Arbeitsgruppe Algebra
Mathematisches Institut
Justus-Liebig-Universität Gießen

Moufang Twin Trees and $\mathbb{Z}$-systems

A step towards the classification of Moufang twin trees

Dissertation

zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften (Dr. rer. nat.)

im Fachbereich
Mathematik und Informatik, Physik, Geographie
der Justus-Liebig-Universität Gießen

vorgelegt von

Herrn **Maximilian Parr**, M.Sc.

Betreuer: PD Dr. Matthias Grüninger

Gießen, 18. Oktober 2021

Acknowledgements

First of all, I want to express my sincere gratitude to my advisors Bernhard Mühlherr and Matthias Grüninger who gave me support and encouragement whenever there was tough progress or while I was struggling to overcome moments of doubt or quailing. I am especially thankful for Matthias' keen insight (not only in the topic of twin trees), his helpful suggestions and ideas to progress with my thesis, and his general kindness.

In addition, I want to give credit to my predecessor Moritz Scholz for his unpublished but nevertheless insightful notes '*Moufang-systems*' about the now-called $\mathbb{Z}$-systems which set a great basis for the work on these systems. I also want to thank Max Horn for initial help with and discussions of the essential paper '*Moufang twin trees of prime order*' by him, Matthias, and Bernhard.

I would like to thank my colleagues of the algebra research group, in particular Sebastian Bischof, Katharina Wendlandt and Lisa Wipper, for a good collaboration in matters of teaching and mutual support.

At last, I am grateful for my family and friends and their encouragement and support in various ways apart from research and work.

Abstract

Based on the ideas presented in the paper '*Moufang twin trees of prime order*' by Matthias Grüninger, Max Horn, and Bernhard Mühlherr ([9]), we generalize their main result [9, Theorem A], that the unipotent horocyclic group of a Moufang twin tree of prime order is nilpotent of class at most two, to a considerably larger class of Moufang twin trees.

Kurzfassung

Ausgehend von den im gemeinsamen Paper '*Moufang twin trees of prime order*' von Matthias Grüninger, Max Horn und Bernhard Mühlherr ([9]) präsentierten Ideen, verallgemeinern wir ihr Hauptresultat [9, Theorem A], dass die unipotente horozyklische Gruppe eines Moufang Zwillingsbaums von Primzahlordnung höchstens Nilpotenzstufe zwei besitzt, auf eine weitaus größere Klasse von Moufang Zwillingsbäumen.

Contents

List of Figures

Introduction

Jacques Tits gradually evolved the theory of buildings in the 1950s and 1960s in an attempt to give a systematic procedure for the geometric interpretation of, at first, semisimple complex Lie groups and, later on, semisimple algebraic groups over an arbitrary field. Inspired by ideas and results of Chevalley in the mids of 1955 which were itself based on works by Bruhat, the theory began fairly quickly to take form and matured until 1965. Tits gave a full account of his developed theory in his famous book '*Buildings of Spherical Type and Finite BN-Pairs*' from 1974 ([26]). Within, a building was treated as a simplicial complex Δ with a family of certain subcomplexes called apartments, whose top-dimensional simplices were named chambers, subject to only a few axioms ([26, 3.1]). This book contains the classification of (thick) irreducible spherical buildings of rank greater or equal to 3 which is one of the most important results in the theory of buildings. The fact that the local structure of a building determines the complete structure ([26, Theorem 4.1.2]) together with the notion of opposite chambers play a crucial role in the establishment of this classification.

A systematic study of non-spherical buildings from the classification point of view started in the 1980s. A most important contribution was provided by Tits in '*Immeubles de type affine*' in 1986 ([24]). In that paper a complete classification of all irreducible Euclidean buildings of rank at least 4 is achieved. The essential tool for the classification is the 'spherical building at infinity'. Subsequent work by Mark A. Ronan in '*A Construction of Buildings with no Rank 3 Residues of Spherical Type*' ([15]) and in the paper '*Building Buildings*' by Ronan and Tits ([17]) suggests that Tits' classification result in [24] is somehow optimal. Indeed, they show that there is no hope for a classification for other families of non-spherical buildings as purely combinatorial objects.

The irreducible spherical buildings of rank 2 are called generalized polygons, and they are too numerous to be classified as they involve buildings of type A_2 which are essentially projective planes. In order to sort out and classify those of algebraic origin, the Moufang condition was introduced

by Tits in his paper '*Endliche Spiegelungsgruppen, die als Weylgruppen auftreten*' from 1977 ([25, 3.3]). This condition is inspired by the notion of a root datum from Borel-Tits theory. It asks that the automorphism group of the building contains all so-called 'root-elations'. It turned out that all thick irreducible spherical buildings of rank $\geqslant 3$ satisfy this condition ([25, Satz 1]), and thus do their irreducible spherical residues of rank 2, also buildings, which are called Moufang (generalized) polygons. These polygons were classified by Tits and Richard M. Weiss in their book '*Moufang Polygons*' from 2002 ([31]) and led to a simplified proof of Tits' classification from 1974 as irreducible spherical higher rank buildings are, in a way, 'amalgamations' of these Moufang polygons.
Under the Moufang condition and in view of the classification of the Moufang polygons, the classification of affine buildings by Tits extends to the rank 3 case if the building at infinity is Moufang.

In the paper '*A Local Approach to Buildings*' from 1981 ([23]), Tits introduced a more 'modern' approach to buildings in form of chamber systems where he dropped the simplicial structure together with the apartments such that only the set of chambers remained. This set comes endowed with a Weyl-group-valued distance function d defined on pairs of chambers that satisfies a few axioms. The development of this viewpoint also took a few years until the late 1980s and was catalysed by the theory of twin buildings that Tits evolved together with Ronan in order to study and classify (the remaining) non-spherical buildings. An early exposition of this theory is given in '*Immeubles jumelés*' by Tits from 1988/89 ([30]). The theory of twin buildings itself was motivated by Tits' own paper '*Uniqueness and presentation of Kac-Moody groups over fields*' from 1987 ([27]) and it became apparent that twin buildings are the geometric framework associated to 'groups of Kac-Moody type' in the same way spherical buildings are associated to algebraic groups. The definition of a twin building as well as a final version of the definition of a building as a chamber system are given in Tits' work '*Twin buildings and groups of Kac-Moody type*' from 1992 ([28, Section 2.1 and 2.2]).
Roughly speaking, a twin building is a pair (Δ_+, Δ_-) of two buildings of the same type together with a codistance function d^* on pairs of chambers not contained in the same building subject to a few axioms. This codistance gives rise to an opposition relation between chambers in Δ_+ and chambers in Δ_- which has similar properties as the opposition relation in a single spherical building. So, even though the individual buildings of a twinning are in general not spherical, a twin building behaves like a spherical building in many regards. Hence, it is not surprising that it became apparent that twin buildings generalize spherical buildings in a natural way ([28, Proposition 1]). In this context, the question, if the

spherical results can be generalized to the twin case, arose naturally. Tits dealt with this question in the remainder of [28] and conjectured, under some restrictions, a possible classification of twin buildings by foundations, i.e. a union of certain spherical rank 2 buildings. Considerable progress on these conjectures was made in '*Local to global structure in twin buildings*' from 1995 by Bernhard Mühlherr and Ronan ([13]) and in '*Locally split and locally finite 2-spherical twin buildings*' by Mühlherr from 1999 ([12]).

In his 1992 paper, Tits also defined an analogue of the Moufang condition for twin buildings ([28, Section 4.3]) as well as a group theoretical datum called RGD-system (short for 'root groups data') to which a twin building is associated (via a twin BN-pair of the same type). There is (almost) a 1-to-1-correspondence between RGD-systems and Moufang twin buildings ([28, Proposition 7]).
Concerning the rank 1 case of the Moufang condition for twin buildings, he also defines Moufang sets which are essentially the split BN-pairs of rank one ([28, Section 4.4]) and turned out to be precisely the Moufang buildings of rank 1. They correspond to specific doubly transitive permutation groups. At that time, those were already classified in the finite case depending on the cardinality of the set those groups act upon. The case of even parity was dealt with in, for instance, '*Finite groups with a split BN-pair of rank 1. I*' by Christoph Hering, William M. Kantor, and Gary M. Seitz ([10]) in 1972. In the same year, the case of odd parity was solved by Ernest Shult in '*On a class of doubly transitive groups*' ([22]). But in general, (possibly thin) buildings of rank 1 are merely sets of cardinality at least 2 without any structure where every pair of chambers constitutes an apartment. Thus a complete classification is out of reach in the infinite case.

The Moufang condition for twin buildings makes especially sense for non-spherical twin buildings of rank 2 which are precisely the twin trees. In their paper '*Twin trees I*' from 1994 ([18]), Ronan and Tits gave a detailed introduction to them and to Moufang twin trees which are the objects of interest of this thesis.
A construction by Tits in '*Arbres jumelés*' from 1995/96 ([29, Section 5.4]) shows that there are uncountably many non-isomorphic 3-regular Moufang twin trees. A classification of all Moufang twin trees in analogue to the classification of Moufang (generalized) Polygons seems therefore not feasible.
Nevertheless, since Moufang twin trees are 'characterized' by an RGD-system, the classification problem of the geometric objects is equivalent to a classification of the group theoretical parameters of such systems. This is exactly the observation which Matthias Grüninger, Max Horn, and

Bernhard Mühlherr used in their paper '*Moufang twin trees of prime order*' from 2016 ([9]) in order to contribute to a potential classification of locally finite Moufang twin trees. In [30, Section 9], Tits gave a general construction to obtain Moufang twin trees which uses a specific subgroup of its automorphism group, the 'unipotent horocyclic group', an important invariant of the Moufang twin tree. The trio proved for any prime p that this subgroup of a $(p+1)$-regular Moufang twin tree is nilpotent of class at most 2.

Adapting their methods, the aim of this thesis is to extent this result to a wider class of Moufang twin trees to (hopefully) come closer to a possible classification of Moufang twin trees.

Overview

In Chapter 1 we only collect elementary definitions and notations concerning groups from different resources. We also state some facts about groups that are used repeatedly in this thesis.

After a few general graph theoretical definitions at the beginning of Chapter 2, a detailed introduction to trees together with the notion of twin trees, mostly based on Ronan and Tits paper ([18]), is given. It is accompanied by examples of a thin and a thick twin tree, the first is a single apartment and shall emphasize how apartments look in an arbitrary twin tree. The second is a rough round-up of the standard example of the twin tree for $\mathrm{GL}_2(\mathbb{K}[t, t^{-1}])$. We define ends of a twinning and construct pairs of half-apartments with the same end to obtain twin apartments consisting of a pair of apartments in which each vertex of one apartment is opposite exactly one vertex in the second one. From there on we introduce the notion of roots, define root groups and Moufang twin trees, and give a short fact about the commutator relations of these root groups.
The section thereafter deals with the connection between Moufang twin trees, RGD-systems and $\mathbb{Z}$-systems. We do not give a general definition of a root group datum, but one adjusted to our case of type $\widetilde{A}_1$. We recall the definition of a $\mathbb{Z}$-system of prime order from [9] and give a more general one. Of particular interest are the irreducible and nilpotent $\mathbb{Z}$-systems.
In the last section, we introduce Moufang sets as well as their construction following De Medts and Segev ([4]) and lay out the example of the projective Moufang set over a field. As a special Moufang set and by a result of Segev and Weiss ([20]), we obtain an irreducible action of a torus on the root groups of this Moufang set. This observation will justify our assumption of irreducibility on a $\mathbb{Z}$-system in subsequent parts of this thesis.

In Chapter 3, we introduce more terminology regarding a $\mathbb{Z}$-system $\Xi = (X, (X_k)_{k\in\mathbb{Z}}, \varsigma, T)$, especially the one of a shift-invariant subgroup which will be in focus most of the time, before we define the normal form of elements in X. It encodes the beginning, ending and width of an element. We then compute some commutator relations between the subgroups $\langle X_i \mid k \leqslant i \leqslant l\rangle$ and $\langle X_j \mid m \leqslant j \leqslant n\rangle$ for integers $k \leqslant l$ and $m \leqslant n$. These relations turn out to be little more charming if the subgroups X_k are assumed to be abelian.
In order to respect the action of T on subgroups of X, we define T-homomorphisms and adjust the isomorphism theorems accordingly, so that facts about (non) finitely generated groups carry over to the corresponding (non) finitely T-generated groups. We define projection maps that are T-homomorphisms and infer some useful consequences if their image contains enough elements. One of those being the fact that we can shorten words.
In the fourth section we distinguish two subsets of a shift-invariant subgroup Y, depending on the normal form of an element. These disjoint sets of 'even' and of 'odd words' have a certain connection to the T-index of Y in X. Under some assumptions, the T-index of Y in X is finite if and only if both subsets are non-empty.
The last part of Chapter 3 is devoted to generating sets of a shift-invariant T-subgroup of X whose $\mathbb{Z}$-system is irreducible. It turns out that shifts of words of minimal width generate such a subgroup. Depending on the facts we prove in later chapters, either one of the two presented descriptions will be more benefiting. We also show that the generating factors of an element can be sorted in a specific way.

The short Chapter 4 deals with a proof that X is T-locally nilpotent under the assumption that the corresponding $\mathbb{Z}$-system is nilpotent.

Closely following the ideas of [9] in Chapter 5 and using the T-index as well as the presented connection to the structure of Y, we step by step derive criteria for a shift-invariant T-subgroup Y to be 'one-sided' normal, i.e. for some $k \in \mathbb{Z}$ we have $Y \cap \langle X_l \mid l \leqslant k\rangle \trianglelefteq X$ or $Y \cap \langle X_l \mid l \geqslant k\rangle \trianglelefteq X$. One of these assumptions is the T-local nilpotency of X. A nice consequence under the same assumptions is the fact that Y is abelian.

In Chapter 6 we continue to follow the strategy of [9]. Alongside a few facts about T-generated, T-locally nilpotent, and nilpotent groups, we show that the quotient of X by some higher commutator group $\delta_k(X)$ is at least T-generated by k elements. In the next step we prove the existence of some k such that the corresponding quotient $X/\delta_k(X)$ is not finitely T-generated. The goal of this chapter is to prove that k equals 1 and that X has infinite T-abelianization.

Chapter 7 deals with the main result that states if $\Xi = (X, (X_k)_{k\in\mathbb{Z}}, \varsigma, T)$ is a nilpotent, irreducible $\mathbb{Z}$-system, then X is nilpotent of class at most 2. We give a few observations on commutators and shift-invariant T-subgroups, before we further adopt the methods from [9], including those involving G-modules, to prove our main theorem. An important step is the result that $[Y, X, X] = [Y, X]$ holds for a normal and shift-invariant T-subgroup Y of X of infinite T-index.

Chapter 1

Preliminary definitions and conventions

We recall some basic definitions, notations, and useful observations concerning groups that will occur in this work. We refer to the chapters 1, 5, and 12 in the book '*A Course in the Theory of Groups*' by Derek J. S. Robinson ([14]) for most of them.

Convention 1.1. We denote by $\mathbb{N}$ the set of natural numbers $\{0, 1, 2, \ldots\}$ and by $\mathbb{P}$ the set of primes. As usual, the set $\{\ldots, -2, -1, 0, 1, 2, \ldots\}$ of all integers is abbreviated by $\mathbb{Z}$.

The symbol G nearly always denotes a *group* with multiplicative binary operation and neutral element $1 = 1_G$. With G^* we label the set of all non-identity elements of G, i.e. $G \setminus \{1\}$. We often omit the symbol $\cdot$ in expressions like $g \cdot h$ and only write gh for $g, h \in G$.

In additive notation, i.e. in case of $\cdot = +$, we never leave out the sign and write $0 = 0_G$ for the identity.

Definition 1.2. As usual, the *order* of an element $g \in G$ is the smallest number $k \geqslant 1$ such that $g^k = 1_G$, or ∞ if this is never the case. We designate an element of order 2 as *involution*.
A group is called a *torsion group* if all of its elements have finite order, and *torsion-free* if all non-trivial elements have infinite order.
The *order* of G is its cardinality $|G|$. If the order of each element of G is a positive integer power of $p \in \mathbb{P}$, then G is called a *p-group*.

A group G is called *abelian* if $gh = hg$ for all $g, h \in G$. For $p \in \mathbb{P}$ an abelian group G is said to be an *elementary abelian* p-group if all elements in G^* have order p.

Convention 1.3. A *field* will always be commutative. If $\mathbb{K}$ is a finite field of order $q \in \mathbb{N}$, then we denote it by $\mathbb{F}_q$.

Definition 1.4. For any subset U of G, we define the *subgroup generated by* U as the intersection of all subgroups containing U and write $\langle U \rangle$. It is the set of all elements of the form $u_1^{\varepsilon_1} u_2^{\varepsilon_2} \cdot \ldots \cdot u_k^{\varepsilon_k}$ with $k \in \mathbb{N}$, $u_i \in U$, and $\varepsilon_i \in \{\pm 1\}$ for all $1 \leqslant i \leqslant k$, where the product equals 1 for $k = 0$ (see [14, 1.3.3]). If U is explicitly given as a subset of certain elements or as a set with conditions, we usually omit the set braces.

