phi_w^-}$ map into $\mathbb{C}^{\Phi^+}$ and $\mathbb{C}^{\Phi_w^-}$ respectively, we can apply our isomorphism ϕ (from corollary 3.41) and pull them into $\mathscr{U}(\mathbb{C})$. This means we can define e_G;

$$e_G(d, n, d') = \phi(e_{\Phi^+}(d))e_N(n)\phi_w(e_{\Phi_w^-}(d')), \tag{4.17}$$

where we define ϕ_w analogously to ϕ. Since ϕ, ϕ_w and e_N all map into subgroups of $\mathfrak{G}(\mathbb{C})$, we see that this functor maps into $G_{\mathbb{C}}$ as required.

We need one last technical proposition of Chevalley before we can prove this gadget is in fact a variety.

Theorem 4.12. *Let $w_0 \in W$ be the unique element of the Weyl group that sends Φ^+ to $-\Phi^+$. Then the morphism*

$$\begin{gathered}\theta : \mathscr{U} \times p^{-1}(w_0) \times \mathscr{U} \to \mathfrak{G} \\ (u, n, v) \mapsto unv\end{gathered} \tag{4.18}$$

defines an isomorphism to an open dense subscheme Ω of $\mathfrak{G}$ with algebra of coordinates of the form

$$\mathcal{O}_\Omega = \mathcal{O}_{\mathfrak{G}}[d^{-1}], \tag{4.19}$$

with d taking value 1 on the lift of w_0 to $\mathfrak{G}(\mathbb{Z})$. We have $\mathcal{O}_\Omega = \mathcal{O}(\mathscr{U}) \otimes \mathbb{Z}[L] \otimes \mathcal{O}(\mathscr{U})$.

Note that by corollary 3.16, such a w_0 does indeed exist.

Theorem 4.13. *The gadget $G = (\underline{G}, G_{\mathbb{C}}, e_G)$ defines a variety over $\mathbb{F}_{1^2}$.*

Proof. Firstly we note that G is a finite graded gadget by its very construction. For the affine group scheme over $\mathbb{Z}$, we take the Chevalley group scheme over $\mathbb{Z}$; $\mathfrak{G}(\mathbb{Z}) = \mathfrak{G}$.

We now look at the gadget $\mathcal{G}(\mathfrak{G})$ over $\mathbb{F}_{1^2}$. We can define this as in example 4.5, with

1. $\underline{\mathfrak{G}}$ is the covariant functor from $\mathbf{AbGroup}^{(2)}$ to $\mathbf{Set}$ given by $\underline{\mathfrak{G}}(D, \varepsilon) = \mathrm{Hom}(\mathcal{O}, \beta[D, \varepsilon])$.
2. $\mathfrak{G}_{\mathbb{C}}$ is the variety $\mathfrak{G}(\mathbb{C})$.
3. e_G is the natural transformation from the functor $\underline{\mathfrak{G}}$ to the functor $\mathrm{Hom}(\mathrm{Spec}\,\mathbb{C}[\_], \mathfrak{G}_{\mathbb{C}})$, given by inclusion.

We now require an immersion of gadgets, $i = (\underline{i}, i_{\mathbb{C}})$. Obviously we may take $i_{\mathbb{C}}$ as identity, which gives us an embedding, and now we need for $\underline{i}$ to be injective. We do not show this here, but it comes from theorem 4.12.

Now, we need to show that for any other affine variety, V and gadget morphism $\rho : G \to \mathcal{G}(V)$, we have there exists some unique morphism $\rho_{\mathbb{Z}} : \mathfrak{G}_{\mathbb{Z}} \to \mathcal{G}(V)$ such that $\rho = \rho_{\mathbb{Z}} \circ i$.

To do this, we follow the same proof structure as that of lemma 4.9. So, consider some affine variety V of finite type over $\mathbb{Z}$ with gadget $\mathcal{G}(V) = (\underline{V}, V_{\mathbb{C}}, e_V)$. Assume that there exists some morphism of gadgets $\phi : G \to \mathcal{G}(V)$ where ϕ is a pair with

$$\begin{aligned} \underline{\phi} &: \underline{G} \to \underline{V}, \\ \phi_{\mathbb{C}} &: \mathfrak{G} \otimes \mathbb{C} \to V_{\mathbb{C}}. \end{aligned} \tag{4.20}$$

But, again as in lemma 4.9, we can see that ϕ induces a map $\widetilde{\phi}$ given by

$$\widetilde{\phi} : \mathcal{O}_{\mathbb{C}}(V) \to \mathcal{O}_{\mathbb{C}}(\mathfrak{G}). \tag{4.21}$$

Since the e functors are natural transformations, and ϕ a morphism, we have for any pair (D, ε) a commutative diagram

$$\begin{array}{ccc} \underline{G}(D,\varepsilon) & \xrightarrow{\underline{\phi}(D,\varepsilon)} & \mathrm{Hom}(\mathcal{O}(V), \beta(D,\varepsilon)) \\ \Big\downarrow{\scriptstyle e_G(D,\varepsilon)} & & \Big\downarrow{\scriptstyle e_V(D,\varepsilon)} \\ \mathrm{Hom}(\mathcal{O}_{\mathbb{C}}(\mathfrak{G}), \mathbb{C}[D,\varepsilon]) & \xrightarrow{\psi} & \mathrm{Hom}(\mathcal{O}_{\mathbb{C}}(V), \mathbb{C}[D,\varepsilon]), \end{array}$$

with $\psi(f) = f \circ \widetilde{\phi}$.