We denote the set $\{u^{-1} \mid u \in U\}$ of inverses of a subset U of G by U^{-1}. If U is a subgroup of G, then clearly $U^{-1} = U$. With this notion we can write $\langle U \rangle = \langle U \cup U^{-1} \rangle =: \langle U, U^{-1} \rangle$ for any $U \subseteq G$ to modify the description of an element $u_1^{\varepsilon_1} u_2^{\varepsilon_2} \cdot \ldots \cdot u_k^{\varepsilon_k}$ as above to an expression $u_1 u_2 \cdot \ldots \cdot u_k$ where $u_i \in U \cup U^{-1}$ for all $1 \leqslant i \leqslant k$. The equality holds as $U \subseteq \langle U, U^{-1} \rangle$ implies $\langle U \rangle \leq \langle U, U^{-1} \rangle$ and $U \cup U^{-1} \subseteq \langle U \rangle$ implies $\langle U, U^{-1} \rangle \leq \langle U \rangle$.

For a subgroup H of G and for $g \in G$ we call the set $gH := \{gh \mid h \in H\}$ a *left coset* and we define $G/H := \{gH \mid g \in G\}$. Note that $gH = g'H \Leftrightarrow g^{-1}g' \in H$. We define right cosets analogously.
More general, the *product* UV of two subsets U and V of G is the set $\{uv \mid u \in U, v \in V\}$.

We say that two subgroups H and K *permute*, if $HK = KH$. This is equivalent for HK to be a subgroup of G. Particularly, we have $HK = \langle H, K \rangle$ (cf. [14, 1.3.13]).

For $g, h \in G$ we set $g^h := h^{-1}gh$ for the *conjugation* of g by h, and define $U^g := \{u^g \mid u \in U\}$ for any non-empty subset $U \subseteq G$. The *normalizer* $N_G(U)$ of U in G is defined as the set $\{g \in G \mid U^g = U\}$ which is a subgroup of G. Observe that $H \leq G$ permutes with $U \leq G$ if $H \leq N_G(U)$.

Definition 1.5. Let $N \leq G$. If $N^g \subseteq N$ for all $g \in G$ or equivalently $gN = Ng$ for all $g \in G$, then we call N a *normal* subgroup of G and write $N \trianglelefteq G$ (see [14, 1.3.15]).

Note that a normal subgroup N permutes with each subgroup H of G, so that $HN = NH$ is always a subgroup of G.

For a subset $\varnothing \neq U \subseteq G$ we define the *normal closure* $\langle U \rangle^G$ of U in G by

$$\langle U \rangle^G := \left\langle U^G \right\rangle = \langle U^g \mid g \in G \rangle \trianglelefteq G$$

(cf. [9, Notation 7.1]). It equals $\bigcap_{U \subseteq N \trianglelefteq G} N$ and is the smallest normal subgroup in G containing U (see [14, Exercises 1.3 *18.]).

Definition 1.6. Let $N \trianglelefteq G$. Then G/N together with the binary operation

$$\circ : G/N \times G/N \to G/N : (gN) \circ (hN) := (g \cdot h)N$$

is a well-defined group, called the *quotient group* of N in G, with identity N and subgroups HN/N for $H \leq G$.

For two groups G and H we denote the *image* of G under a *homomorphism* $f : G \to H$ by $f(G) := \{f(g) \in H \mid g \in G\} \leq H$.
For $U \leq G$ and $V \leq f(G)$ we further define the image of U under f as $f(U) := \{f(u) \in H \mid u \in U\} \leq f(G)$ and the *pre-image* of V under f as $f^{-1}(V) := \{g \in G \mid f(g) \in V\} \leq G$.
The *kernel* of f is the specific pre-image $\ker(f) := f^{-1}(\{1_H\}) \trianglelefteq G$.

We write $G \cong H$ if there exists an isomorphism between G and H. As usual, the *group of all automorphisms* of a group G is denoted by $\mathrm{Aut}(G)$.

The following observation is used in Chapter 5:

Lemma 1.7. *Let $f : G \to H$ be a homomorphism between groups G and H. If $K \leq G$ such that $f(K) = f(G)$, then $K\ker(f) = G$.*

Proof. Since $K \leq G$ and $\ker(f) \trianglelefteq G$, we have $K\ker(f) \leq G$. Let $g \in G$. By assumption there is $k \in K$ such that $f(g) = f(k)$ or, equivalently, $f(k^{-1}g) = 1_H$, i.e. the element $k^{-1}g$ is contained in $\ker(f)$. We infer $g \in K\ker(f)$ and therefore $G \leq K\ker(f)$ which proves equality. $\square$

We recall the First Isomorphism Theorem and its two corollaries, the Second and Third Isomorphism Theorems. They will be adjusted to respect the action of a group T in Chapter 3.

Theorem 1.8. *([14, 1.4.3-1.4.5])*

(I1) *Let $f : G \to H$ be a homomorphism. Then*

$$F : G/\ker(f) \to f(G) : g\ker(f) \mapsto f(g)$$

is an isomorphism, i.e. $G/\ker(f) \cong f(G)$.

(I2) *Let $H \leq G$ and $N \trianglelefteq G$. Then $N \cap H \trianglelefteq H$ and*

$$\varphi : H \to HN/N : h \mapsto hN$$

is an epimorphism with kernel $N \cap H$, i.e. $H/(N \cap H) \cong HN/N$.

(I3) *Let $N, H \trianglelefteq G$ with $N \leq H$. Then $H/N \trianglelefteq G/N$ and*

$$\psi : G/N \to G/H : gN \mapsto gH$$

is an epimorphism with kernel H/N, i.e. $(G/N)/(H/N) \cong G/H$.

Proof. Since $\ker(f) \trianglelefteq G$, the quotient $G/\ker(f)$ is a well-defined group. We observe $f(gk) = f(g)$ for all $k \in \ker(f)$, so that F is also well-defined. The map F is a homomorphism by

$$\begin{aligned} F(g\ker(f) \circ h\ker(f)) &= F((gh)\ker(f)) = f(gh) \\ &= f(g)f(h) = F(g\ker(f))F(h\ker(f)) \end{aligned}$$

for all $g, h \in G$. If $f(g) \in f(G)$, then $F(g\ker(f)) = f(g)$; thus F is surjective. It is injective, since $f(g) = F(g\ker(f)) = F(h\ker(f)) = f(h)$ implies $g^{-1}h \in \ker(f)$ and $g\ker(f) = h\ker(f)$. This proves (I1).

For (I2) note that H normalizes N, so that HN is a subgroup of G with normal subgroup N, as well as itself, so that $N \cap H \trianglelefteq H$. Both quotients HN/N and $H/(N \cap H)$ are therefore well-defined. We have

$$\varphi(gh) = (gh)N = gN \circ hN = \varphi(g) \circ \varphi(h)$$

for all $g, h \in H$ and $\varphi(h) = hN = (hn)N$ for all $(hn)N \in HN/N$. Moreover, we compute

$$\ker(\varphi) = \{h \in H \mid hN = N\} = \{h \in H \mid h \in N\} = N \cap H.$$

Thus φ is an epimorphism with kernel $N \cap H$. The last part of the second statement now follows from the first assertion.

Regarding the third assertion, observe that H/N, G/N and G/H are well-defined by our assumptions. Since $H \trianglelefteq G$, we have $hN^{gN} = (h^g)N \in H/N$ for all $h \in H$ and $g \in G$ which yields $H/N \trianglelefteq G/N$. Note that $gN = g'N$ implies $g^{-1}g' \in N \subseteq H$ and $gH = g'H$; hence ψ is a well-defined map. It is clearly surjective. Similar to the corresponding part of (I1), we see that it is a homomorphism. As $N \leq H$ and thus

$$\ker(\psi) = \{gN \in G/N \mid gH = H\} = \{gN \in G/N \mid g \in H\} = H/N,$$

the remaining claim follows by (I1), again. □

As a direct consequence of (I3), the *canonical projection* $\rho : g \mapsto gN$ for a normal subgroup N of G is an epimorphism from $G \cong G/\{1\}$ to G/N with kernel $N \cong N/\{1\}$.

Definition 1.9. A *(left) action* of a group G on a non-empty set X is a map

$$\lambda : G \times X \to X : (g, x) \mapsto g.x := \lambda(g, x)$$

such that for all $g, h \in G$ and $x \in X$ we have

$$1.x = x \quad \text{and} \quad (gh).x = g.(h.x).$$

A right action is defined in an analogous way.

Regarding an action of G on any element x or any non-empty subset U of X, we will use the term *G-orbit* of x for the orbit set $G(x) := \{g.x \mid g \in G\}$ and G-orbit of U for $G(U) := \{g.u \mid g \in G, u \in U\}$, respectively.

The set U is called *G-invariant* (and the action of G on U invariant) if $G(U) \subseteq U$. In the special case of U being a G-invariant group we call U a *G-group*. If U is a G-invariant subgroup of a (possibly non G-invariant) group H, we write $U \leq_G H$ and say that U is a *G-subgroup* of H.

The subgroup

$$\mathrm{St}_G(U) := \{g \in G \mid g.u = u \text{ for all } u \in U\}$$

consists of all elements in G that *fix* the set U point-wise. It is called the (point-wise) *stabilizer* of U in G. If U is a singleton $\{u\}$, we just write $\mathrm{St}_G(u)$.

An action is said to be *transitive* on X if $G(x) = X$ for one (and hence any) $x \in X$, *free* (or *semiregular*) if $\mathrm{St}_G(x) = \{1\}$ for all $x \in X$, and *regular* (or *sharply transitive*) if it is transitive and free. An action is *doubly transitive* if it is transitive on pairs (x, y) with $x \neq y \in X$.

If a group G acts regularly on a set X, then $|G| = |X|$ (cf. [14, 1.6.1(iii)]).

Beginning in Chapter 3, we will frequently use a left action of a group T on a group G by automorphisms.

The definition of commutators (of certain weight) and groups generated by those elements are fundamental to this work:

Definition 1.10. The *commutator* of two elements $g, h \in G$ is the product

$$[g, h] := g^{-1}g^h = g^{-1}h^{-1}gh.$$

If $g_1, g_2, \ldots, g_n \in G$ (for some $n \geqslant 3$), we define the sometimes so-called *simple commutator of weight n* by

$$[g_1, g_2, \ldots, g_n] := [\ldots[[g_1, g_2], g_3], \ldots, g_n].$$

For two non-empty subsets $U, V \subseteq G$ we denote by $[U, V]$ the *group generated by the set of all commutators* $[u, v]$ with $u \in U$ and $v \in V$. Observe that $[U, V] = [V, U]$, since $[u, v] = [v, u]^{-1}$ for all $u \in U$ and $v \in V$ (see [14, 5.1.5(i)]). The group $\delta(G) := [G, G]$ is called the *derived subgroup* and we will call $G_{ab} := G/\delta(G)$ the *abelianization* of G. Note that this group is the largest abelian quotient of G (cf. [14, p. 124]).
We define $[U_1, U_2, \ldots, U_n]$ for $n \geqslant 3$ and non-empty subsets U_1, $U_2, \ldots, U_n$ of G in the obvious way.

Lemma 1.11. *Let $f : G \to H$ be a homomorphism between a group G and an abelian group H. Then $\delta(G) \leq \ker(f)$.*

Proof. Let $g, h \in G$. Since $\ker(f)$ is a subgroup of G, it suffices to show that the generators $[g, h]$ of $\delta(G)$ are contained in $\ker(f)$. We compute

$$f([g,h]) = f(g^{-1}h^{-1}gh) = f(g)^{-1}f(h)^{-1}f(g)f(h) = [f(g), f(h)] = 1_H,$$

and infer $[g, h] \in \ker(f)$. □

The following general observation is frequently utilized to compute the commutator of products or the product of commutators:

Lemma 1.12. *([14, 5.1.5(ii)]) Let G be a group and $x, y, z \in G$. Then*

$$[xy, z] = [x, z]^y[y, z] \quad \textit{and} \quad [x, yz] = [x, z][x, y]^z.$$

Proof. By inserting $1 = zy(zy)^{-1}$, we get

$$[xy, z] = y^{-1}x^{-1}z^{-1}xyz = y^{-1}x^{-1}z^{-1}xzyy^{-1}z^{-1}yz = [x, z]^y[y, z].$$

For the second equality, we obtain

$$[x, yz] = x^{-1}z^{-1}y^{-1}xyz = x^{-1}z^{-1}xzz^{-1}x^{-1}y^{-1}xyz = [x, z][x, y]^z$$

by simply inserting $1 = xz(xz)^{-1}$. □

The preceding lemma directly implies

Corollary 1.13. *Let G be a group and $H, K, N \leq G$ such that H and K normalize N. Then $[H, N] \leq N$, $[H, K]$ normalizes N, and $[HN, KN] \leq [H, K]N$.*

Proof. Let $h \in H$, $k \in K$, and $n, n' \in N$. As H normalizes N, we have $(n^{-1})^h \in N$ and thus $[h, n] = (n^{-1})^h n \in N$, so that $[H, N] \leq N$.

For the second statement we just observe that $n^{[h,k]} = (((n^{h^{-1}})^{k^{-1}})^h)^k \in N$ by our premise. In particular, $[H, K]N$ is a subgroup of G.

The last part follows by applying Lemma 1.12 twice. Indeed, we compute

$$\begin{aligned} [hn, kn'] &= [hn, n'][hn, k]^{n'} = [h, n']^n[n, n']([h, k]^n[n, k])^{n'} \\ &= [h, n']^n[n, n'][h, k]^{nn'}[n, k]^{n'}, \end{aligned}$$

where the first, second and fourth (conjugated) commutators are contained in N by the first part. Hence, we infer

$$[hn, kn'] = m[h, k]m' = [h, k]m''m' \in [H, K]N$$

for some $m, m', m'' \in N$ which implies $[HN, KN] \leq [H, K]N$. □

We have another direct consequence (see, for example, the book '*Finite Groups*' by Daniel Gorenstein ([8, Chapter 2, Theorem 2.1(iii)]) that holds for arbitrary groups).

Corollary 1.14. *Let G be a group and $\varnothing \neq U, V \subseteq G$. Then $[U, V] \trianglelefteq \langle U, V\rangle$. In particular, we have $[U, G] \trianglelefteq G$.*

Proof. By changing the statement of the lemma above into the identity

$$[x, y]^z = [x, z]^{-1}[x, yz],$$

whereby $x \in U$ and $y, z \in V$, we observe $[x, y]^z \in [U, V]$ and thus $V \subseteq N_G([U, V])$. We get $U \subseteq N_G([U, V])$ by interchanging U and V; hence $\langle U, V\rangle \subseteq N_G([U, V])$.
Since $[U, V] \leq [\langle U, V\rangle, \langle U, V\rangle] \leq \langle U, V\rangle$, we infer $[U, V] \trianglelefteq \langle U, V\rangle$. □

We denote the *center* of a group G by

$$Z(G) := \{g \in G \mid \forall h \in G : [g, h] = 1\} \trianglelefteq G.$$

More general, for any non-empty subset U of G we define the *centralizer* of U in G to be the subgroup

$$C_G(U) := \{g \in G \mid \forall u \in U : [g, u] = 1\}.$$

The definitions of some series hereinafter are especially important for Chapters 4 and 6. Hence, we collect them here:

Definition 1.15. Let $n \in \mathbb{N}$. A group G is called *solvable* (or *soluble*) if it has an *abelian series*

$$G = G_n \trianglerighteq G_{n-1} \trianglerighteq \ldots \trianglerighteq G_1 \trianglerighteq G_0 = \{1\},$$

i.e. a finite series of subgroups $G_i \leq G$ such that G_{i+1}/G_i is abelian for all $0 \leqslant i \leqslant n-1$. The smallest number n with this property is called the *derived length* of G. We recursively define

$$\delta_0(G) := G \quad \text{and} \quad \delta_d(G) := [\delta_{d-1}(G), \delta_{d-1}(G)] \quad \text{for} \quad d \geqslant 1.$$

In particular, we have $\delta_1(G) = \delta(G)$. The descending series

$$\delta_0(G) \trianglerighteq \delta_1(G) \trianglerighteq \delta_2(G) \trianglerighteq \ldots$$

is called the *derived series*.

A group G is solvable if and only if $\delta_d(G) = \{1\}$ for some $d \in \mathbb{N}$. If G is solvable, then d is actually the derived length of G (see [14, 5.1.8]).

The first part of the following lemma can be found in the book '*The Theory of Nilpotent Groups*' by Anthony E. Clement, Stephen Majewicz, and Marcos Zyman ([2, Proposition 1.2]).

Lemma 1.16. *Let $f : G \to H$ be a homomorphism between two groups G and H.*

(i) If $A, B \leq G$, then we have $f([A, B]) = [f(A), f(B)] \leq \delta(H)$ and, in particular, $f(\delta(G)) = \delta(f(G))$.

(ii) If f is surjective, then $f(\delta_d(G)) = \delta_d(H)$ for all $d \geqslant 0$.

Proof. Since f is a homomorphism, a product $\prod_{i=1}^{n}[a_i, b_i]^{\varepsilon_i}$ of commutators, where $a_i \in A, b_i \in B$, and $\varepsilon_i \in \{\pm 1\}$ for $1 \leqslant i \leqslant n$ for some $n \in \mathbb{N}$, is mapped to the product $\prod_{i=1}^{n}[f(a_i), f(b_i)]^{\varepsilon_i}$, i.e. $f([A, B]) \subseteq [f(A), f(B)]$. The other inclusion follows by the same argument as $h \in [f(A), f(B)]$ is a product $\prod_{j=1}^{m}[f(a'_j), f(b'_j)]^{\varepsilon'_j}$ with $f(a'_j) \in f(A), f(b'_j) \in f(B)$, and $\varepsilon'_j \in \{\pm 1\}$ for all $1 \leqslant j \leqslant m$ for some $m \in \mathbb{N}$.
Clearly, we have $[f(A), f(B)] \leq \delta(H)$ by $f(A), f(B) \leq H$.

Now, let f be surjective. We prove the second assertion via induction on d. There is nothing to show for $d = 0$ as $f(G) = H$. For $d = 1$ we have $f(\delta(G)) = \delta(f(G)) = \delta(H)$ by (i).
Assume that $f(\delta_d(G)) = \delta_d(H)$ holds for some $d \geqslant 2$. Then

$$\begin{aligned} f(\delta_{d+1}(G)) &= f([\delta_d(G), \delta_d(G)]) = [f(\delta_d(G)), f(\delta_d(G))] \\ &= [\delta_d(H), \delta_d(H)] = \delta_{d+1}(H), \end{aligned}$$

what proves the claim. □

Definition 1.17. Let $n \in \mathbb{N}$. A group G is called *nilpotent* if it has a *central series*

$$G = G_n \trianglerighteq G_{n-1} \trianglerighteq \ldots \trianglerighteq G_1 \trianglerighteq G_0 = \{1\},$$

i.e. a finite series of subgroups $G_i \trianglelefteq G$ such that $G_{i+1}/G_i \leq Z(G/G_i)$ for all $0 \leqslant i \leqslant n-1$. The smallest number n with this property is called the *nilpotent class* of G. We use the notation $nc(G) = n$.

For all $0 \leqslant i \leqslant n-1$ note that

$$G_{i+1}/G_i \leq Z(G/G_i) \;\Leftrightarrow\; [G_{i+1}/G_i, G/G_i] = \{G_i\} \;\Leftrightarrow\; [G_{i+1}, G] \leq G_i$$

holds. In particular, the last relation implies $[G_{i+1}, G] \leq G_{i+1}$ and $G_{i+1} \trianglelefteq G$. So we could replace the defining property for nilpotency by the equivalent statement on the right-hand side.

The *lower central series* of a group G is the descending series

$$\gamma_0(G) \trianglerighteq \gamma_1(G) \trianglerighteq \gamma_2(G) \trianglerighteq \dots,$$

where the groups are defined as follows:

$$\gamma_0(G) := G \quad \text{and} \quad \gamma_d(G) := [\gamma_{d-1}(G), G] \quad \text{for} \quad d \geqslant 1.$$

The *upper central series* of a group G is the ascending series

$$\zeta_0(G) \trianglelefteq \zeta_1(G) \trianglelefteq \zeta_2(G) \trianglelefteq \dots,$$

where $\zeta_0(G) := \{1\}$ and $\zeta_i(G)$ is defined to be the pre-image of the center $Z\left(G/\zeta_{i-1}(G)\right)$ under the canonical projection $G \to G/\zeta_{i-1}(G)$ for all $i \geqslant 1$, i.e.

$$\zeta_i(G) = \{x \in G \mid \forall y \in G : [x, y] \in \zeta_{i-1}(G)\}.$$

In particular, $\zeta_1(G) = Z(G)$.

We can directly formulate another useful statement:

Lemma 1.18. *Let G and H be groups and $d \in \mathbb{N}$.*

(i) If $K \leq G$ and $N \trianglelefteq G$, then we have $\gamma_d(K) \leq \gamma_d(KN) \leq \gamma_d(K)N$.

(ii) If $f : G \to H$ is an epimorphism, then $f(\gamma_d(G)) = \gamma_d(H)$.

Proof. We use induction on d for both parts and start with assertion (i). The base case follows by $\gamma_0(K) = K \leq KN$ and $\gamma_0(KN) = KN = \gamma_0(K)N$. For $d = 1$ we observe $\gamma_1(K) = \delta(K) \leq \delta(KN) = \gamma_1(KN)$ as well as $\delta(KN) \leq \delta(K)N = \gamma_1(K)N$ by Corollary 1.13.
Let $d \geqslant 2$. Then the induction hypothesis for the first inequality implies

$$\gamma_{d+1}(K) = [\gamma_d(K), K] \leq [\gamma_d(KN), K] \leq [\gamma_d(KN), KN] = \gamma_{d+1}(KN),$$

whereas the second inequality yields

$$\begin{aligned} \gamma_{d+1}(KN) = [\gamma_d(KN), KN] &\leq [\gamma_d(K)N, KN] \\ &\leq [\gamma_d(K), K]N = \gamma_{d+1}(K)N, \end{aligned}$$

by using Corollary 1.13 again.

Now we turn to the second part. The base step equals the one in Lemma 1.16 (ii), so we only need to perform the induction step for $d \geqslant 2$. We compute

$$f(\gamma_{d+1}(G)) = f([\gamma_d(G), G]) = [f(\gamma_d(G)), f(G)] = [\gamma_d(H), H] = \gamma_{d+1}(H)$$

by Lemma 1.16 (i), our induction hypothesis, and the first part of this lemma. $\square$

A group G is nilpotent if and only if there exists an $s \in \mathbb{N}$ such that either $\gamma_s(G) = \{1\}$ or, also equivalent, there exists an $s \in \mathbb{N}$ such that $\zeta_s(G) = G$ (see, for instance, [14, 5.1.9]). Moreover, the nilpotent class of a nilpotent group G coincides with the length of lower and upper central series.

We say that a series *terminates* if it is descending and ends in $\{1\}$ or is ascending and ends in G.

Note that a nilpotent group G is also solvable, since $G_{i+1}/G_i \leq Z(G/G_i)$ for all $0 \leqslant i \leqslant n-1$ implies that G_{i+1}/G_i is abelian for all $0 \leqslant i \leqslant n-1$.

Each group $\delta_i(G)$, $\gamma_i(G)$ and $\zeta_i(G)$, $i \in \mathbb{N}$, is actually *characteristic* in G, i.e. invariant under all automorphisms of G (see [14, p. 28 and p. 125]).

Let P be a property a group can have, i.e. finite, abelian, nilpotent, a p-group, ... A group G is called *locally* P if every finitely generated subgroup of G has property P.

We will use the following definition in the last chapter to prove our main result Theorem 7.7:

Definition 1.19. A (right) G-*module* is an abelian group M together with a (right) action of G on M such that

$$(ab).g = (a.g)(b.g)$$

for all $g \in G$ and $a, b \in M$. Furthermore (see [9, Remark 9.3]), we define the *module commutator*

$$[m, g] := m^{-1}(m.g)$$

for $g \in G$ and $m \in M$ as well as the group

$$[M, G] := \langle [m, g] \mid m \in M, g \in G \rangle \leq M.$$

Even though M is abelian, we use a multiplicative notion.

Chapter 2

From Moufang twin trees to $\mathbb{Z}$-systems

In the first part of this chapter we pave the way for Definition 2.33 of a $\mathbb{Z}$-system. Thus, we draw the connection between Moufang twin trees as rank 2 cases of non-spherical twin buildings having the Moufang property and these systems of groups satisfying a handful of axioms. Our definition of a $\mathbb{Z}$-system is adapted to Definition 3.2 of a $\mathbb{Z}$-system of prime order appearing in the paper '*Moufang twin trees of prime order*' by Matthias Grüninger, Max Horn, and Bernhard Mühlherr from 2016 ([9]). Their concept of a $\mathbb{Z}$-system can be viewed as weaker version of the concept of an RGD-system of type $\widetilde{A}_1$. The latter itself is a tool for constructing Moufang twin trees.

We will state all definitions in the middle of this chapter. For now, let us start by an introduction to Moufang twin trees.

2.1 Trees and twin trees

After we have set the stage for groups in the first chapter, we now tend to the geometrical objects of interest: twin trees.

The main reference for the basic theory of twin trees is the paper '*Twin Trees I*' by Mark A. Ronan and Jacques Tits from 1994 ([18]). In this paper they give a detailed overview of the 2-rank case of non-spherical twin buildings freed from some technical complications of the general definition of twin buildings.
A general description of a twin building can be found, for example, in the comprehensive book '*Buildings*' by Peter Abramenko and Kenneth

S. Brown (see [1, Definition 5.133]) or in Ronan's book '*Lectures on Buildings*' (cf. [16, Chapter 11, Section 1]).

Besides the paper of Ronan and Tits mentioned above there are also the course notes '*Arbres Jumelés*' by Jacques Tits from 1995/96 ([29]) which contain more advanced results on twin trees. The content of this section is built upon both references.

We start by recalling the combinatorial definition of a tree as a graph (see the book '*Graphentheorie*' by Reinhard Diestel ([6, pp. 2-15])):

Definition 2.1. Let V be a non-empty set and E be a subset of the set of all subsets of V whose elements are sets containing two elements of V. Then $\Gamma = (V, E)$ is called a *graph* with V its set of *vertices* and E its set of *edges*. We may abbreviate the notion of an edge $e = \{v, v'\}$ by writing $e = vv'$ (or $e = v'v$). Vertices v and v' with $vv' \in E$ are called *adjacent*.
The *valency* of a vertex v is the (possibly infinite) number of vertices adjacent to v. A graph whose vertices all have finite valency is called *locally finite*. If further all valencies are equal, say to $n \in \mathbb{N}$, then the graph is called *n-regular*.

Let $k \in \mathbb{N}$ and $v, v' \in V$. A *walk* P of *length* k from v to v' (in Γ) is a sequence

$$P = v_0 e_0 v_1 e_1 v_2 \dots v_{k-1} e_{k-1} v_k$$

with $v_i \in V$ for $0 \leqslant i \leqslant k$, $e_i = \{v_i, v_{i+1}\} \in E$ for all $0 \leqslant i \leqslant k-1$, and $v_0 = v$ as well as $v_k = v'$. We then call v the *initial* and v' the *terminal* vertex of P or both simply the *endpoints* of P. If the vertices (and therefore the edges) are pairwise distinct, we call P a *path* from v to v'. We may omit the edges (or the vertices) in a given sequence in this case.
A graph is *connected* if there is a path for each pair of distinct vertices connecting them.
We call a walk *closed* if its endpoints coincide. A closed walk $v_0 e_0 \dots e_{k-1} v_0$ of length $k \geqslant 3$ is called a *cycle* if $v_0 e_0 v_1 \dots v_{k-1}$ is a path.

Infinite walks and *infinite paths* are defined as infinite sequences of vertices and edges with similar properties as their finite analogues. They either have exactly one or zero endpoints.

Definition 2.2. An *isomorphism* from a graph $\Gamma = (V, E)$ onto a graph $\Gamma' = (V', E')$ is a bijection $\varphi : V \to V'$ such that $vw \in E \Leftrightarrow \varphi(v)\varphi(w) \in E'$. In particular, this induces a bijection between the set of edges. If $\Gamma = \Gamma'$, then an isomorphism is called an *automorphism*.
The set of automorphism with composition as operation yields the group $\mathrm{Aut}(\Gamma)$ of automorphism of Γ.

Definition 2.3. A *tree* is a connected graph without cycles.

Note that this definition of a graph does neither allow multiple edges between vertices nor loops on vertices unlike a more general notation appearing, for example, in Jean-Pierre Serre's book '*Trees*' ([21, Chapter I.2, Definition 1]).

Remark 2.4. It is readily verified that a graph $\Gamma = (V, E)$ is a tree if and only if for each pair of distinct vertices there exists a unique path between them (cf. [6, Theorem 0.5.1]).
Therefore, given two vertices v and v', the length of a path from v to v' is a uniquely determined, non-negative integer $l_v^{v'}$ and gives rise to the *distance function*

$$d_\Gamma : V \times V \to \mathbb{N} : (v, v') \mapsto l_v^{v'},$$

i.e. to a function satisfying $d_\Gamma(v, v') = 0 \Leftrightarrow v = v'$, $d_\Gamma(v, v') = d_\Gamma(v', v)$ for all $v, v' \in V$, and $d_\Gamma(v, v'') \leqslant d_\Gamma(v, v') + d_\Gamma(v', v'')$ for all $v, v', v'' \in V$.

On the contrary, in [29, 1.1] by Tits or in the first chapter of the joint paper [18] with Ronan, a tree is defined as follows:

Definition 2.5. A *tree* is a pair $\Delta = (T, d)$ consisting of a non-empty set T, whose elements will be called vertices, together with a symmetric function $d : T \times T \to \mathbb{N}$ (in the sense that $d(x, y) = d(y, x)$ for all $x, y \in T$), called *distance*, satisfying the following properties, where $x, y \in T$ with $d(x, y) = n$:

(T0) $d(x, y) = 0 \Leftrightarrow x = y$.

(T1) If $z \in T$ with $d(y, z) = 1$, then $d(x, z) = n \pm 1$.

(T2) If $n > 0$, then there exists a unique $z \in T$ with $d(y, z) = 1$ and $d(x, z) = n - 1$.

It is easy to verify that the distance function d_Γ in the remark above satisfies these conditions. Hence a tree defined as a connected graph with no cycles is a tree in this sense.

Conversely, if we join each pair (x, y) of vertices of Δ with $d(x, y) = 1$, those vertices are called adjacent or neighbours, then we obtain a connected graph with vertex set T and one edge $\{x, y\}$ for each pair of adjacent vertices. This graph has no cycles (we may assume that it has at least three vertices), since, by using (T1) as well as (T2) on adjacent vertices x and y, the distance $d(y, z)$ for vertices $z \neq x$ adjacent to y increases by 1. Inductively, the distance increases successively if we take successively neighbours of neighbours away from x. So, no such vertex will be adjacent to x again and no path can become a cycle by adjoining an edge having x as one vertex.

Thus both definitions of a tree are indeed equivalent. Since the definition using the distance is more suited by having general buildings in mind, we will use it onwards.

In view of this an automorphism of a tree Δ is a bijection on the vertex set T which preserves distances.

We are only interested in infinite trees, i.e. in trees whose vertices have valency at least 2. Therefore we slightly alter the second property above to secure that each vertex has at least two neighbours (see [29, 1.1]):

(T1') If $z \in T$ with $d(y,z) = 1$, then $d(x,z) = n \pm 1$; there is $z \in T$ with $d(y,z) = 1$ such that $d(x,z) = n+1$.

Convention 2.6. For the remainder of this work, we will mean an infinite tree subject to the axioms (T0), (T1') and (T2) when speaking of a tree.

Remark 2.7. Trees are *chamber complexes* of rank 2, i.e. finite-dimensional simplicial complexes whose maximal simplices, called *chambers*, are of the same dimension and each two can be connected by a sequence of chambers in which consecutive ones are distinct and share a common face of codimension 1 (cf. [1, Appendix A.1] for more details).

From the simplicial viewpoint, a building is a simplicial complex together with a family of subcomplexes, called apartments, that satisfy certain axioms. It can be shown that infinite trees are exactly the rank-2 non-spherical buildings of type $\widetilde{A}_1$ (resp. $I_2(\infty)$) (see [1, Definition 4.1 and Proposition 4.44] and the discussion in between). The apartments are then *Coxeter complexes* associated to the *infinite dihedral group*

$$D_\infty = \langle \{s,t\} \mid s^2 = t^2 = 1 \rangle,$$

a *Coxeter group*. (The right-hand side is a presentation with $\{s,t\}$ the set of generators and defining relations $s^2 = t^2 = 1$.) It is the affine Weyl group of type $\widetilde{A}_1$ that appears in the context of Lie theory (see [1, Section 0.1]).

Following [29] and [18], we now define a twin tree (note that there is a sign-typo at the end of (AJ2) in Tit's course notes):

Definition 2.8. Let $\Delta_+ = (T_+, d_+)$ and $\Delta_- = (T_-, d_-)$ be two trees. A *codistance* between Δ_+ and Δ_- is a symmetric function

$$d^* : (T_+ \times T_-) \cup (T_- \times T_+) \to \mathbb{N},$$

such that for each pair $(x,y) \in (T_+ \times T_-) \cup (T_- \times T_+)$ the following conditions are satisfied, where $s \in \{+,-\}$:

(TT1) If $z \in T_+ \cup T_-$ with $d_s(y,z) = 1$, then $d^*(x,z) = d^*(x,y) \pm 1$.

(TT2) If $d^*(x,y) > 0$, then there exists a unique vertex z in the same tree Δ_s that contains y such that $d_s(y,z) = 1$ and $d^*(x,z) = d^*(x,y) + 1$.

The triple $(\Delta_+, \Delta_-, d^*)$ is called a *twin pair of trees* (or simply a *twin tree*) or a *twinning* of Δ_+ and Δ_-. The individual trees are called the *halves* of the twin tree.

Given two vertices x and y contained in different halves of the twin tree, we say that x and y are *opposite* if $d^*(x,y) = 0$. Note that each vertex has at least one opposite vertex by (T1') (or by Corollary 2.15 below).

If the valency of each vertex of a (twin) tree is at least 3 (resp. equals 2), then the (twin) tree is called *thick* (resp. *thin*).

Definition 2.9. An *isomorphism* between two twin trees $(\Delta_+, \Delta_-, d^*)$ and $(\Delta'_+, \Delta'_-, \delta^*)$ is a pair $\varphi = (\varphi_+, \varphi_-)$ of isomorphisms φ_+ from Δ_+ to Δ'_+ and φ_- from Δ_- to Δ'_- that preserves codistances, i.e. $d^*(x,y) = \delta^*(\varphi_+(x), \varphi_-(y))$ for all $x \in T_+$ and $y \in T_-$. If both twin trees $(\Delta_+, \Delta_-, d^*)$ and $(\Delta'_+, \Delta'_-, \delta^*)$ are identical, then we call such a pair an *automorphism*.