We need to show that $\phi_{\mathbb{C}}(\mathcal{O}(V)) \subset \mathcal{O}(\mathfrak{G})$. So, we can take $h \in \mathcal{O}(V)$ as last time, with $\widetilde{\phi}(h) = h'$.

Now, we can consider theorem 4.12. Note that elements of $\mathcal{O}_{\mathbb{C}}(\mathfrak{G})$ have trivial d in the algebra $\mathcal{O}_{\mathfrak{G}}[d^{-1}]$, and so their intersection is precisely $\mathcal{O}(\mathfrak{G})$. This means that if we can show that h' restricted to $\Omega \subset \mathfrak{G}$ is a member of $\mathcal{O}(\Omega) \subset \mathcal{O}_{\mathbb{C}}(\Omega)$, then $h' \in \mathcal{O}(\mathfrak{G})$.

To do this then, we consider the group $N_{\mathbb{Z}^*,-1}$ as defined in definition 3.24. Since this is isomorphic to the semidirect product of a torus and Weyl group W, we can take a lift of our w_0 back into $N_{\mathbb{Z}^*,-1}$ to give $p^{-1}(w_0) = w_0'\mathscr{T}$ for some torus.

Now, consider again our functor $\underline{G}(D, \varepsilon)$ and look at elements of the form

$$g \in C := \underline{\mathbb{A}}^{\Phi^+} \times p^{-1}(w_0) \times \underline{\mathbb{A}}^{\Phi^+}. \tag{4.22}$$

From our choice w_0', we can identify cosets of the form $\mathrm{Hom}(L, D)w_0'$. Now, L is a free abelian group, and so we may pick generators $l_1, \ldots, l_m$. These can be used to identify uniquely every $f \in \mathrm{Hom}(L, D)$, by their images. For example,

$$f(l) = f(l_{i_1} \ldots l_{i_j}) = f(l_{i_1}) \ldots f(l_{i_j}), \tag{4.23}$$

and so $f(l)$ is completely determined by the image of the generators under f. So, for every f we can identify a tuple $(d_1, \ldots, d_j)$ elements that are the images of the generators of L.

Also, remember that

$$\underline{\mathbb{A}}^{\Phi^+}(D) = D^{\Phi^+}, \tag{4.24}$$

which gives us that for any map y from $\Phi \cup \{1, \ldots, j\}$ to the abelian group D, we can define an element $g(y)$ in C, by

$$g(y) = (y_r)_{r \in \Phi^+} \times (y_l)_{l \in \{1,\ldots,j\}} \times (y_{-r})_{r \in \Phi^+}. \tag{4.25}$$

Obviously $g(y) \in \underline{G}(D, \varepsilon)$ and so we may consider this image under $\underline{\phi}$ to give an element in $\mathrm{Hom}(\mathcal{O}(V), \beta[D, \varepsilon])$, and we consider this mapping h, which shows

$$\underline{\phi}(D, \varepsilon)(g(y))(h) \in \beta[D, \varepsilon] \subset \mathbb{C}[D, \varepsilon]. \tag{4.26}$$

The last subset was by applying the obvious morphism $e_V(D, \varepsilon)$.

Now this is the same, by the commutative diagram, as the map

$$\psi(e_G(D,\varepsilon)(g(y)))(h), \tag{4.27}$$

where as before, $h \in \mathcal{O}(V)$.

Now, as we showed, we only need consider h' on $\Omega \subset \mathfrak{G}$; $h'\big|_\Omega$. Since we are working with rings over varieties, $h'\big|_\Omega$ is a polynomial.

Since we have that $\mathcal{O}(\Omega) = \mathcal{O}(\mathscr{U}) \otimes \mathbb{Z}[L] \otimes \mathcal{O}(\mathscr{U})$, our polynomial will have unknowns for each tensored term. The $\mathcal{O}(\mathscr{U})$ terms both have basis Φ^+, and $\mathbb{Z}[L]$ has basis $\{1,\dots,j\}$, but since it is a torus, every element must also be invertible. Hence we have that the polynomial is represented by

$$P(t_r, u_i, u_i^{-1}) \text{ with } r \in \Phi, i \in \{1,\dots,j\}, \tag{4.28}$$

since $2|\Phi^+| = |\Phi|$. Now, we consider the group D to be the group of maps from the set $\Phi \cup \{1,\dots,j\}$ to the cyclic group of order n for some $n \in \mathbb{Z}$, Cartesian product with the cyclic group of order 2 generated by ε. This means that we may represent D in the following way

$$D = (\mathbb{Z}/n\mathbb{Z})^{\Phi\cup\{1,\dots,j\}} \times \mathbb{Z}/2\mathbb{Z}.$$

Now, we can construct an algebra homomorphism

$$\theta_n : \mathbb{C}[t_r, u_i, u_i^{-1}] \to \mathbb{C}[(\mathbb{Z}/n\mathbb{Z})^{\Phi\cup\{1,\dots,j\}}]. \tag{4.29}$$

The images of t_r and u_i are pretty obvious; we let η be a generator of $\mathbb{Z}/n\mathbb{Z}$ and denote by η_k the element of $(\mathbb{Z}/n\mathbb{Z})^{\Phi\cup\{1,\dots,j\}}$ with an η in the k^{th} position, and 0 elsewhere. Then $\theta_n(t_r) = \eta_r$ and $\theta_n(u_i) = \eta_i$, and extend θ_n linearly to give a homomorphism.