Example 2.10. An easy example is the thin twin tree. Take two paths without endpoints, i.e. two 2-regular trees $\Delta_+ = (T_+, d_+)$ and $\Delta_- = (T_-, d_-)$, and choose any pair $(x_0, y_0) \in T_s \times T_{-s}$, where $s \in \{+, -\}$ and $-s = -$ if $s = +$ and $-s = +$ if $s = -$. We index the vertices in both trees depending on their distance to x_0 and y_0, respectively. So the vertices distinct from x_0 are denoted by $x_{\pm 1}, x_{\pm 2}, x_{\pm 3}, \ldots$ in Δ_s and vertices distinct from y_0 by $y_{\pm 1}, y_{\pm 2}, y_{\pm 3}, \ldots$ in Δ_{-s} such that vertices with a common sign lay in the same direction and $d_s(x_0, x_i) = |i| = d_{-s}(y_0, y_i)$ for all $i \in \mathbb{Z} \setminus \{0\}$.
Now, we define a symmetric map $d^* : (T_+ \times T_-) \cup (T_- \times T_+) \to \mathbb{N}$ via

$$d^*(x_0, y_0) := 0 =: d^*(y_0, x_0) \quad \text{and} \quad d^*(x_i, y_j) := |i-j| =: d^*(y_j, x_i)$$

for all $i, j \in \mathbb{Z} \setminus \{0\}$.
To verify that d^* is indeed a codistance, let x_i be a vertex in T_s and y_j be a vertex in T_{-s}. If z is a neighbour of y_j, then $z \in \{y_{j-1}, y_{j+1}\}$ and

$$d^*(x_i, z) = |i - (j \pm 1)| = |i - j \mp 1| \leqslant |i-j| + 1$$

as well as

$$d^*(x_i, z) = |i - (j \pm 1)| = |i - j \mp 1| \geqslant \Big||i-j| - 1\Big|$$

by the triangle inequalities. The latter expression equals 1, if $|i-j| = 0$, and $|i-j| - 1$, if $|i-j| \geqslant 1$. Since $|i-j| \neq |i-j| \pm 1$ for all $i, j \in \mathbb{Z}$, we infer

$$d^*(x_i, z) = |i-j| \pm 1 = d^*(x_i, y_j) \pm 1,$$

and (TT1) is shown. For (TT2), let $d^*(x_i, y_j) > 0$. Then

$$d^*(x_i, y_{j-1}) = |i - j + 1| = i - j + 1$$

and

$$d^*(x_i, y_{j+1}) = |i - j - 1| = i - j - 1.$$

Hence there is one unique neighbour of y_j such that the codistance increases. This proves that $(\Delta_+, \Delta_-, d^*)$ is a twin tree.

Example 2.11. In their last remark in [18, Section 1], Ronan and Tits point out a non-thin example. It is the twin tree attached to the general linear group $\mathrm{GL}_2(\mathbb{F}_2[t, t^{-1}])$ of the free module of rank 2 over the ring $\mathbb{F}_2[t, t^{-1}]$ of Laurent polynomials over the field $\mathbb{F}_2$, i.e. the group of all invertible (2×2)-matrices with entries in the given ring. It turns out that the constructed twin tree is 3-regular:

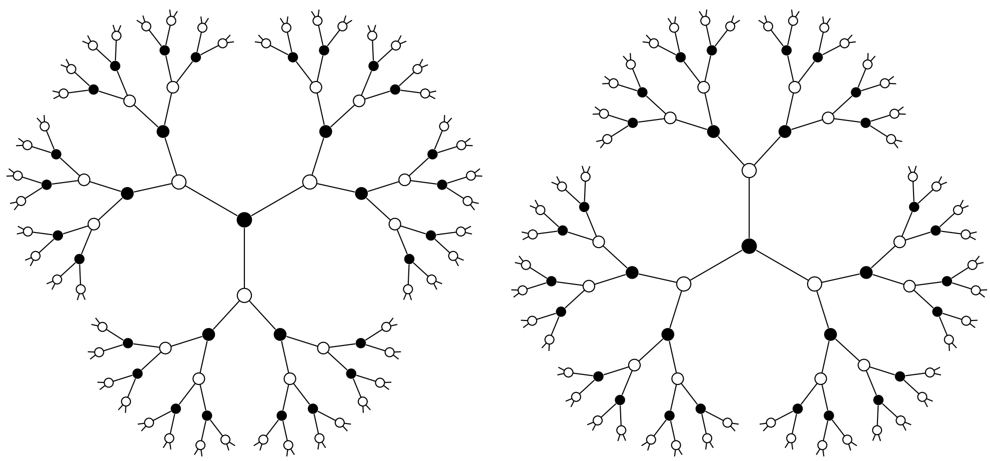

Figure 2.1: Bruhat-Tits trees for $\mathrm{GL}_2(\mathbb{F}_2(t))$
Source: own representation, inspired by [21, Chapter II.2, Example 2.4.1]

The construction of the twin tree for $\mathrm{GL}_2(\mathbb{K}[t, t^{-1}])$, where $\mathbb{K}$ is an arbitrary field, is the standard example and outlined in detail in [18, Section 2]. The halves Δ_+ and Δ_- of the twin tree for $\mathrm{GL}_2(\mathbb{K}[t, t^{-1}])$ are the Bruhat-Tits trees for $\mathrm{GL}_2(\mathbb{K}(t)) \geq \mathrm{GL}_2(\mathbb{K}[t, t^{-1}])$, where $\mathbb{K}(t)$ is the field of rational functions in one variable over $\mathbb{K}$, and are obtained by using the discrete valuations with uniformizer t and t^{-1}, respectively (see also [1, Section 6.9] or [21, Chapter II, Section 1.1 and 1.6]). The vertex sets of those trees are sets of classes of $\mathbb{K}[t]$-lattices resp. $\mathbb{K}[t^{-1}]$-lattices and the distances between two lattice classes (of the same tree) are defined via quotients of certain representative lattices. Two such classes are adjacent if and only if their quotient is isomorphic to the residue field $\mathbb{K}$. Ronan and Tits show in [18, Proposition 2.2] that their defined codistance gives indeed rise to a twin tree.

We say that two vertices of a tree are of *the same type* if and only if the distance between them is even. This partitions the set of vertices of a tree in two disjoint sets. This observation extends to twin trees if we say that vertices of even codistance are of the same type. By [18, Proposition 1], two trees of a thick twin tree are isomorphic, since vertices of the same type have same valency.

Remark 2.12. Concerning the previous example, we want to clarify that in [1] and [21], only the subgroup $\mathrm{SL}_2(\mathbb{K}[t,t^{-1}])$ of $\mathrm{GL}_2(\mathbb{K}[t,t^{-1}])$ containing matrices with determinant 1 is considered. It is a 'minimal version' of a Kac-Moody group (see the paper '*Uniqueness and presentation of Kac-Moody groups over fields*' by Tits ([27, Section 3.10(d)])) and its elements preserve the type of the vertices of the twin tree (cf. [18, Corollary 4.3]). We refer to [1, Section 8.11] for a short introduction to groups of Kac-Moody type together with many references.

Definition 2.13. An infinite sequence $A = \dots x_{i-1}x_ix_{i+1}x_{i+2}\dots$ of pairwise distinct vertices x_i with $i \in \mathbb{Z}$ in a tree is called an *apartment*. It is an infinite path without endpoints.
A *half-apartment* in a tree is an infinite sequence $x_kx_{k+1}x_{k+2}x_{k+3}\dots$ of pairwise distinct vertices x_i with $k \leqslant i \in \mathbb{Z}$, i.e. an infinite path without a terminal vertex.
We say that two half-apartments have the same end, if their intersection is a half-apartment. This gives rise to an equivalence relation on the set of half-apartments of a tree and we call the equivalence classes *ends of the tree*.

Equivalently, we may view (half-)apartments as infinite sequences of pairwise distinct edges; especially if we think of a tree as a chamber complex.

Note that an end e and an vertex x determine a unique half-apartment which is denoted by (xe). Furthermore, any apartment has two ends and any two ends e and f determine a unique apartment. We denote it by (ef).

Before we give the construction of a pair $(e_+, e_-) \in \Delta_+ \times \Delta_-$ of ends in a twin tree $(\Delta_+, \Delta_-, d^*)$, we state two useful observations of Ronan and Tits that provide the basis for the construction below. The first one directly follows from (TT2) and latter is a consequence of the preceding one:

Lemma 2.14. *([18, Proposition 3.1]) Let $s \in \{+,-\}$, x be a vertex in Δ_s, and $P = y_0y_1y_2\dots$ be a path in Δ_{-s} such that $d^*(x,y_0) = n$ and $d^*(x,y_1) = n-1$ for some $n \geqslant 1$. Then the codistance from x decreases monotonically along P (until it reaches the vertex y_n opposite x if P is long enough).*

Corollary 2.15. *([18, Corollary 3.2]) Let $s \in \{+, -\}$, x be a vertex in Δ_s, and P be an infinite path in Δ_{-s}. If P is a half-apartment, then the codistance from x either reaches 0 at some vertex of P or increases monotonically along P. Particularly, if P is an apartment, then $d^*(x, y) = 0$ for some vertex y of P.*

Construction 2.16. Let $(\Delta_+, \Delta_-, d^*)$ be a twin tree and let $s \in \{+, -\}$. A pair $(x, y) \in T_s \times T_{-s}$ of non-opposite vertices determines one end e_s of Δ_s and on end e_{-s} of Δ_{-s} as follows:

There exists a unique half-apartment $y_0 y_1 y_2 \ldots$ in Δ_{-s} starting in $y = y_0$ along which the codistance from x increases by (TT2). We call its end e_{-s}. Analogously, there is a unique half-apartment $x_0 x_1 x_2 \ldots$ in Δ_s starting in $x = x_0$ along which the codistance from y increases and whose end will be denoted by e_s.

The codistance between vertices on these two half-apartments is easy to calculate (cf. [18, Lemma 3.3]). With the notion above we have

$$d^*(x_i, y_j) = d^*(x, y) + i + j$$

for all $i, j \in \mathbb{N}$.

If we have ends like this determined by another pair of non-opposite vertices such that one of these ends equals e_s, then the other end necessarily equals e_{-s} (see [18, Proposition 3.4]). Hence, one identifies these ends and talks of the *ends of the twin tree* or *ends of the twinning*. We say that the half-apartments (xe_s) and (ye_{-s}) have the same end e and may write (xe) and (ye) instead.

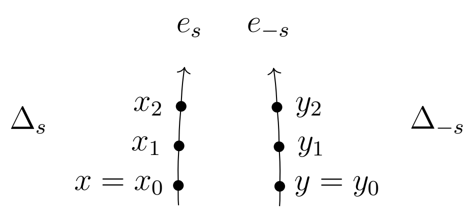

Figure 2.2: Two half-apartments with the same end e.

Note that this construction does not yield all ends of Δ_+ or Δ_-. For example, if both trees are 3-regular, they have countable many vertices (take a vertex x and the countable union of the disjoint sets of vertices with distance $n \in \mathbb{N}$ from x which have cardinality 1 if $n = 0$ or else $3 \cdot 2^{n-1}$), and thus determine countably many ends of the twinning. But one single tree has uncountably many ends (take a vertex x and identify each half-apartment $xx_1x_2x_3\ldots$, which determines a unique end of the tree, with the sequence $(z_i)_{n\in\mathbb{N}}$ with $z_0 \in \{0,1,2\}$ and $z_i \in \{0,1\}$ for $i \geqslant 1$; this yields indeed all possible sequences including the binary ones which are uncountable).

Definition 2.17. Given a twin tree $(\Delta_+, \Delta_-, d^*)$, let A_+ and A_- be apartments of Δ_+ and Δ_-, respectively. The pair (A_+, A_-) is called a *twin apartment* if each vertex of one apartment is opposite exactly one vertex of the other apartment.

The construction of a pair of half-apartments above can be extended to the construction of a twin apartment as shown in the proof of [18, Proposition 3.5] which states that each pair of opposite edges is contained in a unique twin apartment and that its apartments have the same two ends. First, we give the definition of opposite edges:

Definition 2.18. Let $s \in \{+,-\}$ and ϵ_s be an edge with vertices x and x' in Δ_s and ϵ_{-s} be an edge with vertices y and y' in Δ_{-s}. Such a pair of edges is called *opposite* if x is opposite one vertex y' and x' is opposite the other vertex y.

Construction 2.19. Now, let $\epsilon_s = \{x, x'\}$ be an edge in Δ_s opposite an edge $\epsilon_{-s} = \{y, y'\}$ in Δ_{-s} with x opposite y'. Then $d^*(x,y) = 1 = d^*(x',y')$ and the pairs (x,y) and (x',y') determine half-apartments $x_0x_{-1}x_{-2}\ldots$ and $y_1y_2y_3\ldots$ with the same end e, where $x_0 = x$ and $y_1 = y$, as well as half-apartments $x_1x_2x_3\ldots$ and $y_0y_{-1}y_{-2}\ldots$ with the same end f, where $x_1 = x'$ and $y_0 = y'$, respectively. We adjoin (xe) and $(x'f)$ to an apartment A_s in Δ_s as well as (ye) and $(y'f)$ to an apartment A_{-s} in Δ_{-s}, i.e. we obtain two apartments with the same two ends.
The following figure shows the twin apartment constructed above, where the orange lines indicate which vertices are opposite:

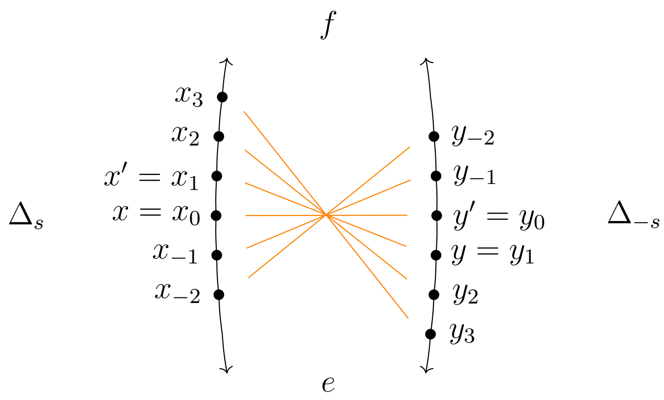

Figure 2.3: The twin apartment containing a pair $(\epsilon_s, \epsilon_{-s})$ of opposite edges.

Furthermore, it follows from the proof of the last cited proposition in Ronan and Tits' paper that $d^*(x_i, y_j) = |i - j|$ for all $i, j \in \mathbb{Z}$, so that x_i is opposite y_j if and only if $i = j$. Thus, the pair (A_s, A_{-s}) is indeed a twin apartment. Uniqueness can be shown by applying Lemma 2.14 to get a second vertex opposite either x, x', y or y', contradicting the defining property of a twin apartment.

On the contrary, each twin apartment contains infinitely many pairs of opposite edges. If we take an arbitrary twin apartment and choose a pair of opposite edges, then the constructed twin apartment above coincides with the given one. Hence all twin apartments are of this form.

Note that a twin apartment is a thin twin tree as described in Example 2.10.

It is a fact that in a thick twin tree, the set of twin apartments uniquely determines the codistance and thus the twin tree, as it determines the pairs of opposite vertices (cf. [18, Proposition 3.6]).

At last we define when a twin tree is called Moufang (see [18, p. 475-476]):

Definition 2.20. Let (A_+, A_-) be a twin apartment with ends e and f in a twin tree $(\Delta_+, \Delta_-, d^*)$ and write $A_s = (e_s f_s)$ for $s \in \{+, -\}$. Let x be a vertex of A_s and y be the unique vertex of A_{-s} opposite x. We then obtain one *root* $\alpha := (xe_s) \cup (ye_{-s})$ containing e and its *opposite root* $-\alpha := (xf_s) \cup (yf_{-s})$ containing f. In other words, we denote half a twin apartment with the term root. The set $\partial\alpha := \{x, y\}$ is called the *boundary* of α (and is equal to the boundary of its opposite).

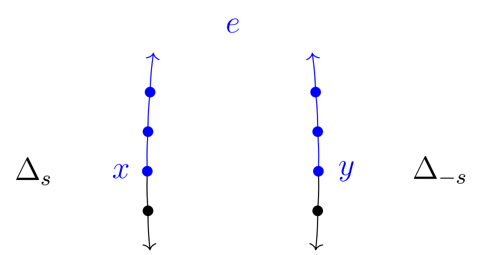

Figure 2.4: A root α with boundary $\{x, y\}$ containing e.

Definition 2.21. Given a root α of $(\Delta_+, \Delta_-, d^*)$, we denote by U_α the group of automorphisms of $(\Delta_+, \Delta_-, d^*)$ fixing α and every edge that contains a vertex of $\alpha \setminus \partial\alpha$. We call U_α a *root group* if it acts transitively on the set of twin apartments containing α. If U_β is a root group for all roots β of the twin tree, then we call the twin tree *Moufang*.

Since a thin twin tree is itself a twin apartment, all groups U_α, with α a root of the twin tree, are trivial, but act transitively, so that any thin twin tree is an example for a Moufang twin tree.

As hinted in Example 2.11, a twin tree is much more rigid than its halves. The rigidity theorem of Ronan and Tits states this general fact:

Theorem 2.22. *([18, Theorem 4.1]) Let $(\Delta_+, \Delta_-, d^*)$ be a thick twin tree, $s \in \{+, -\}$, and $(\epsilon_s, \epsilon_{-s})$ be a pair of opposite edges. The only automorphism fixing one of those edges and the set of all edges having at least one vertex in common with the other one is the identity.*

As a direct implication thereof, we see that the action of U_α on the set of twin apartments containing a root α is free. Hence, this action is regular when U_α is a root group.

In view of [18, Lemma 4.4] which states that there is a bijection between the set of twin apartments that contain α and the set of edges not in α but containing a boundary vertex of α, we may equivalently observe the local action of U_α on the direct neighbourhood of a vertex in $\partial\alpha$, i.e. on the set of vertices adjacent to a boundary vertex that lie not in α.

For a twin tree to be Moufang it is sufficient that U_α is a root group for all roots α of a given twin apartment (see [18, Proposition 4.5]). Particularly, the group generated by these root groups acts transitively on the set of all twin apartments.

Since our definition of a $\mathbb{Z}$-system uses commutator relations inspired by properties of root groups of a Moufang twin tree, we state an appropriate result after a short preparation:

Let (A_+, A_-) be a twin apartment of a thick Moufang twin tree, let $s \in \{+, -\}$, let $\ldots, x_{-1}, x_0, x_1, \ldots$ be the vertices of A_s, and $\ldots, y_{-1}, y_0, y_1, \ldots$ be the vertices of A_{-s}, respectively, such that x_i is opposite y_i for all $i \in \mathbb{Z}$. Then $x_0 x_{-1} x_{-2} \ldots$ and $y_0 y_1 y_2 \ldots$ as well as $x_0 x_1 x_2 \ldots$ and $y_0 y_{-1} y_{-2} \ldots$ determine the same ends e and f, respectively. For $i \in \mathbb{Z}$ let α_i denote the root $(x_i e) \cup (y_i e)$ and $-\alpha_i = (x_i f) \cup (y_i f)$ its opposite root. Further, let U_i be the root group for α_i.

Lemma 2.23. *([18, Corollary 4.7]) Let $m+1 \leqslant n$. Then the commutator group $[U_m, U_n]$ is contained in the product $U_{m+1}U_{m+2}...U_{n-1}$ (which equals 1 if $m+1 = n$).*

For the root groups V_i of the opposite roots $-\alpha_i$ the statement also holds, but we need to exchange m and n, i.e. we have $[V_n, V_m] \leq V_{n+1}V_{n+2}...V_{m-1}$ for $n+1 \leqslant m$.

Example 2.24. In their last example in [18, Section 4], Ronan and Tits explicitly describe the root groups of the standard example of the (thick) twin tree for $\mathrm{GL}_2(\mathbb{K}[t, t^{-1}])$. With the notion introduced before the lemma, we have

$$U_i = \left\{ \begin{pmatrix} 1 & ct^i \\ 0 & 1 \end{pmatrix} \,\middle|\, c \in \mathbb{K} \right\} \quad \text{and} \quad V_i = \left\{ \begin{pmatrix} 1 & 0 \\ ct^{-i} & 1 \end{pmatrix} \,\middle|\, c \in \mathbb{K} \right\}.$$

Hence, those groups are isomorphic to the additive group $(\mathbb{K}, +)$ via

$$c \mapsto \begin{pmatrix} 1 & ct^i \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad c \mapsto \begin{pmatrix} 1 & 0 \\ ct^{-i} & 1 \end{pmatrix},$$

respectively. In particular, these root groups are abelian.

2.2 RGD- and $\mathbb{Z}$-systems

In his course '*Immeubles Jumelés*' from 1988/89 ([30, Section 9]), Jacques Tits provides a procedure to obtain all Moufang twin trees. The given construction is based on a system $(G; (U_\alpha)_{\alpha\in\Phi})$ consisting of a group G and a family of subgroups U_α, which are indexed by roots of a root system Φ, subject to a few conditions that involve commutator relations of certain pairs of subgroups from the family. Within this course, Tits also discusses if a classification of these systems is feasible which would yield a classification of all Moufang twin trees (and therefore Moufang sets in particular). In his paper '*Twin Buildings and Groups of Kac-Moody Type*' from 1992 ([28]), he introduced this system as RGD-system where RGD stands for 'root groups data'.
Nowadays, an RGD-system is usually defined as a triple that augments the pair $(G; (U_\alpha)_{\alpha\in\Phi})$ above by another subgroup which turns out to be the intersection of all normalizers $N_G(U_\alpha)$ (see [1, Definition 7.82, Remark 7.84, and Subsection 8.6.1]).

We are interested in Moufang twin trees that are associated to RGD-systems of certain type, namely of type $\widetilde{A}_1$. Since the root system of this type has a concrete description which we give in the first definition below, the general definition of an RGD-system can be considerably simplified to suit our case.

Definition 2.25. ([9, Definition 2.1]) For each $z \in \mathbb{Z}$ we define $\sigma_z := 1$ for $z \leqslant 0$ and $\sigma_z := -1$ for $z > 0$. Then the set $\Phi := \mathbb{Z} \times \{-1, 1\}$ is the *root system* (of type $\widetilde{A}_1$), $\Phi^+ := \{(z, \sigma_z) \mid z \in \mathbb{Z}\}$ is the set of *positive roots* and $\Phi^- := \Phi \setminus \Phi^+$ the set of *negative roots*.
For $\alpha = (z, \sigma) \in \Phi$ we set $-\alpha := (z, -\sigma)$. Furthermore, for $i = 0, 1$ we define $r_i \in \mathrm{Sym}(\Phi)$ via $(z, \sigma) \mapsto (2i - z, -\sigma)$ and set $\alpha_i := (i, \sigma_i) \in \Phi^+$. Observe that $\alpha_0.r_0 = -\alpha_0$ and $\alpha_1.r_1 = -\alpha_1$.

The following figure shows the root system, where white nodes are positive and black nodes are negative roots:

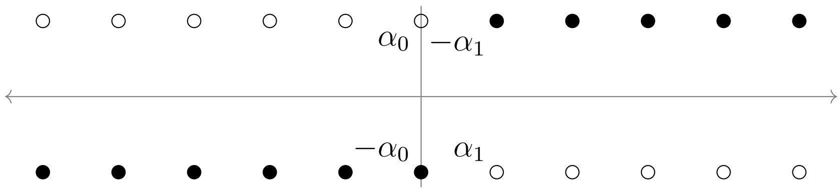

Figure 2.5: The root system of type $\widetilde{A}_1$.
Source: own representation, inspired by [9, Fig.1]

The permutation r_0 acts as the point reflection in $(0,0)$ and r_1 as the composition of r_0 followed by the translation by $(2,0)$.

Convention 2.26. For the remainder, let Φ always denote a root system of type $\widetilde{A}_1$ with the prominent roots $\alpha_i = (i, \sigma_i)$ and permutations $r_i : \Phi \to \Phi : (z, \sigma) \mapsto (2i - z, -\sigma)$ for $i = 0, 1$.

Definition 2.27. ([9, Definition 2.2]) An *RGD-system* (of type $\widetilde{A}_1$) is a triple $\mathcal{R} = (G, (U_\alpha)_{\alpha\in\Phi}, H)$ consisting of a group G, a family $(U_\alpha)_{\alpha\in\Phi}$ of subgroups of G, and a subgroup H of G such that the following conditions hold:

(R1) $|U_\alpha| > 1$ for all $\alpha \in \Phi$;

(R2) For all $z < z' \in \mathbb{Z}$ and all $\sigma \in \{-1, 1\}$ we have

$$\left[U_{(z,\sigma)}, U_{(z',\sigma)}\right] \subseteq \left\langle U_{(n,\sigma)} \,\middle|\, z < n < z' \right\rangle.$$

(R3) For $i = 0, 1$ there exists a function $m_i : U^*_{\alpha_i} \to G$ such that

$$m_i(u) \in U_{-\alpha_i} u U_{-\alpha_i} \quad \text{and} \quad m_i(u) U_\alpha m_i(u)^{-1} = U_{\alpha.r_i}$$

for all $u \in U^*_{\alpha_i}$ and for all $\alpha \in \Phi$.
Moreover, we have $m_i(u)^{-1} m_i(u') \in H$ for all $u, u' \in U^*_{\alpha_i}$.

(R4) For $i = 0, 1$ the group $U_{-\alpha_i}$ is not contained in $U_+ := \langle U_\alpha \mid \alpha \in \Phi^+ \rangle$.

(R5) G is generated by H and the family $(U_\alpha)_{\alpha\in\Phi}$.

(R6) $H \leq N_G(U_\alpha)$ for all $\alpha \in \Phi$.

In [1], Abramenko and Brown outline the correspondence between Moufang twin buildings and RGD-systems in general, i.e. how an RGD-system yields a Moufang twin building and vice versa. We refer to Section 8.3 with Proposition 8.22 via Section 8.6 and its Example 8.47 up to Section 8.9 cumulating in Theorem 8.81 for a detailed discussion.

Convention 2.28. If not explicitly stated otherwise, then by an RGD-system we will always mean one of type $\widetilde{A}_1$ in the following.

Example 2.29. ([9, Example 2.4]) Let $G = \mathrm{SL}_2(\mathbb{K}[t, t^{-1}])$ and let $z \in \mathbb{Z}$. We define $U_{(z,1)} := U_z$ and $U_{(z,-1)} := V_z$ to be the root groups of Example 2.24. We further set

$$H := \left\{ \begin{pmatrix} c^{-1} & 0 \\ 0 & c \end{pmatrix} \middle| \, c \in \mathbb{K}^* \right\} \leq \bigcap_{\alpha\in\Phi} N_G(U_\alpha).$$

Then $(G, (U_\alpha)_{\alpha\in\Phi}, H)$ is an RGD-system with trivial commutator relations.

Grüninger, Horn and Mühlherr say, in analogue to the projective plane, that a tree is of *order* $q \in \mathbb{N}$ if it is regular of degree $q+1$ (cf. [9, Remark 2.5(i)]).

In order to classify all Moufang twin trees of prime order p (note that these twin trees are thick), they want to classify all RGD-systems in which all subgroups U_α are of order p. They further conclude that these classifications are equivalent; and thus, as the finite Moufang sets are known, it remains to classify all possible commutator relations between them (see the discussion in [9, Section 2 and the first part of Section 3]).

Remark 2.30. Let $\mathcal{R} = (G, (U_\alpha)_{\alpha\in\Phi}, H)$ be an RGD-system and $\mathcal{T}_\mathcal{R}$ the associated Moufang twin tree. The *unipotent horocyclic group*, a subgroup of $\mathrm{Aut}(\mathcal{T}_\mathcal{R})$ and an important invariant of the Moufang twin tree, is the group $U_{++} := \langle U_{(z,1)} \mid z \in \mathbb{Z}\rangle$ (cf. [9, Remark 2.5(ii)]).

For the classification of the RGD-systems of interest, the three authors above introduce the following concept:

Definition 2.31. ([9, Definition 3.2]) Let $p \in \mathbb{P}$. A $\mathbb{Z}$-*system* of order p is a pair $(X, (x_k)_{k\in\mathbb{Z}})$ consisting of a group X and a family $(x_k)_{k\in\mathbb{Z}}$ of elements in X with the following properties:

(Z1) The group X is generated by the family $(x_k)_{k\in\mathbb{Z}}$.

(Z2) The subgroup $\langle x_k \mid m \leqslant k \leqslant n\rangle$ is of cardinality p^{n-m+1}.

(Z3) There exists an automorphism τ of X such that $\tau(x_k) = x_{k+2}$ for all $k \in \mathbb{Z}$.

Furthermore, the elements of the family satisfy the condition

(Z4) For all integers $m < n$ we have $[x_m, x_n] \in \langle x_k \mid m < k < n\rangle$,

which is (ZS5) of [9, Lemma 4.2].

The main result of their work is

Theorem 2.32. *([9, Theorem 3.4]) Let $(X, (x_k)_{k\in\mathbb{Z}})$ be a $\mathbb{Z}$-system of prime order. Then X is nilpotent of class at most 2.*

Since, given a Moufang twin tree of prime order and its associated RGD-system $(G, (U_\alpha)_{\alpha\in\Phi}, H)$, the group $U_{++} = \langle U_{(z,1)} \mid z \in \mathbb{Z}\rangle$ coincides with the group $X = \langle\langle x_k\rangle \mid k \in \mathbb{Z}\rangle$, they infer that the unipotent horocyclic group of the Moufang twin tree is nilpotent of class at most 2, again. Thus, their theorem narrows down the search for possible candidates for the invariant of certain locally finite Moufang twin trees.

In the present work we generalize their result to a considerably larger class of $\mathbb{Z}$-systems by introducing a more general concept of a $\mathbb{Z}$-system.

Our main result is similar to the main result of [9] stated above. We will present it as Theorem 7.7 which states that, given a soon to be defined quadruple $(X, (X_k)_{k\in\mathbb{Z}}, \varsigma, T)$, the group X, the analogue of the group of a $\mathbb{Z}$-system of prime order, is nilpotent of class at most 2.

An action of a group G on a group K is called *irreducible* (and K sometimes *G-irreducible*) if the only G-invariant subgroups of K are $\{1\}$ and K.

We adjust the properties (Z1) up to (Z4) of a $\mathbb{Z}$-system of prime order to obtain the definition of a general $\mathbb{Z}$-system. This was similarly done at the beginning of the unpublished manuscript '*Moufang-systems*' by my predecessor Moritz Scholz ([19]) for his so-called 'Moufang-systems'. His manuscript is based on an earlier version of our main source [9].

Property (M5) below is partially inspired by a convention in Scholz' notes (see [19, p. 6]).

Definition 2.33. A *$\mathbb{Z}$-system* is a quadruple $\Xi = (X, (X_k)_{k\in\mathbb{Z}}, \varsigma, T)$ consisting of a group X, a family of non-trivial subgroups $(X_k)_{k\in\mathbb{Z}}$, an automorphism $\varsigma \in \mathrm{Aut}(X)$, and a group T such that the following conditions hold:

(M1) $X = \langle X_k \mid k \in \mathbb{Z}\rangle$.

(M2) For all $m \leqslant n \in \mathbb{Z}$ the map

$$\begin{aligned} \rho_{m,n} : X_m \times X_{m+1} \times \ldots \times X_n &\to X_{m,n} \\ (x_m, x_{m+1}, \ldots, x_n) &\mapsto x_m x_{m+1} \cdot \ldots \cdot x_n \end{aligned}$$

is a bijection, where $X_{m,n} := \langle X_k \mid m \leqslant k \leqslant n\rangle$.

(M3) For all $m < n \in \mathbb{Z}$ we have $[X_m, X_n] \leq X_{m+1,n-1}$.

(M4) For all $k \in \mathbb{Z}$ we have $\varsigma(X_k) = X_{k+2}$.

(M5) T is a subgroup of $\mathrm{Aut}(X)$ normalized by ς, i.e. $T^\varsigma = T$, and X_k is T-invariant for all $k \in \mathbb{Z}$.

The members X_k of the family are called *rooted groups* of Ξ, the automorphism ς is called the *shift operator* of Ξ, and the group T is called the *torus* of Ξ. Furthermore, for $n \in \mathbb{Z}$ we set

$$X_{-\infty,n} := \langle X_k \mid k \leqslant n\rangle \quad \text{and} \quad X_{n,\infty} := \langle X_k \mid k \geqslant n\rangle.$$

A $\mathbb{Z}$-system is called *irreducible* if T acts irreducibly on X_k for all $k \in \mathbb{Z}$. If all rooted groups are nilpotent, we call a $\mathbb{Z}$-system *nilpotent*.

Note that if such a system is irreducible, then X_k is the smallest T-invariant subgroup containing any of its elements $x_k \neq 1$.

Remark 2.34. Let $\Xi = (X, (X_k)_{k\in\mathbb{Z}}, \varsigma, T)$ be a $\mathbb{Z}$-system. Then

$$\Xi^{pc} := (X, (X_{k+1})_{k\in\mathbb{Z}}, \varsigma, T) \quad \text{and} \quad \Xi^{inv} := \left(X, \left(X_k^{inv}\right)_{k\in\mathbb{Z}}, \varsigma^{-1}, T\right),$$

where $X_k^{inv} := X_{-k}$ for $k \in \mathbb{Z}$, are $\mathbb{Z}$-systems, again.
It is clear that Ξ^{pc}, where 'pc' suggestively stands for parity change, is a $\mathbb{Z}$-system. We show that Ξ^{inv}, where 'inv' indicates the inverse indexing, satisfies all conditions for a $\mathbb{Z}$-system, since we will use this system in a proof later on. Observe that

$$X_{m,n}^{inv} = \left\langle X_k^{inv} \,\middle|\, m \leqslant k \leqslant n \right\rangle = \langle X_{-k} \mid -n \leqslant -k \leqslant -m \rangle = X_{-n,-m}$$

for all $m \leqslant n$. Since

$$\begin{aligned} f_{m,n} : X_m^{inv} \times \ldots \times X_n^{inv} &\to X_{-n} \times \ldots \times X_{-m}, \\ (x_{-m}, \ldots, x_{-n}) &\mapsto (x_{-n}, \ldots, x_{-m}) \end{aligned}$$

is bijective for all $m \leqslant n$, the maps

$$\begin{aligned} \rho_{m,n}^{inv} := \rho_{-n,-m} \circ f_{m,n} : X_m^{inv} \times \ldots \times X_n^{inv} &\to X_{m,n}^{inv}, \\ (x_{-m}, \ldots, x_{-n}) &\mapsto x_{-n} \cdot \ldots \cdot x_{-m} \end{aligned}$$

are bijective for all $m \leqslant n$, again. Moreover, we have

$$\left[X_m^{inv}, X_n^{inv}\right] = [X_{-m}, X_{-n}] = [X_{-n}, X_{-m}] \leq X_{-(n-1),-(m+1)} = X_{m+1,n-1}^{inv}$$

for all $m < n$ and

$$\varsigma^{-1}\left(X_k^{inv}\right) = \varsigma^{-1}(X_{-k}) = X_{-(k+2)} = X_{k+2}^{inv}$$

for all $k \in \mathbb{Z}$. At last we note that the groups X_k^{inv} are still T-invariant and $T^{\varsigma} = T \Leftrightarrow T = T^{\varsigma^{-1}}$.