Now, we already have some information about $\theta_n(P)$; since we have $\underline{\phi}(D,\varepsilon)(g(y))(h) \in \beta(D,\varepsilon)$, and so $\theta_n(P) \in \mathbb{Z}[(\mathbb{Z}/n\mathbb{Z})^{\Phi\cup\{1,\dots,j\}}]$; i.e. it has coefficients in $\mathbb{Z}$ for every n. From this, we need to deduce that P has coefficients in $\mathbb{Z}$.

Well, this is in fact much easier than in theorem 4.9, because here we can just consider a very large n. Indeed, define $\deg(t_r)_b$ to be the maximum degree of t_r in the b^{th} term of P. We define similarly $\deg(u_i)_b$. Then note that the image of u_i and t_r are linearly independent under θ_n (provided n is not a factor of $\deg(t_r)_b$ or $\deg(u_i)_b$ of any b). Hence for large enough n (under the same provisions of

coprimality between n and degrees of the unknowns), each term of P is linearly independent under θ_n.

So, we may choose a large n, such that every degree of u_i and t_r is coprime to it in every term. Such an n would be

$$n = \prod_{b,r,i} \deg(t_r)_b \deg(u_i)_b + 1 \tag{4.30}$$

if we had no u_i^{-1} terms. Writing such an n in this case just gets very ugly, but it is not hard to do once you have the order of the elements of D.

This means that every term is linearly independent and no term vanishes under θ_n. Hence all coefficients in $\theta_n(P)$ are identical to those in P, and so, since $\theta_n(P) \in \mathbb{Z}[(\mathbb{Z}/n\mathbb{Z})^{\Phi \cup \{1,\dots,j\}}]$, we have that $P \in \mathbb{Z}[t_r, u_i, u_i^{-1}]$, so $h'\big|_\Omega \in \mathcal{O}_\Omega$ as required; the $\mathbb{Z}$-morphism can now just be taken as a restriction of ϕ as in lemma 4.9. $\square$

Chapter 5

Conclusion

In this paper we have introduced classical algebraic geometry, and some basic theory of Lie algebras. We have studied the structure of Chevalley groups, and then considered them as group schemes. This additional observation allowed us to show that Chevalley group schemes can be defined as varieties over $\mathbb{F}_{1^2}$.

But where could one go from here? Well, in their paper, Connes and Consani do go on to briefly discuss schemes over $\mathbb{F}_1$, which are obviously the natural structure to consider after one looks at varieties. In defining this, they extend the general functor used in the definition of a variety from **AbGroup** to the scheme of abelian monoids. We still do not know fully how many properties of classical schemes roll over into schemes over $\mathbb{F}_1$.

There are still plenty of other things that could be done in this field; a possibility is looking into Kac-Moody groups. These could be defined as varieties over $\mathbb{F}_1$ in the same way as Chevalley groups can be.

The Cartan matrix for a simple Lie algebra is an $n \times n$ matrix with entries of the form

$$a_{ij} = 2\frac{(r_i, r_j)}{(r_i, r_i)} \tag{5.1}$$

for $r_i, r_j \in \Delta$. We can give a generalized version.

Definition 5.1. *A generalized Cartan matrix is an $n \times n$ matrix (a_{ij}) such that*

1. $a_{ii} = 2$.
2. $a_{ij} \leq 0$ *for all* $i \neq j$.
3. $a_{ij} = 0$ *if and only if* $a_{ji} = 0$.
4. $a_{ij} \in \mathbb{Z}$.

It is possible, for a standard Cartan matrix, to recover the Lie algebra it represents (this is discussed in [10]), so one naturally asks can we do something similar for a generalized Cartan matrix. The answer is yes, and the structures gained from such a venture are known as Kac-Moody algebras. They are Lie algebras, but are not quite a nice as the ones we have studied, since in general they are infinite dimensional. However, analogues of things like the root space decomposition still hold. From this we can get analogues of Chevalley groups, called Kac-Moody groups. Due to this connection, it could well be that we can define Kac-Moody group schemes as varieties over some extension of $\mathbb{F}_1$.

It is also suggested that the field of one element may be the key to solving the elusive Riemann hypothesis.

Definition 5.2. *The Riemann zeta function $\zeta(s)$ is given by*

$$\zeta(s) = \sum_{n=1}^{\infty} n^{-s}. \tag{5.2}$$

The Riemann hypothesis claims that if $\zeta(a) = 0$ then the real part of a is equal to $\frac{1}{2}$. This hypothesis has been around since 1859, when it was mentioned almost as an afterthought in Riemann's paper "On The Number Of Primes Less Than A Given Magnitude". Some progress has been made, in particular in 1948 André Weil proved an analogue of the hypothesis over function fields. It is conjectured that $\operatorname{Spec}\mathbb{Z}$ is a curve over $\operatorname{Spec}\mathbb{F}_1$, which would mean that $\operatorname{Spec}\mathbb{Z} \times_{\operatorname{Spec}\mathbb{F}_1} \operatorname{Spec}\mathbb{Z}$ makes sense, and could potentially provide a surface upon which to try and replicate Weil's proof. Unfortunately, due to the need for brevity in this paper we do not attempt such a proof here.