As desired, we can directly observe that our definition generalized the $\mathbb{Z}$-systems of prime order:

Lemma 2.35. *Any $\mathbb{Z}$-system of prime order is an irreducible and nilpotent $\mathbb{Z}$-system.*

Proof. Let $(X, (x_k)_{k\in\mathbb{Z}})$ be a $\mathbb{Z}$-system of prime order and $T = \{\mathrm{id}_X\}$. The groups $X_k := \langle x_k \rangle$, with $k \in \mathbb{Z}$, have order p by (Z2), are cyclic, hence abelian and thus nilpotent, and they generate X by (Z1). As they do not contain any proper non-trivial subgroup, the action of T on X_k is trivially irreducible. Since the automorphism τ in (Z3) sends generators to generators, it maps X_k onto X_{k+2} for all $k \in \mathbb{Z}$. It also normalizes T. Hence, the preliminaries as well as (M1), (M4) and (M5) above are satisfied by the tuple $(X, (X_k)_{k\in\mathbb{Z}}, \tau, T)$. In addition, (M3) follows by (Z4) together with Lemma 1.12, the general fact that $x^y = x[x, y]$ for any $x, y \in X$, and an induction to see that $[x_m^i, x_n^j] \in \langle x_k \mid m < k < n\rangle$ for all $i, j \in \mathbb{N}$. As both

$$X_{m,n} := \langle X_k \mid m \leqslant k \leqslant n\rangle \quad \text{and} \quad X_m \times X_{m+1} \times \ldots \times X_n$$

have order p^{n-m+1} for all $m \leqslant n \in \mathbb{Z}$ by (Z2), we can define a bijection $\rho_{m,n}$ from $X_m \times X_{m+1} \times \ldots \times X_n$ to $X_{m,n}$ conform to (M2) for all $m \leqslant n$. $\square$

Example 2.36. A Moufang twin tree of prime order yields an irreducible and nilpotent $\mathbb{Z}$-system.
Indeed, let $(G, (U_\alpha)_{\alpha\in\Phi}, H)$ be the associated RGD-system. By the lemma above, it suffices to show that this RGD-system yields a $\mathbb{Z}$-system of prime order. Each group U_α is of prime order. For $k \in \mathbb{Z}$ we set $X_k := U_{(k,1)}$ as well as $X := \langle X_k \mid k \in \mathbb{Z}\rangle$. By using the maps m_i in (R3) to obtain elements $s_i \in G$ such that U_α is conjugated onto $U_{\alpha.r_i}$ for all $\alpha \in \Phi$, we define an automorphism $\tau \in \mathrm{Aut}(X)$ via $x \mapsto x^{s_0 s_1}$ that sends X_k to X_{k+2} for all $k \in \mathbb{Z}$ (see [9, Lemma 3.1(ii)]). At last, we choose two non-trivial elements $x_0 \in X_0$ and $x_1 \in X_1$ and set $x_{2k} := \tau^k(x_0)$ and $x_{2k+1} := \tau^k(x_1)$ for all $k \in \mathbb{Z}$. Then $X = \langle\langle x_k\rangle \mid k \in \mathbb{Z}\rangle$ and the pair $(X, (x_k)_{k\in\mathbb{Z}})$ fulfils (Z1) and (Z3) as well as (Z2) by [9, Lemma 3.1(i)].

Convention 2.37. Let $\Xi = (X, (X_k)_{k\in\mathbb{Z}}, \varsigma, T)$ denote a $\mathbb{Z}$-system. We assume that T acts from the left on X (maybe via a homomorphism $\lambda : T \to \mathrm{Aut}(X)$, but then we will omit the notion of λ).

2.3 An example based on Moufang sets

The last example of this section shows that the assumption of irreducibility of a (nilpotent) $\mathbb{Z}$-system is quite natural and does not reduce our study to the trivial case. Primary, we make use of the papers '*A course on Moufang sets*' by Tom De Medts and Yoav Segev from 2009 ([4]) and '*On the action of the Hua subgroups in special Moufang sets*' by Segev and Richard M. Weiss from 2008 ([20]).

For it to be stated, we need the terminology of a Moufang set that appeared a few times up to now and has its origin in Tits' article [28, Section 4.4]. These sets correspond to rank 1 Moufang buildings.

Our definition of a Moufang set is based on the one appearing in [4]. Moreover, we may adapt facts about Moufang sets directly from the earlier papers '*Moufang sets and Jordan division algebras*' by De Medts and Weiss from 2006 ([5]) and '*Identities in Moufang sets*' by De Medts and Segev from 2008 ([3]), respectively.

Definition 2.38. ([4, Section 1.2]) A *Moufang set* is a set X of cardinality at least 3 together with a collection $(U_x)_{x\in X}$ of groups satisfying the following two properties:

(S1) For all $x \in X$ the group U_x is a subgroup of $\mathrm{Sym}(X)$, fixes x, and acts regularly on $X \setminus \{x\}$.

(S2) For all $x \in X$ the group U_x permutes the family $(U_y)_{y\in X}$ by conjugation.

We denote the Moufang set by the pair $\mathbb{M} := (X, (U_x)_{x\in X})$. The groups U_x are called *root groups* of $\mathbb{M}$ and the group $G^\dagger := \langle U_x \mid x \in X\rangle$ is called the *little projective group* of $\mathbb{M}$.

Remark 2.39. The first assertion implies that $U_x \neq U_y$ for all $x \neq y \in X$ and that $G^\dagger$ is transitive one X. In fact, it acts doubly transitively on X, since $U_x \leq \mathrm{St}_{G^\dagger}(x)$ is transitive on $X \setminus \{x\}$ for all $x \in X$ (cf. the book '*Permutation groups*' by Dixon and Mortimer [7, Exercise 2.1.3]).
The second assertion implies that the root groups are isomorphic to each other.

The upcoming lemma shows that we may replace (S2) by another property (S2'). It is used, for example, in the definition of a Moufang set in [5] and [20]. We use the notion of a right action for the action of G on X in the remainder.

Lemma 2.40. *Under (S1) the condition (S2) is equivalent to*

(S2') For all $y \in X$ and for all $g \in G^\dagger$ we have $U_y^g = U_{y.g}$.

Proof. Let $x \in X$ be arbitrary. First we show (S2) $\Rightarrow$ (S2'). Let $y \in X$ and $g \in U_x$. By (S2) there exists $x(y, g) \in X$ such that $U_g^y = U_{x(y,g)}$. As $G^\dagger$ is transitive on X, we observe

$$U_y^g \leq (\mathrm{St}_{G^\dagger}(y))^g = \mathrm{St}_{G^\dagger}(y.g),$$

so that U_y^g fixes $y.g$. We conclude $U_y^g = U_{x(y,g)} = U_{y.g}$ with (S1). As $G^\dagger$ is generated by the root groups, we infer $U_y^g = U_{y.g}$ for all $y \in X$ and $g \in G^\dagger$. Conversely, assume that (S2') holds and take $x \in X$. Then $U_y^g = U_{y.g}$ for all $y \in X$ and $g \in U_x$. Now, let $g \in U_x$ be arbitrary. By the first observation in the preceding remark, we have $U_x^g = U_y^g \Leftrightarrow U_x = U_y \Leftrightarrow x = y$ for $x, y \in X$. Moreover, since U_x fixes x and is transitive on $X \setminus \{x\}$, we have $U_{y.g^{-1}}^g = U_{(y.g^{-1}).g} = U_y$ for all $y \in X$. Hence each element of U_x permutes the family $(U_y)_{y\in X}$. $\square$

Following the main construction of De Medts and Segev in [4, Section 3], there is a way to obtain a Moufang set from a group U and a permutation τ of U^*:

Construction 2.41. Let U be a group (with additive notion as usual in the context of Moufang sets) and let S denote the disjoint union of U and $\{\infty\}$, where ∞ is just a new symbol.
For $a \in U$ we define a permutation $\pi_a \in \mathrm{Sym(S)}$ by $\infty \mapsto \infty$ and $s \mapsto s+a$ for all $s \in U$. We set $U_\infty := \{\pi_a \mid a \in U\}$. Observe that this subgroup of $\mathrm{Sym}(S)$ fixes ∞ and acts regularly on $U = S \setminus \{\infty\}$. We use the notation of a right action for permutations in $\mathrm{Sym}(S)$ for now.
Let τ be a permutation of U^*. We extend τ to a permutation of S via $0.\tau = \infty$ and $\infty.\tau = 0$. Furthermore, we define $U_0 := U_\infty^\tau$ as well as $U_a := U_0^{\pi_a}$ for all $a \in U^*$. Note that U_0 fixes 0, and thus U_a fixes a for all $a \in S \setminus \{\infty, 0\}$.
At last, we define $\mathbb{M}(U,\tau) := (S,(U_s)_{s\in S})$ and the group $G^\dagger := \langle U_\infty, U_0\rangle = \langle U_s \mid s \in S\rangle$.

Remark 2.42. Let U be of order at least 2. Then the group U_s acts regularly on $S \setminus \{s\}$ for all $s \in S$.
Indeed, this is clear for $s = \infty$. For $s = 0$, if there exists $u \in U_0^*$ fixing some $t \in S$, then there is $a \in U^*$ such that $t = t.u = t.\pi_a^\tau$. Equivalently, π_a fixes $t.\tau^{-1}$, i.e. $t = \infty.\tau = 0$. Therefore, the action of U_0 on $S \setminus \{0\}$ is free. If $x, y \in S \setminus \{0\}$, then for $b := -(x.\tau^{-1}) + y.\tau^{-1}$ we have $x.\pi_b^\tau = y$; hence, U_0 acts regularly on $S \setminus \{0\}$.
Analogously, it follows for each $a \in U^*$ that the group U_a acts regularly on $S \setminus \{a\}$, since we observe $x - a \neq 0 \neq y - a$ for $x, y \in S \setminus \{a\}$ and use the regularity of U_0 to obtain a unique element $u \in U_0$ such that $(x-a).u = y-a \Leftrightarrow (x.\pi_a^{-1}).u = y.\pi_a^{-1} \Leftrightarrow x.u^{\pi_a} = y$.

Hence, $\mathbb{M}(U,\tau)$ fulfils (S1), if U is non-trivial. Below we will give a criterion for $\mathbb{M}(U,\tau)$ to be a Moufang set.

Example 2.43. Let $(\Delta_+, \Delta_-, \delta^*)$ be a thick Moufang twin tree. Choose a twin apartment $A = (A_+, A_-)$ with ends e and f as well as a pair (x, y) of opposite vertices within A. Let α denote a root with boundary $\{x, y\}$ containing an end, say e, of the twin apartment. Let x_∞ be the neighbouring vertex of x in α and x_0 be the vertex adjacent to x in $-\alpha$.

Since the action of the root group $U := U_\alpha$ is regular on the set V of vertices adjacent to x excluding x_∞, we have $|U| = |V| \geqslant 2$ and may identify a group element $a \in U$ with the unique vertex $x_a \in V$ such that $x_a = x_0.a$. We also identify the vertex x_∞ with the new symbol ∞.
Let $W := V \cup \{x_\infty\}$. First we show that W together with the family of root groups indexed by the roots with extremity x form a Moufang set.

By [18, Lemma 4.4], we can equivalently index these groups by the set W, since each edge on x not in α lies in a unique twin apartment containing α, and so in a unique root that contains this edge and its other vertex $w \in W \setminus \{x_\infty\}$. Using [18, Corollary 4.9] which states that if $v \in V$ and if β is a root containing v with extremity x, then the group of permutations induced on V by the root group U_β depends only on x and v and not on the particular choice of β, this is indeed justified. In particular, we have $U = U_{x_\infty}$.

Since the root group U_w acts regularly on $W \setminus \{w\}$ and fixes w for all $w \in W$, the pair $(W, (U_w)_{w\in W})$ naturally satisfies (S1). Let $w \neq w' \in W$ and $u \in U_w$. It is clear that $U_w^u = U_w$. If $u' \in U_{w'}$, then u'^u fixes $w'.u$ as well as the root β containing $w'.u$ and all edges containing a vertex in $\beta \setminus \partial\beta$, so that $U_{w'}^u \subseteq U_{w'.u}$. Similarly, if $u' \in U_{w'.u}$, then $u'^{u^{-1}}$ fixes w' as well as the root and edges fixed by $U_{w'}$, so that $U_{w'.u}^{u^{-1}} \subseteq U_{w'} \Leftrightarrow U_{w'.u} \subseteq U_{w'}^u$ and both groups are equal. By transitivity of U_w on $W \setminus \{w\}$, for $w'' \neq w$ we have $w''.u^{-1} \neq w$ such that $(w''.u^{-1}).u = w''$; hence $U_{w''.u^{-1}}^u = U_{w''}$. Moreover, we observe $U_{w'}^u = U_{w''}^u \Leftrightarrow U_{w'} = U_{w''}$. Thus, u permutes the family $(U_w)_{w\in W}$ by conjugation and (S2) follows.

We have shown that $(W, (U_w)_{w\in W})$ is actually a Moufang set.

Now, let $S := U \cup \{\infty\}$. Following Construction 2.41 above, we define a pair $\mathbb{M}(U, \tau) = (S, (U_s)_{s\in S})$ and proof that it is also a Moufang set. It is essentially the same as the one we obtain directly, since we transfer properties of $(W, (U_w)_{w\in W})$ by the identification $\iota : W \to S$ above that sends x_∞ to ∞ and $v \in V$ to $a \in U$ where a is uniquely determined by $v = x_0.a$.
In alignment with the construction, we use the additive notion for the group operation in the (original) root groups for the remainder of this example.

For each $a \in U$ the restriction of the automorphism a to W is a permutation of W with $x_\infty \mapsto x_\infty$ and $x_s \mapsto x_{s+a}$ for all $s \in U$. This permutation induces a permutation $\pi_a = \iota^{-1} a \iota$ on S with $\infty \mapsto \infty$ and $s \mapsto s + a$ for all $s \in U$ via the identification mentioned above. We set $U_\infty := \{\pi_a \mid a \in U\} \leq \mathrm{Sym}(S)$.

Analogously to the proof of [18, Proposition 4.5], there exists an automorphism of the Moufang twin tree whose restriction to the set W is a permutation interchanging x_0 and x_∞. In particular, this automorphism conjugates U_{x_∞} onto U_{x_0}. Indeed, let $g \in U^*$. Using the transitivity of $U_{-\alpha}$, there are elements $h(g), h(-g) \in U_{-\alpha}^*$ such that $x_0.g = x_\infty.h(g)$ and $x_0.(-g) = x_\infty.h(-g)$.

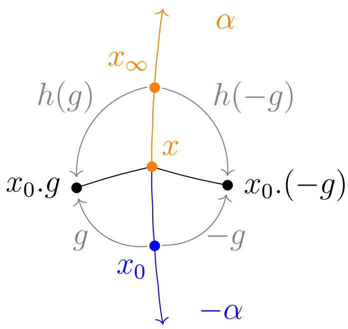

Figure 2.6: Interchanging two neighbours of a boundary vertex in a given twin apartment.

Since $U_{-\alpha}$ fixes x_0, we compute

$$\begin{aligned} x_0.(h(-g)+g-h(g)) &= (x_0.h(-g)).(g-h(g)) = x_0.(g-h(g)) \\ &= (x_0.g).(-h(g)) = (x_\infty.h(g)).(-h(g)) \\ &= x_\infty.(h(g)-h(g)) = x_\infty.0 = x_\infty \end{aligned}$$

as well as

$$\begin{aligned} x_\infty.(h(-g)+g-h(g)) &= (x_\infty.h(-g)).(g-h(g)) = (x_0.(-g)).(g-h(g)) \\ &= (x_0.(-g+g)).(-h(g)) = (x_0.0).(-h(g)) \\ &= x_0.(-h(g)) = x_0. \end{aligned}$$

The element $\gamma := h(-g)+g-h(g) \in \mathrm{Sym}(W)$ now gives rise via conjugation by ι to a permutation τ of S that interchanges 0 and ∞. Using it, we define the permutation groups $U_0 := U_\infty^\tau$ and $U_a := U_0^{\pi_a}$ for all $a \in U^*$. Thus, we get a pair $\mathbb{M}(U,\tau) := (S,(U_s)_{s\in S})$ and a group $G^\dagger := \langle U_s \mid s \in S\rangle$ as laid down in the construction. As a direct consequence of the previous remark, the pair $\mathbb{M}(U,\tau)$ satisfies (S1). Using ι, we verify that $U_s = U^\iota_{\iota^{-1}(s)} = U^\iota_{x_s}$ holds for all $s \in S$. If $s=\infty$, then

$$U^\iota_{x_\infty} = \{a^\iota \mid a \in U_{x_\infty} = U\} = \{\pi_a \mid a \in U\} = U_\infty.$$

For $s=0$, by inserting $\iota\iota^{-1}$ we have

$$\begin{aligned} U^\iota_{x_0} &= \left(U^\gamma_{x_\infty}\right)^\iota = \{(a^\gamma)^\iota \mid a \in U\} = \{(-\gamma)^\iota \pi_a \gamma^\iota \mid a \in U\} \\ &= \{\pi_a^\tau \mid a \in U\} = U_0. \end{aligned}$$

In the remaining cases we observe

$$U^\iota_{x_s} = \left(U^s_{x_0}\right)^\iota = (-s)^\iota U^\iota_{x_0} s^\iota = U_0^{\pi_s} = U_s.$$

At last, let $s, s' \in S$, $u \in U_{s'}$ and $z \in U_{x_{s'}}$ with $u = z^\iota$. Then the proven equality $U_s = U_{x_s}^\iota$ for all $s \in S$ and the fact that z permutes the family $(U_w)_{w \in W}$ by conjugation yield

$$\begin{aligned} U_s^u &= \left(U_{x_s}^\iota\right)^u = u^{-1}\iota^{-1}U_{x_s}\iota u = \iota^{-1}z^{-1}U_{x_s}z\iota = \left(U_{x_s}^z\right)^\iota \\ &= U_{x_s.z}^\iota = U_{\iota(x_s.z)} = U_{\iota(\iota^{-1}(s).z)} = U_{s.(\iota^{-1}z\iota)} = U_{s.u}. \end{aligned}$$

Like above, it can now be readily verified that $\mathbb{M}(U, \tau)$ satisfies (S2) and is indeed a Moufang set as claimed.