Bibliography

[1] R. W. Carter, *Simple Groups of Lie Type* Wiley, 1989.

[2] R. W. Carter, On The Representation Theory of the Finite Groups of Lie Type over an Algebraically Closed Field of Characteristic 0
Vol 77 *Algebra IX* Ed. A. I. Kostrikin and I. R. Shafarevich
Springer, 1992.

[3] A. Connes, C. Consani,
On the Notion of Geometry Over $\mathbb{F}_1$ http://www.alainconnes.org/docs/ak.pdf

[4] M. Demazure, A. Grothendieck *Schemas en Groupes (SGA 3), Tome 1.* Springer-Verlag, 1970.

[5] D. S. Dummit, R. M. Foote *Abstract Algebra, Third Edition.* John Wiley & Sons, Inc. 2004.

[6] D. Eisenbud, J. Harris, *The Geometry of Schemes (GTM 197)* Springer, 2000.

[7] K. Erdmann, M. J. Wildon, *Introduction to Lie Algebras* Springer, 2006.

[8] A. Gathmann, **Algebraic Geometry**
Notes for a class taught at the University of Kaiserslautern 2002/2003
http://www.mathematik.uni-kl.de/ gathmann/class/alggeom-2002/main.pdf [Online; accessed 24-February-2013]

[9] A. Grothendieck, J. Dieudenné *Éléments de Géométrie Algébrique I: Les language en Schémas.* Springer-Verlag, 1971.

[10] J. E. Humphreys. *Introduction to Lie Algebras and Representation Theory (GTM 9).* Springer, 1972.

[11] J. E. Humphreys, *Reflection Groups and Coxeter Groups* Cambridge University Press, 1990.

[12] M. Kapranov, A. Smirnov
Cohomology Determinants and Reciprocity Laws: The Number Field Case
http://www.neverendingbooks.org/DATA/KapranovSmirnov.pdf [Online; accessed 15-September-2013]

[13] D. Mumford, *The Red Book of Varieties and Schemes* Springer-Verlag, 1988.

[14] V. I. Piercey, The Functor of Points
http://math.arizona.edu/∼vpiercey/FunPoints.pdf [Online; accessed 02-September-2013]

[15] K. Ueno, *Algebraic Geometry 1* Translations of Mathematical Monographs, Volume 185 AMS, 1997.

[16] R.G. Underwood, *An Introduction to Hopf Algebras* Springer, 2011.

[17] W. C. Waterhouse, *Introduction to Affine Group Schemes (GTM 66)* Springer-Verlag New York Inc., 1979.

[18] E. Webber, Chevalley Groups
http://gr-tes.epfl.ch/~testerma/doc/projets/WeberChevalleygroups.pdf [Online; accessed 05-August-2013]

Appendix A

Category Theory

Here we give a (very) brief non-rigorous introduction to category theory. In many ways, category theory can be thought of as an abstraction of set theory, and a useful tool in much of modern day mathematics. A category is pair of collections; the objects of a category and the morphisms between them.

Definition A.1. *A category* $\mathscr{C}$ *is a pair* $(\mathrm{Ob}(\mathscr{C}), \mathrm{Hom}(\mathscr{C}))$*, where* $\mathrm{Ob}(\mathscr{C})$ *is a collection of objects and* $\mathrm{Hom}(\mathscr{C})$ *is the collection of morphisms between every two objects, which are subject to some axioms we state later. For every* $A \in \mathrm{Ob}(\mathscr{C})$*, we require the existence of the unique identity morphism* id_A *in* $\mathrm{Hom}(\mathscr{C})$*.*

We call a category small if both $\mathrm{Ob}(\mathscr{C})$ *and* $\mathrm{Hom}(\mathscr{C})$ *are proper sets.*

These morphisms are written in same way as functions. For example, we can have the category of groups; **Group**, where the morphisms are group homomorphisms. In this way $f : \mathbb{Z} \to \mathbb{Z}$, $f(x) = 2x$ belongs to Hom(**Group**), where we consider $\mathbb{Z}$ a group under addition. Similarly, we may define the category **Set**, where the objects are sets and the morphisms are just functions.

We refer during the paper to quite a few different categories, whose names are self-explanatory, but which we list here for completeness:

Group	is the category of Groups.
AbGroup	is the category of finite Abelian Groups.
$\mathbf{AbGroup^{(n)}}$	is the category of finite pointed Abelian Groups, where the point has order exactly n.
K-Algebra	is the category of K-Algebras.
Scheme	is the category of Schemes.