Definition 2.44. ([5, Definition 3.1 and Definition 3.2]) Let $\mathbb{M}(U, \tau)$ be as in the construction above and let $a \in U^*$. We define the following permutations of S:

$$\begin{aligned} \mu(a) &:= \pi^\tau_{(-a).\tau^{-1}}\pi_a \left(\pi^\tau_{a.\tau^{-1}}\right)^{-1} = \pi^\tau_{(-a).\tau^{-1}}\pi_a\pi^\tau_{-(a.\tau^{-1})}, \\ h_a &:= \tau\pi_a\tau^{-1}\pi^{-1}_{a.\tau^{-1}}\tau\pi^{-1}_{(-(a.\tau^{-1})).\tau}. \end{aligned}$$

For convenience we set $\mu(0) := 0$ as well as $h_0 := 0$.

If $a \in U^*$, then the map $\mu(a)$ is the unique permutation in $U_0^*\pi_a U_0^*$ of S that interchanges 0 and ∞ (see [3, Lemma 3.3(2)]). The maps h_a are called *Hua maps* and a short calculation shows that they fix both 0 and ∞.

We have the following criterion due to De Medts and Weiss:

Theorem 2.45. *([5, Theorem 3.2]) $\mathbb{M}(U, \tau)$ is a Moufang set if and only if the restriction of each Hua map to U is an automorphism of U.*

If $\mathbb{M} = \mathbb{M}(U, \tau)$ is a Moufang set, we define the *Hua subgroup* of $\mathbb{M}$ by

$$H(\mathbb{M}) := \langle \mu(a)\mu(b) \mid a, b \in U^* \rangle$$

(cf. [3, Definition 4.2.1]). Note that $H(\mathbb{M}) \leq \mathrm{St}_{G^\dagger}(\{0, \infty\})$. Actually, equality holds by [5, Theorem 3.1(ii)], so that the Hua maps are really contained in the Hua subgroup.

There is a last to be defined property for a Moufang set and a result by Segev and Weiss about those objects, before we begin the pending example:

Definition 2.46. ([20, Definition 1·1]) A Moufang set $\mathbb{M}(U, \tau)$ is called *special* if $(-a).\tau = -(a.\tau)$ for all $a \in U^*$.

Theorem 2.47. *([20, Theorem 1·2]) Let $\mathbb{M}(U, \tau)$ be a special Moufang set with Hua subgroup H. Let $W \leq U$ be non-trivial and H-invariant. Then U is an elementary abelian 2-group or $W = U$.*

Note that the authors write 'either U is an elementary abelian 2-group, or $W = U$'. But both situations can happen simultaneously as the example below will show for $\mathbb{K} = \mathbb{F}_2$.

We give the example of the so-called *projective* Moufang set over a field $\mathbb{K}$. Its name originates from the underlying set which can be identified with the projective line over $\mathbb{K}$ (see the cited example hereafter).

Example 2.48. ([4, Section 5.1]) Let $\mathbb{K}$ be an arbitrary field and let U denote the additive group $(\mathbb{K}, +)$. We define $\infty := 0^{-1}$ (as well as $\infty^{-1} = 0$) and $S := \mathbb{K} \cup \{\infty\}$. We further use the convention $a + \infty = a \cdot \infty = -\infty = \infty$ for all $a \in U$. We can then extend the involution $\tau : U^* \to U^* : x \mapsto -x^{-1}$ to a involution in $\mathrm{Sym}(S)$ also denoted by τ. We obtain a pair $\mathbb{M}(U, \tau) = (S, (U_s)_{s\in S})$ as seen in Construction 2.41.

Let $a \in U^*$ and $s \in S$. First, we calculate

$$\begin{aligned} s.\pi^{-1}_{a.\tau^{-1}} &= s - a.\tau^{-1} = s - \left(-a^{-1}\right) = s + a^{-1} \quad \text{and} \\ s.\pi^{-1}_{(-(a.\tau^{-1})).\tau} &= s - \left(-\left(a.\tau^{-1}\right)\right).\tau = s - \left(-\left(-a^{-1}\right)\right).\tau \\ &= s - a^{-1}.\tau = s - (-a) = s + a. \end{aligned}$$

Hence, the Hua maps are given by

$$\begin{aligned} s.h_a &= s.\tau\pi_a\tau^{-1}\pi^{-1}_{a.\tau^{-1}}\tau\pi^{-1}_{(-(a.\tau^{-1})).\tau} \\ &= \left(-s^{-1}\right).\pi_a\tau^{-1}\pi^{-1}_{a.\tau^{-1}}\tau\pi^{-1}_{(-(a.\tau^{-1})).\tau} \\ &= \left(-s^{-1} + a\right).\tau^{-1}\pi^{-1}_{a.\tau^{-1}}\tau\pi^{-1}_{(-(a.\tau^{-1})).\tau} \\ &= \left(-\left(-s^{-1} + a\right)^{-1}\right).\pi^{-1}_{a.\tau^{-1}}\tau\pi^{-1}_{(-(a.\tau^{-1})).\tau} \\ &= \left(-\left(-s^{-1} + a\right)^{-1} + a^{-1}\right).\tau\pi^{-1}_{(-(a.\tau^{-1})).\tau} \\ &= \left(-\left(-\left(-s^{-1} + a\right)^{-1} + a^{-1}\right)^{-1}\right).\pi^{-1}_{(-(a.\tau^{-1})).\tau} \\ &= -\left(-\left(-s^{-1} + a\right)^{-1} + a^{-1}\right)^{-1} + a \\ &= a - \left(a^{-1} - \left(a - s^{-1}\right)^{-1}\right)^{-1}. \end{aligned}$$

The latter equals $a^2 s$ when applying the *Hua identity*; this identity can be found, for example, in the article '*Jordan algebras and their applications*' by Kevin McCrimmon ([11, Expression (3.3)]) where b needs to be replaced by $-b^{-1}$ to obtain after some rearrangements the desired equality.
Since each Hua map fixes ∞, their restriction to U is a bijection and by

$$(u + u').h_a = a^2(u + u') = a^2 u + a^2 u' = u.h_a + u'.h_a$$

for all $u, u' \in U$, we see that it is an automorphism of U. Thus, Theorem 2.45 implies that $\mathbb{M}(U, \tau)$ is a Moufang set. As explicitly shown in the cited source, its little projective group is the projective special linear group $\mathrm{PSL}_2(\mathbb{K}) = \mathrm{SL}_2(\mathbb{K})/Z(\mathrm{SL}_2(\mathbb{K}))$.

This Moufang set is special as

$$(-a).\tau = -(-a)^{-1} = -(-a^{-1}) = -(a.\tau)$$

holds for all $a \in U^*$. Now, the Hua subgroup H of the special Moufang set $\mathbb{M}(U, \tau)$ acts irreducibly on U by Theorem 2.47, unless $\mathbb{K}$ is of characteristic 2 and the subset $\{k^2 \mid k \in \mathbb{K}\}$ of all squares forms a non-trivial, proper, and H-invariant subfield of $\mathbb{K}$, and thus a non-trivial, proper, and H-invariant subgroup of the elementary abelian 2-group U. For example, this is the case for the imperfect field $\mathbb{F}_2(t)$ of rational functions.

Chapter 3

$\mathbb{Z}$-systems and shift-invariant subgroups

3.1 Further definitions

We introduce some notions and definitions regarding an arbitrary $\mathbb{Z}$-system $\Xi = (X, (X_k)_{k\in\mathbb{Z}}, \varsigma, T)$. First, we define shift-invariance of a subgroup as seen in [9, Definition 6.1]:

Definition 3.1. A subgroup $Y \leq X$ is called *shift-invariant*, if $\varsigma(Y) = Y$.

Remark 3.2. A shift-invariant subgroup Y is just S-invariant for $S := \langle \varsigma \rangle$.

Let Y be a shift-invariant subgroup. For $k \in \mathbb{Z}$ or $m < n \in \mathbb{Z} \cup \{-\infty, \infty\}$ we define $Y_k := Y \cap X_k$ and $Y_{m,n} := Y \cap X_{m,n}$, respectively.

The subgroups Y_k and $Y_{m,n}$ satisfy $\varsigma(Y_k) = Y_{k+2}$ and $\varsigma(Y_{m,n}) = Y_{m+2,n+2}$ by (M4). If we use (M3), then we observe $[Y_m, Y_n] \leq Y_{m+1,n-1}$ for $m < n$.

A short induction shows that $\varsigma^k(Y) = Y$ for all $k \in \mathbb{Z}$.

Remark 3.3. Let $k \in \mathbb{N}$ and $Y \leq X$ be a shift-invariant subgroup. Then $\delta_k(Y)$ and $\zeta_k(Y)$ are shift-invariant again. Furthermore, if $Z \trianglelefteq X$ is another shift-invariant subgroup with $Z \trianglelefteq Y$, then Y/Z is shift-invariant with respect to the automorphism induced on the group X/Z.
Since $\delta_k(Y)$ and $\zeta_k(Y)$ are characteristic and $\varsigma \in \mathrm{Aut}(X)$, this is clear in the first two cases. For Y/Z observe that $\varsigma(yZ) = \varsigma(y)Z \in Y/Z$ for all $y \in Y$.

Recall that we denote a T-invariant subgroup H of X by $H \leq_T X$ according to Definition 1.9. Since T acts via automorphisms on X (from the left), similar to the previous remark we see that if $H \trianglelefteq_T G \leq_T X$, then

G/H is T-invariant, again. We give some further terminology regarding T-groups.

Definition 3.4. For any $M \subseteq X$ we define the *subgroup T-generated by M* as the intersection of all T-subgroups that contain M and write $\langle M \rangle_T$. In short we may call this the *T-span* of M. Equivalently, it is the smallest T-invariant subgroup containing M. Note that, if M is a T-group, then $\langle M \rangle_T = M$.
We say that a subgroup $H \leq_T X$ is *T-generated* by a subset M of X, if $H = \langle M \rangle_T$. If the generating set M is finite, we say that H is *finitely T-generated*. An element $m \in M$ is called a *T-generator* of H.

We say that X is *T-locally nilpotent*, if every finitely T-generated subgroup of X is nilpotent.

If $N \trianglelefteq X$ is T-invariant and X/N is finitely T-generated, we say that N has or is of *finite T-index* in X. If the quotient is not finitely T-generated, we say N is of *infinite T-index*.

Remark 3.5. If $T = \mathrm{Inn}(X)$ is the subgroup of $\mathrm{Aut}(X)$ consisting of all inner automorphisms of X, i.e. automorphisms c_g, for $g \in X$, defined by $c_g(x) = x^g$ for all $x \in X$, then $\langle M \rangle_T = \langle M \rangle^X$ is the normal closure of M in X (cf. Definition 1.5).

Remark 3.6. Since the rooted groups are non-trivial and T-invariant, each of them is T-generated by at least one element. Hence, the group $X = \langle X_k \mid k \in \mathbb{Z} \rangle$ is not finitely T-generated.

In fact, the subgroup T-generated by M and the subgroup generated by the T-orbit of M coincide:

Lemma 3.7. *Let M be a subset of X. Then $\langle M \rangle_T = \langle T(M) \rangle$.*

Proof. The group on the right-hand side clearly contains M. An element g of $\langle T(M) \rangle$ is a product $m_1 m_2 \cdot \ldots \cdot m_k$ with $k \in \mathbb{N}$ and $m_i \in T(M) \cup T(M)^{-1}$ for all $1 \leqslant i \leqslant k$. Applying $t \in T$ to g and using the fact $t.m_i \in T(M) \cup T(M)^{-1}$ for all $1 \leqslant i \leqslant k$, we infer $t.g \in \langle T(M) \rangle$. Hence, the subgroup $\langle T(M) \rangle$ is T-invariant and the inclusion $\langle M \rangle_T \subseteq \langle T(M) \rangle$ follows.

Note for the reverse inclusion that $T(M)$ is contained in the subgroup $\langle M \rangle_T$ by its T-invariance, so that $\langle T(M) \rangle \subseteq \langle M \rangle_T$. $\square$

Remark 3.8. Let Y be a shift- and T-invariant subgroup of X. It follows that for all $k \in \mathbb{Z}$ and $m \leqslant n \in \mathbb{Z}$ the groups X_k and $X_{m,n}$, and therefore Y_k and $Y_{m,n}$ are T-invariant. In particular, if the $\mathbb{Z}$-system is irreducible, then the subgroups Y_k with k even (resp. odd) are either trivial or equal to X_k.

Remark 3.9. Let A and B be two T-invariant subgroups of X. Then the normal closure $\langle A\rangle^X$ of A in X as well as the group $[A, B]$ generated by all commutators $[a, b]$ with $a \in A$ and $b \in B$ are T-invariant subgroups.
The second statement is clear, because $t.[a, b] = [t.a, t.b] \in [A, B]$ for all $t \in T$, $a \in A$, and $b \in B$. For the first statement, let $t \in T$, $a \in A$, and $x \in X$. Since T acts by automorphisms, there exists a unique $z \in X$ such that $t.x = z$, so that

$$t.(a^x) = (t.x)^{-1}(t.a)(t.x) = (t.a)^z \in A^z \subseteq \langle A\rangle^X.$$

In particular, the characteristic subgroups $\delta(X_k)$ as well as $\delta_i(X)$ and $\gamma_i(X)$ are T-invariant for all $k \in \mathbb{Z}$ and $i \in \mathbb{N}$.

We have the following observation that we will use frequently:

Lemma 3.10. *Let $\Xi = (X, (X_k)_{k\in\mathbb{Z}}, \varsigma, T)$ be an irreducible and nilpotent $\mathbb{Z}$-system. Then the rooted groups are abelian.*

Proof. Since Ξ is irreducible, the previous remark implies that the commutator groups $\delta(X_k)$ are either trivial or equal to X_k, i.e. the rooted groups are either abelian or perfect. By nilpotency, the subgroups X_k cannot be perfect, and thus are abelian. $\square$

Let G be a group. An element $g \in G$ is called *divisible* by a positive integer j if $g = h^j$ for some $h \in G$. We call g *uniquely j-divisible* if there is a unique element h with $g = h^j$.
A group G is called *divisible* if each $g \in G$ is divisible by every positive integer and *uniquely divisible* if each $g \in G$ is uniquely j-divisible for all $j > 0$.

Remark 3.11. In fact, the rooted groups of an irreducible and nilpotent $\mathbb{Z}$-system are vector spaces over prime fields.
Indeed, let A be an abelian group and j be a positive integer. First note that the set $O_j = \{a \in A \mid o(a) \text{ divides } j\} = \{a \in A \mid a^j = 1\}$ forms a characteristic subgroup, since

$$(ab^{-1})^j = a^j(b^j)^{-1} = 1^j = 1 = \phi(1) = \phi(a^j) = \phi(a)^j$$

for all $a, b \in O_j$ and $\phi \in \mathrm{Aut}(A)$. Hence such groups are T-invariant subgroups of X_k. If X_k is torsion, then by irreducibility the orders of all non-trivial elements divide each $j > 1$, which can only be the case if $X_k = O_p$ for some prime p and $O_j = \{1\}$ for all $j \neq p$, so that X_k is an elementary abelian p-group (see [14, 4.1.1]).

Furthermore, the subsets $H_j = \{a^j \mid a \in A\}$ of A are characteristic subgroups as well. Hence, if X_k is torsion-free, then X_k is divisible by irreducibility and the elements $h \in H_j$ (for j fixed) are uniquely determined, i.e. each $x \in X_k$ is uniquely j-divisible for all $j > 0$ and X_k is a vector spaces over $\mathbb{Q}$ as a uniquely divisible group by [14, 4.1.5].

3.2 Normal form and commutator relations

Throughout this section, let $\Xi = (X, (X_k)_{k\in\mathbb{Z}}, \varsigma, T)$ be an arbitrary $\mathbb{Z}$-system. We start by collecting a few basic properties. The following result is an analogue to the statement (ZS6) of Lemma 4.2 in [9] (see also [19, Lemma 2.1]).