So what restrictions do we give to the morphisms of a category? Well, to start we need some binary operations on subsets of $\mathrm{Hom}(\mathscr{C})$. For every $A, B \in \mathrm{Ob}(\mathscr{C})$, we consider all morphisms between A and B and denote this collection by $\mathrm{Hom}(A, B)$. Now for every three objects A, B, C, we define a binary operation $\circ : \mathrm{Hom}(A, B) \times \mathrm{Hom}(B, C) \to \mathrm{Hom}(A, C)$, and we call this composition of morphisms. Indeed, in all the categories we discuss in this project, this operation will be identically the familiar composition of functions. We also require that this composition be associative and that the identity morphism id acts as the left and right identity on $\mathrm{Hom}(A, A)$.

A category $\mathscr{C}$ is locally small if for every C, D objects of $\mathrm{Ob}(\mathscr{C})$, $\mathrm{Hom}(C, D)$ is a proper set. All categories we consider in this paper are locally small.

Definition A.2. *Given a category $\mathscr{C}$, an initial object I is an object such that $|\mathrm{Hom}(I, C)| = 1$ for all $C \in \mathscr{C}$. A terminal object is an object T such that $|\mathrm{Hom}(C, T)| = 1$ for all $C \in \mathscr{C}$. If an object is both initial and terminal then we call it a zero object.*

Example A.3. *In the category **Set**, the empty set is the unique initial object, and the singleton sets are all terminal objects. In the category **Group**, the groups of one element are zero objects.*

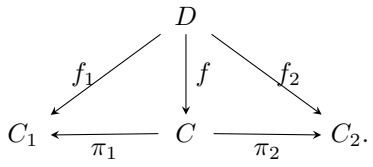
If a category has a zero object Z, then we define the zero morphism for any object C to be the unique $0_C : C \to Z$.

From a category $\mathscr{C}$ we can create a new category $\mathscr{C}^V$ known as the opposite category. The objects of this new category are the same, but the morphisms are now all reversed, so their origin is now their target and vice versa. Note that if $I \in \mathscr{C}$ is initial, then $I \in \mathscr{C}^V$ is terminal, and vice versa. We discuss these very briefly in section 2.3.

We will also require the notion of a product in a category. This generalises the cartesian product of sets and the direct product of groups.

Definition A.4. *Let $\mathscr{C}$ be a category and C_1, C_2, and D be objects of $\mathscr{C}$. We call C the product of C_1 and C_2 if there exist two morphisms $\pi_1 : C \to C_1$ and $\pi_2 : C \to C_2$ such that for any morphisms*

$f_1 : D \to C_1$ and $f_2 : D \to C_2$, there exists a unique morphism $f : D \to C$ such that the following diagram commutes



We call π_1 and π_2 projections and write $C = C_1 \times C_2$. We write $f = \{f_1, f_2\}$.

This can be extended in the obvious way to give a product for any family of objects indexed by some indexing set. However, it is worth noting that such products may not necessarily exist.

We can now discuss maps between different categories. We call these maps functors.

Definition A.5. *Given two categories $\mathscr{C}$ and $\mathscr{D}$, a covariant functor between them is a pair of maps $(F_{\mathrm{Ob}}, F_{\mathrm{Hom}})$, where F_{Ob} maps $\mathrm{Ob}(\mathscr{C})$ to $\mathrm{Ob}(\mathscr{D})$ by giving every A in $\mathrm{Ob}(\mathscr{C})$ an associated $F_{\mathrm{Ob}}(A)$ in $\mathrm{Ob}(\mathscr{D})$.*

Similarly, F_{Hom} maps $\mathrm{Hom}(\mathscr{C})$ to $\mathrm{Hom}(\mathscr{D})$ by giving every $f \in \mathrm{Hom}(\mathscr{C})$ an associated $F_{\mathrm{Hom}}(f)$ in $\mathrm{Hom}(\mathscr{D})$ such that $F_{\mathrm{Hom}}(id_A) = id_{F_{\mathrm{Ob}}(A)}$ and $F_{\mathrm{Hom}}(f \circ g) = F_{\mathrm{Hom}}(f) \circ F_{\mathrm{Hom}}(g)$. If $f : A \to B$, then $F_{\mathrm{Hom}} f : F_{\mathrm{Ob}} A \to F_{\mathrm{Ob}} B$.

So functors are maps between categories that map objects to objects, morphisms to morphisms, preserve identity morphisms and preserve composition of morphisms. For ease of notation we shall drop the clumsy subscript on the functor pair and write both of them as F. This will not cause confusion, as it will be obvious from the context whether we are acting on an object or a morphism. Similarly, we drop the "Ob" prefix when talking about objects in a category; we now simply refer to an object or a morphism in $\mathscr{C}$.

It is also worth noting that there also exist contravariant functors; functors which reverse the direction of morphisms. We give the definition here.

Definition A.6. *A contravariant functor F between two categories $\mathscr{C}$ and $\mathscr{D}$ associates for every object C in $\mathscr{C}$, an object $F(C)$ in $\mathscr{D}$. Also, for every morphism $f : C \to C'$ in $\mathscr{C}$, f maps to a new morphism $F(f) : F(C') \to F(C)$ such that $F(id_A) = id_{F(A)}$ and $F(f \circ g) = F(g) \circ F(f)$.*

Example A.7. *Consider the previous two categories we have mentioned;* **Group** *and* **Set**. *We can define a functor F from* **Group** *to* **Set** *by mapping a group to its underlying set, forgetting*

the group structure entirely. We do the same thing to the morphisms of **Group** *(which are group homomorphisms); we just map them to their set map counterparts. It is easy to see that this satisfies the definition of a functor. Because all this functor did was "forget" some structure, we call it a forgetful functor.*

A functor F from $\mathscr{C}$ to $\mathscr{D}$ is called faithful if F preserves all morphisms between any two objects of $\mathscr{C}$ under its mapping into $\mathscr{D}$. For example, in the example above, the functor is faithful, since group homomorphisms embed under F into the functions on sets.

The obvious question to ask now is when are maps between functors "nice"? Well, these are characterised by natural transformations.

Definition A.8. *Consider two categories $\mathscr{C}$ and $\mathscr{D}$, and two functors F and G between them. Then consider a collection of morphisms N between objects of $\mathscr{D}$ such that for every A in $\mathscr{C}$ we have n_A in N with $n_A : F(A) \to G(A)$. Also assume that for another morphism $f : A \to B$, we have that*

$$n_B \circ F(f) = G(f) \circ n_A. \tag{A.1}$$

Then we say that N is a natural transformation from F to G.

Note that this can also be expressed in a diagram.

$$\begin{array}{ccc} F(A) & \xrightarrow{n_A} & G(A) \\ {\scriptstyle F(f)}\downarrow & & \downarrow{\scriptstyle G(f)} \\ F(B) & \xrightarrow[n_B]{} & G(B) \end{array}$$

If equation A.1 holds, then the above diagram will commute (you can take either set of arrows from $F(A)$ to get to $G(B)$ and from the same a in $F(A)$ you will reach the same b in $G(B)$) and we call this a commutative diagram.

This allows us to define a functor category.

Definition A.9. *Given a locally-small category $\mathscr{C}$ and an arbitrary category $\mathscr{D}$, we define the functor category $\mathscr{D}^{\mathscr{C}}$, where the objects are functors $F : \mathscr{C} \to \mathscr{D}$, and the morphisms are natural transformations.*

The final thing we do in this appendix is state a very important category theoretical lemma, which will crop up in a few places in this project.

Theorem A.10 (Yoneda's Lemma)**.** *Let $\mathscr{C}$ be a locally-small category with X, Y objects of $\mathscr{C}$. Then we have:*

1. *If $\mathscr{F}$ is a contravariant functor from $\mathscr{C}$ to* **Set***, then natural transformations from* $\mathrm{Hom}(\_, X)$ *to $\mathscr{F}$ correspond to the elements of $\mathscr{F}(X)$.*

2. *Assume that* $\mathrm{Hom}(\_, X)$ *and* $\mathrm{Hom}(\_, Y)$*, functors from $\mathscr{C}$ to* **Set***, are isomorphic. Then we have $X \cong Y$. Generally, we have that maps from* $\mathrm{Hom}(\_, X)$ *and* $\mathrm{Hom}(\_, Y)$ *are the same as maps from X to Y. This gives us that a functor $\mathfrak{F} : \mathscr{C} \to \mathbf{Set}^{\mathscr{C}^V}$ that maps X to the functors mapping X' to* $\mathrm{Hom}(X', X)$ *defines an equivalence between $\mathscr{C}$ and some subcategory of functors.*

In particular, we have that given a natural transformation $\Phi : \mathrm{Hom}(\_, X) \to \mathscr{F}$, the corresponding element of $\mathscr{F}(X)$ is $\Phi_X(\mathrm{id}_X)$.

Appendix B

Lie Algebras

Definition B.1. *A Lie algebra $\mathfrak{L}$ over a field $\mathbb{F}$ is a vector space equipped with an alternating bilinear operation $[\cdot,\cdot]:\mathfrak{L}\times\mathfrak{L}\to\mathfrak{L}$ that satisfies the Jacobi identity*

$$[[uv],w]+[[wu],v]+[[vw],u]=0\ \forall u,v,w\in\mathfrak{L}. \tag{B.1}$$

A subalgebra $\mathfrak{M}$ is a vector subspace of $\mathfrak{L}$, which is closed under the lie bracket.

Throughout this paper we assume that all our Lie algebras are finite dimensional, unless otherwise stated.

Example B.2. *Any vector space with the trivial Lie bracket i.e. $[x,y]=0\ \forall x,y$, is a Lie algebra.*

Example B.3. *Take V a finite dimensional vector space over a field $\mathbb{F}$ and consider the $\mathbb{F}$-linear endomorphisms of V, denoted $\mathrm{End}(V)$. These can be represented as $n\times n$ matrices with entries in $\mathbb{F}$. This is an algebra under the standard addition and mutliplication of matrices. However, it can also be defined as a Lie algebra by giving $[x,y]=xy-yx\ \forall x,y\in\mathrm{End}(V)$.*

We now state and prove some basic facts about Lie algebras, most importantly the existence of a Cartan subalgebra for semisimple Lie algebras over $\mathbb{C}$. In fact from now on, we shall assume our base field is $\mathbb{C}$ unless otherwise stated.

Definition B.4. *An ideal $\mathfrak{I}$ of a Lie algebra $\mathfrak{L}$ is a subalgebra such that if $x\in\mathfrak{I}$, $y\in\mathfrak{L}$ then $[x,y]\in\mathfrak{I}$. Note that since we have $[x,y]=-[y,x]$ we have no distinction between left and right ideals.*

Definition B.5. *A Lie algebra homomorphism ρ from a Lie algebra $\mathfrak{L}$ to a Lie algebra $\mathfrak{M}$ is a linear map such that ρ commutes with the Lie bracket, i.e. $\rho([x,y]) = [\rho(x), \rho(y)]$. A derivation D is a linear map that acts as follows; $D[x,y] = [D(x), y] + [x, D(y)]$.*

Example B.6. *An important example in the study of Lie algebras is the adjoint homomorphism. For a Lie algebra $\mathfrak{L}$ and an element of the algebra x, we have*

$$\begin{aligned} \mathrm{ad}(x) : \mathfrak{L} &\to \mathfrak{L} \\ y &\mapsto [x,y]. \end{aligned} \tag{B.2}$$

This gives us a natural inclusion of $\mathfrak{L}$ into the algebra of linear endomorphisms of itself. It is also worth noting that $\mathrm{ad}(x)$ *is also a derivation since by the Jacobi identity we have*

$$\mathrm{ad}(x)[y,z] = [x,[y,z]] = -[y,[z,x]] - [z,[x,y]] = [[x,y],z] + [y,[x,z]] = [\mathrm{ad}(x)(y), z] + [y, \mathrm{ad}(x)(z)].$$

Lie algebra ideals and homomorphisms act analogously to those in other structures; direct analogues of the isomorphism theorems apply to Lie algebras, e.g. the kernels of homomorphisms are ideals and a Lie algebra modulo the kernel of a homomorphism acting on it is isomorphic to the image of that homomorphism.

Definition B.7. *We say that that a Lie algebra $\mathfrak{L}$ is abelian if $[x,y] = 0\ \forall x, y \in \mathfrak{L}$. It is semisimple if it has no non-zero solvable ideals. It is nilpotent if all but finitely many terms in its lower central series are zero; i.e. if the sequence $\mathfrak{L} > [\mathfrak{L},\mathfrak{L}] > [\mathfrak{L},[\mathfrak{L},\mathfrak{L}]] > [\mathfrak{L},[\mathfrak{L},[\mathfrak{L},\mathfrak{L}]]] > ...$ terminates.*

Definition B.8. *A linear transformation T of a vector space V is said to be diagonalisable (or semisimple) if there exists some basis B of V such that the matrix representing T with respect to B is diagonal. T is said to be nilpotent if $T^r = 0$ for some $r \in \mathbb{N}$.*

Definition B.9. *Consider a subalgebra $\mathfrak{C}$ of a Lie algebra $\mathfrak{L}$. If $\mathfrak{C}$ is abelian, every element of it is semisimple, and it is maximal with respect to these two properties, then we call $\mathfrak{C}$ a Cartan subalgebra of $\mathfrak{L}$.*

We are almost ready to prove that Cartan subalgebras exist, but we require one more theorem. This is Engel's theorem, and to prove this we will need a few preliminary lemmas.

Lemma B.10. *Let $\mathfrak{L}$ be a Lie subalgebra of* $\mathrm{gl}(V)$ *such that for all $x \in \mathfrak{L}$, x is a nilpotent linear*

transformation. Then we have that there exists some non-zero $v \in V$ such that $x(v) = 0$ for every $x \in \mathfrak{L}$.

Proof. We do this by induction on the dimension of $\mathfrak{L}$. If the dimension of $\mathfrak{L}$ is one, then $\mathfrak{L}$ is generated as a Lie algebra by some $z \in \mathrm{gl}(V)$. By hypothesis, we have that z is nilpotent and so $z^r = 0$, which gives that, in particular, for any non-zero $w \in V$, we have

$$z(z^{r-1}(w)) = 0. \tag{B.3}$$

If we pick r minimal so that $z^r = 0$, one of these w must have the property that $z^{r-1}(w) \neq 0$. Such a w and set $v = z^{r-1}(w)$, we see that for any scalar α, $\alpha z(v) = 0$, giving us the basis case for our induction. So, we may assume that the dimension of $\mathfrak{L}$ is greater than one.

So, consider a maximal Lie subalgebra of $\mathfrak{L}$, denoted by A. We now take the quotient $\mathfrak{L}' := \mathfrak{L}/A$, and define a map

$$\begin{aligned} &\rho : A \to \mathrm{gl}(\mathfrak{L}'), \text{ defined by} \\ &\rho(a)(x + A) = [a, x] + A. \end{aligned} \tag{B.4}$$

Now, we can show that this is well defined. Consider some $y = x + a'$ for some $a' \in A$. Then $[a, x] = [a, y - a'] = [a, y] - [a, a']$, which is equivalent to $[a, x]$ modulo A.

Now, ρ is also a Lie algebra homomorphism. To show this, we make the following standard argument, making use of the definition of the Lie bracket in $\mathrm{gl}(V)$, and the Jacobi identity.

$$\begin{aligned} [\rho(a), \rho(b)](x + A) &= (\rho(a)\rho(b) - \rho(b)\rho(a))(x + A) \\ &= \rho(a)([b, x] + A) - \rho(b)([a, x] + A) \\ &= [a, [b, x]] - [b, [a, x]] + A \\ &= [[a, b], x] + A \\ &= \rho[a, b](x + A). \end{aligned} \tag{B.5}$$

Since ρ is a homomorphism, we have $\rho(A)$ is a proper Lie subalgebra of $\mathrm{gl}(\mathfrak{L}')$. Note that $\rho(a)$ is induced from $\mathrm{ad}(a)$, which is nilpotent. Hence $\rho(a)$ is nilpotent. This gives us that we can use our inductive hypothesis on A.

Hence we have that there exists some $y + A \in \mathfrak{L}'$ such that $\rho(a)(y + A) = 0$ for all $a \in A$. Equivalently, $[y, a] \in A$ for all a. So, we can define $A \oplus \langle y \rangle$. The maximality of A gives us that

$A \oplus \langle y \rangle = \mathfrak{L}$, and since A is obviously an ideal of $A \oplus \langle y \rangle$, we have A is an ideal of $\mathfrak{L}$.

We can now apply our hypothesis to $A \subseteq \mathrm{gl}(V)$, so there is some non-zero $w \in V$ such that $a(w) = 0$ $\forall a \in A$. Hence the set

$$W = \{v \in V : a(v) = 0 \ \ \forall a \in A\} \tag{B.6}$$

is non-zero, and since all $a \in A$ are linear transformations, this gives us that W is a subspace of $\mathrm{gl}(V)$. Also, W is $\mathfrak{L}$-invariant, since for $x \in \mathfrak{L}$ we have $ax(w) = [a, x](w) - xa(w) = [a, x](w) \in W$, for since A is an ideal, $[a, x] \in A$.

This gives us that $y(W) \subseteq W$ for our nilpotent y. Hence we have a non-zero $v \in W$ such that $y(v) = 0$. Because $\mathfrak{L} = A \oplus \langle y \rangle$, we may say for any $x \in \mathfrak{L}$, $x = a + \alpha y$ for some $a \in A$ and scalar α. Hence, $x(v) = 0$, proving the claim. □

We can now make our next claim, which is central to the proof of Engel's theorem.

Proposition B.11. *Consider some vector space V. Then if $\mathfrak{L}$ is a Lie subalgebra of* $\mathrm{gl}(V)$ *such that every element of $\mathfrak{L}$ is nilpotent, there is a basis of V such that in this base, every $x \in \mathfrak{L}$ is strictly upper triangular.*

Proof. We prove this by induction on the dimension of V. Note that if the dimension of V is zero, this is trivially true, so we assume the dimension is greater than or equal to one. Now, from lemma B.10, we have that there exists some non-zero $v \in V$ such that $x(v) = 0$ for all $x \in \mathfrak{L}$. In a similar method to the previous lemma, we take $U := \langle v \rangle$ and let $V' = V/U$. A linear transformation of V, $x \in \mathfrak{L}$ induces a transformation x' on V' in the obvious way, $\rho(x) = x + U$. Note that this is obviously a homomorphism.

Hence, $\rho(L)$ is a subalgebra of $\mathrm{gl}(V')$, and every element of $\rho(L)$ is nilpotent. Since the dimension of V' is one less than V, we may apply our inductive hypothesis and find a basis of V' which makes $\rho(\mathfrak{L})$ strictly upper triangular. Let this basis be the set

$$\{v_1 + U, v_2 + U, ..., v_{n-1} + U\}.$$

But note that since $x(v) = 0$, the set $\{v, v_1, ..., v_{n-1}\}$ is a basis for V and $\mathfrak{L}$ is upper triangular with respect to it. □

We are now in a situation where we can prove Engel's theorem.

Theorem B.12 (Engel's Theorem)**.** *A Lie algebra $\mathfrak{L}$ is nilpotent if and only if for all $x \in \mathfrak{L}$, the map* $\mathrm{ad}(x)$ *is nilpotent.*

Proof. Assume $\mathfrak{L}$ is nilpotent, and terminates at the n^{th} term of the lower central series. Then $[x_0, [x_1, [..., x_{n-1}]]...] = 0$ for all x_i. By definition of the adjoint homomorphism therefore we have that $\mathrm{ad}(x_0)\mathrm{ad}(x_1)...\mathrm{ad}(x_{n-1}) = 0$ for all $x_0 \dots x_{n-1} \in \mathfrak{L}$. But this is true for all $x_i \in \mathfrak{L}$, in particular for $x_i = x$ and so we get, indeed, that $\mathrm{ad}(x)$ is nilpotent.

For the other direction, we define $\mathfrak{L}' := \mathrm{ad}\mathfrak{L}$. Our hypothesis says that every element of $\mathfrak{L}'$ is a nilpotent linear transformation, and so proposition B.11 gives us that there is a basis of $\mathfrak{L}$ where $\mathfrak{L}'$ is strictly upper triangular. Hence it follows that $\mathfrak{L}'$ is nilpotent, which gives that $\mathfrak{L}$ is nilpotent. □

Lemma B.13. *Let $\mathfrak{L}$ be a Lie algebra, and take some nilpotent $x \in \mathfrak{L}$. Then the map* $\mathrm{ad}x$ *is nilpotent.*

Proof. We use an identical argument as in the first part of the proof of Engel's theorem. □

Theorem B.14. *Let $\mathfrak{L}$ be a semisimple Lie algebra. Then it contains a Cartan subalgebra.*

Proof. Omitted. See [7], p. 95. □