Lemma 3.12. *Let $x \in X^*$. Then there exist $m \leqslant n \in \mathbb{Z}$, $x_m \in X_m^*$, $x_n \in X_n^*$ and for $m < k < n$ elements $x_k \in X_k$ such that*

$$x = x_m x_{m+1} \cdot \ldots \cdot x_n.$$

Moreover, the integers m and n as well as the elements $x_m, x_{m+1}, \ldots, x_n$ are uniquely determined by x.

Proof. Let $x \in X^*$. By assertion (M1), there exist finitely many indices $k_1, \ldots, k_r$ such that x is a product of elements from $X_{k_i}^*$ for $1 \leqslant i \leqslant r$. Set $m := \min\{k_1, \ldots, k_r\}$ and $n := \max\{k_1, \ldots, k_r\}$, so that $x \in X_{m,n}^*$. Since (at least) one factor in the product above is contained in X_m^* resp. X_n^*, we have $x \notin X_{m+1,n}$ resp. $x \notin X_{m,n-1}$. Using property (M2), there are $x_k \in X_k$ for all $m \leqslant k \leqslant n$ such that $x = \prod_{k=m}^{n} x_k$, and the factors x_k in this decomposition are uniquely determined. We infer $x_m \neq 1 \neq x_n$ by our observation. If there exist $m \neq m'$ or $n \neq n'$ such that $x \in X_{m',n'}$ has a decomposition $x = \prod_{k=m'}^{n'} y_k$ with $y_{m'} \neq 1 \neq y_{n'}$, then $x \in X_{\min\{m,m'\},\max\{n,n'\}}$ has two distinct decompositions, contradicting (M2), since we observe $x_l \neq 1 = y_l$ or $x_{l'} = 1 \neq y_{l'}$ for $l = m$ or $l = n$. Hence, the indices m and n are also uniquely determined. □

Hence, the following definition is valid (similar to [9, Definition 4.3]):

Definition 3.13. If $x \in X^*$ can be uniquely written as $x = x_m x_{m+1} \cdot \ldots \cdot x_n$ with unique integers $m \leqslant n$ like in the previous lemma, then we call the decomposition $x_m x_{m+1} \cdot \ldots \cdot x_n$ the *normal form* of x and set $\mu(x) := m$ as well as $\nu(x) := n$. The *width* of x is defined by $\omega(x) := \nu(x) - \mu(x) + 1$.

For later purpose, we set $\mu(1) := \infty$, $\nu(1) := -\infty$, and $\omega(1) := 0$. Further, let $\mu(U) := \{\mu(u) \mid u \in U\}$ and $\nu(U) := \{\nu(u) \mid u \in U\}$ for $U \subseteq X^*$.

We sometimes use the notation $x = x_m x_{m+1} \cdot \ldots \cdot x_n$ for $x \in X$ where the decomposition is not necessarily the normal form of x, i.e. $\omega(x) < n-m+1$. But in cases of being the normal form we will state it clearly; for example, by using $\mu(x)$ or $\nu(x)$ instead of m or n, respectively, to indicate the first or last letter of the normal form.

Also note that if $x \in X$ has a certain width, then any shift $\varsigma^k(x)$, for $k \in \mathbb{Z}$, has the same width, since ς is an automorphism.

The third property of a $\mathbb{Z}$-system has as a direct consequence:

Lemma 3.14. *([19, Lemma 2.3]) Let $m < n \in \mathbb{Z}$ and let $x_i \in X_i^*$ for $i = m, n$. Then we have*

$$x_n x_m = x_m x_n [x_m, x_n]^{-1} \in x_m x_n X_{m+1,n-1}.$$

In particular, if $n - m = 1$, then $[x_n, x_m] = 1$ and X_m commutes with X_n.

Proof. The claim follows by the fact that

$$x_n x_m = x_m x_n x_n^{-1} x_m^{-1} x_n x_m = x_m x_n [x_n, x_m] = x_m x_n [x_m, x_n]^{-1}$$

and that the commutator is contained in $X_{m+1,n-1}$ by (M3). □

Recall Lemma 1.12 which shows how the commutator of products is computed. Together with (M3) it gives rise to useful commutator relations between two subgroups $X_{k,l}$ and $X_{m,n}$ with $k \leqslant l$ and $m \leqslant n$. The first and separately stated observation of the following lemma is due to Scholz (see [19, Lemma 2.2]).

Lemma 3.15. *Let $m, n \in \mathbb{Z}$ with $m \leqslant n$. Then $[X_{m,n}, X_{n+1}] \leq X_{m+1,n}$. More generally, we have*

$$[X_{m,n}, X_l] \leq X_{m+1,l-1}$$

for all integers $l \geqslant n + 1$.

Proof. Since the first statement is the general case for $l = n+1$, we perform an induction on $L := l - n \geqslant 1$.

For the base step $L = 1$, we argue by an induction on $N := n - m \in \mathbb{N}$. If $N = 0$, then by Lemma 3.14 and (M3) we have

$$\{1\} = [X_n, X_{n+1}] = [X_{m,n}, X_{n+1}] = [X_m, X_{n+1}] \leq X_{m+1,n}.$$

Note that $X_{m+1,n} = \langle \varnothing \rangle = \{1\}$, so we observe equality here.

Now, assume that the induction hypothesis holds for $N > 0$. Let $m \leqslant n \in \mathbb{Z}$ such that $N + 1 = n - m$ and take $x \in X_{m,n}$ and $y \in X_{n+1}$. If one of these elements is trivial, then $[x, y]$ is trivial and obviously contained in $X_{m+1,n}$.
So, let x and y be non-trivial. By Lemma 3.12 there exists $x_k \in X_k$ for $m \leqslant k \leqslant n$ such that $x = x_m \cdot ... \cdot x_n$ (not necessarily the normal form of x). We compute, using the first equation in Lemma 1.12,

$$[x, y] = [x_m \cdot ... \cdot x_n, y] = [x_m, y]^{x_{m+1}\cdot ... \cdot x_n}[x_{m+1} \cdot ... \cdot x_n, y],$$

where the first factor is contained in $X_{m+1,n}$ by (M3) and the second factor is an element of $X_{m+2,n}$ by our hypothesis as $n - (m + 1) = N$. Thus, we conclude $[x, y] \in X_{m+1,n}X_{m+2,n} \subseteq X_{m+1,n}$ which proves the first claim, since those elements generate $[X_{m,n}, X_l]$.

For the induction step, let $L = l - n$ be greater than 1 and assume that $[X_{m',n'}, X_{l'}] \leq X_{m'+1,l'-1}$ holds for all $m' \leqslant n' \in \mathbb{Z}$ and all integers l' with $1 \leqslant l' - n' < L$. It remains to show that $[X_{m,n}, X_l] \leq X_{m+1,l-1}$ for all $m \leqslant n$. But $X_{m,n} \subseteq X_{m,n+1}$ for all $m \leqslant n$, so we directly infer

$$[X_{m,n}, X_l] \subseteq [X_{m,n+1}, X_l] \leq X_{m+1,l-1}$$

by our induction hypothesis, since $l - (n + 1) < L$. □

Remark 3.16. It follows from the lemma above that $X_{n+1} \subseteq N_X(X_{m,n})$ for all $m \leqslant n \in \mathbb{Z}$, i.e. X_{n+1} normalizes $X_{m,n}$. Indeed, let $x \in X_{m,n}$ and $y \in X_{n+1}$. Then

$$x^y = x[x, y] \in xX_{m+1,n} \subseteq X_{m,n}X_{m+1,n} \subseteq X_{m,n},$$

which implies $X^y_{m,n} \subseteq X_{m,n}$. Similarly, we have $X^{y^{-1}}_{m,n} \subseteq X_{m,n}$ which is equivalent to $X_{m,n} \subseteq X^y_{m,n}$. As equality holds, we conclude $y \in N_X(X_{m,n})$.

To put it in other words (while using the injectivity of $\rho_{m,n+1}$ from (M2) to observe $X_{n+1} \cap X_{m,n} = \{1\}$), the group $X_{m,n+1}$ is the *(internal) semidirect product* of $X_{m,n}$ and X_{n+1}, written $X_{m,n} \rtimes X_{n+1}$.

With this in mind, the statement of Lemma 3.15 can almost be extended to include all edge cases $n = l$: If $x \in X_{m,l}$ has normal form $x_m \cdot ... \cdot x_l$ and $y \in X_l$, then by applying Lemma 1.12 repeatedly and using (M3) as well as the fact that $X_l \subseteq N_X(X_{k,l-1})$ for all $k \leqslant l - 1$, we observe

$$\begin{aligned}[x, y] &= [x_m \cdot ... \cdot x_l, y] = [x_m, y]^{x_{m+1}\cdot ... \cdot x_l}[x_{m+1} \cdot ... \cdot x_l, y] = \dots \\ &= \underbrace{[x_m, y]^{x_{m+1}\cdot ... \cdot x_l}}_{\in X^{x_{m+1}\cdot ... \cdot x_l}_{m+1,l-1} = X_{m+1,l-1}} \cdot \underbrace{[x_{m+1}, y]^{x_{m+2}\cdot ... \cdot x_l}}_{\in X^{x_{m+2}\cdot ... \cdot x_l}_{m+2,l-1} = X_{m+2,l-1}} \cdot ... \cdot \underbrace{[x_{l-1}, y]^{x_l}}_{=1} \cdot \underbrace{[x_l, y]}_{\in \delta(X_l)}.\end{aligned}$$

So, either $m = n = l$ and $[x, y] \in \delta(X_l)$, which is no additional information, or $m < n = l$ and $[x, y] \in X_{m+1,l-1}\delta(X_l) \subseteq X_{m+1,l}$. Thus the lemma above is also true for $m \leqslant n = l$ in case of $\delta(X_l) = \{1\}$:

Corollary 3.17. *Let $m \leqslant n \leqslant l$ be integers. If the rooted groups are abelian, then*

$$[X_{m,n}, X_l] \leq X_{m+1,l-1}.$$

There is a symmetric result for lower indexed rooted groups:

Lemma 3.18. *Let $m \leqslant n \in \mathbb{Z}$. Then $[X_{m-1}, X_{m,n}] \leq X_{m,n-1}$. Again, we more generally have*

$$[X_k, X_{m,n}] \leq X_{k+1,n-1}$$

for all integers $k \leqslant m - 1$.

Proof. We could similarly argue as seen in Lemma 3.15 via an induction on the distance of indices $K := m - k \geqslant 1$. But it is a more elegant solution to look at the $\mathbb{Z}$-system

$$\Xi^{inv} = \left(X, \left(X_k^{inv}\right)_{k\in\mathbb{Z}}, \varsigma^{-1}, T\right)$$

defined in Remark 2.34. Let $m \leqslant n$. We apply Lemma 3.15 to see that

$$\left[X_{m,n}^{inv}, X_l^{inv}\right] \leq X_{m+1,l-1}^{inv}$$

for all $l \geqslant n + 1$. For $k = -l, m' = -n$ and $n' = -m$, this is equivalent to

$$[X_k, X_{m',n'}] = [X_{-n,-m}, X_{-l}] \leq X_{-(l-1),-(m+1)} = X_{k+1,n'-1}$$

for all $m' - 1 \geqslant k$. This shows the claimed relation for Ξ. $\square$

Remark 3.19. In analogy to Remark 3.16 above, by using Lemma 3.18, we see that $X_{m-1,n}$ is the semidirect product $X_{m-1} \ltimes X_{m,n}$ of $X_{m,n}$ and X_{m-1} for all $m \leqslant n$.

With Lemma 3.18 and the fact that $\delta(X_k)$ and X_n normalize $X_{k+1,n-1}$, but essentially similar to the discussion of the edge cases above, either $k = m < n$ and we infer

$$\begin{aligned}[y, x] &= [y, x_{k+1} \cdot ... \cdot x_n][y, x_k]^{x_{k+1}\cdot...\cdot x_n} = [y, x_{k+1} \cdot ... \cdot x_n]x_n^{-1}[y, x_k]zx_n \\ &= [y, x_{k+1} \cdot ... \cdot x_n][y, x_k](z')^{x_n} = [y, x_{k+1} \cdot ... \cdot x_n][y, x_k]z'' \\ &\in X_{k+1,n-1}\delta(X_k)X_{k+1,n-1} = \delta(X_k)X_{k+1,n-1} \subseteq X_{k,n-1}\end{aligned}$$

for $x \in X_{k,n}$, $x_i \in X_i$ for $k \leqslant i \leqslant n$ such that $x = x_k \cdot ... \cdot x_n$, some elements $z, z', z'' \in X_{k+1,n-1}$, and $y \in X_k$, or $k = m = n$ and we only see $[y, x] \in \delta(X_k)$.

In the abelian case, we have a corollary analogue to the last one:

Corollary 3.20. *Let $k \leqslant m \leqslant n$ be integers. If the rooted groups are abelian, then*

$$[X_k, X_{m,n}] \leq X_{k+1,n-1}.$$

Whether our rooted groups are abelian or not, we have a third

Lemma 3.21. *Let $b, m, n \in \mathbb{Z}$ such that $m < b < n$. Then*

$$[X_b, X_{m,n}] \leq X_{m+1,n-1}.$$

Proof. Let $y \in X_b$ and $x = x_m \cdot \ldots \cdot x_n \in X_{m,n}$. With Lemma 1.12 we can factorize $[y, x]$ to

$$[y, x_n][y, x_{n-1}]^{x_n} \cdot \ldots \cdot [y, x_b]^{x_{b+1} \cdot \ldots \cdot x_n} \cdot \ldots \cdot [y, x_{m+1}]^{x_{m+2} \cdot \ldots \cdot x_n}[y, x_m]^{x_{m+1} \cdot \ldots \cdot x_n}.$$

Since $[g, h]^{-1} = [h, g]$ for all $g, h \in X$, we infer that each commutator is contained in $X_{b+1,n-1}$, X_b or $X_{m+1,b-1}$ by property (M3). Hence, they are all elements of $X_{m+1,n-1}$ by our choice of b. As X_n normalizes $X_{m+1,n-1}$ by Remark 3.16, the conjugates are contained in $X_{m+1,n-1}$ as well. Thus, the claimed relation follows. □

Combining the last three lemmata, we derive an essential and very useful fact that we utilize quite often in the remainder:

Proposition 3.22. *Let $k \leqslant l$ and $m \leqslant n$ be integers. Then*

(i) $[X_{k,l}, X_{m,n}] \leq X_{\min\{k,m\}+1,\max\{l,n\}-1}$ *if $k \neq m$ and $l \neq n$.*

(ii) $[X_{k,l}, X_{m,n}] \leq \delta(X_k)X_{k+1,\max\{l,n\}-1}$ *if $k = m$ and $l \neq n$.*

(iii) $[X_{k,l}, X_{m,n}] \leq X_{\min\{k,m\}+1,n-1}\delta(X_n)$ *if $k \neq m$ and $l = n$.*

(iv) $[X_{k,l}, X_{m,n}] \leq \delta(X_k)X_{k+1,n-1}\delta(X_n)$ *if $k = m$ and $l = n$.*

Proof. Let $x = x_k \cdot \ldots \cdot x_l \in X_{k,l}$ and $y = y_m \cdot \ldots \cdot y_n \in X_{m,n}$ be two elements (not necessarily in normal form). Up to relabelling the indices and switching the entries of the commutator, we may assume $l \leqslant n$ to significantly reduce the number of cases. We did manually check the sixteen cases subdivided into three distinct classes (four cases if $l < m$, also four more cases if $l = m$, and eight cases if $m < l$) to obtain the result beforehand, but we give another proof here by distinguishing the cases given in the proposition.

First, by applying Lemma 1.12 repeatedly as seen before, we get a decomposition

$$[x, y] = [x_k \cdot \ldots \cdot x_l, y] = [x_k, y]^{x_{k+1} \cdot \ldots \cdot x_l}[x_{k+1}, y]^{x_{k+2} \cdot \ldots \cdot x_l} \cdot \ldots \cdot [x_{l-1}, y]^{x_l}[x_l, y].$$

We start with the case $k = m$ and $l = n$, i.e. we consider $[X_{k,n}, X_{k,n}]$. If $k = n$, then we have $[X_{k,n}, X_{k,n}] = [X_n, X_n] = \delta(X_n)$. If $k = m < n$, then

$$\begin{aligned}[x_k, y]^{x_{k+1}\cdot\ldots\cdot x_n} &\in (\delta(X_k)X_{k+1,n-1})^{x_{k+1}\cdot\ldots\cdot x_n} = \delta(X_k)^{x_{k+1}\cdot\ldots\cdot x_n}X_{k+1,n-1} \\ &\subseteq \delta(X_k)X_{k+1,n-1}X_{k+1,n-1} = \delta(X_k)X_{k+1,n-1} \subseteq X_{k,n-1}\end{aligned}$$

by Remark 3.19, $[x_{k+i}, y]^{x_{k+i+1}\cdot\ldots\cdot x_n} \in X_{k+1,n-1}^{x_{k+i+1}\cdot\ldots\cdot x_n} = X_{k+1,n-1}$ for all $1 \leqslant i \leqslant l-k-1$ by Lemma 3.21, and $[x_n, y] \in X_{k+1,n-1}\delta(X_n)$ by Remark 3.16. Thus, $[x, y] \in \delta(X_k)X_{k+1,n-1}\delta(X_n)$; but by definition we already know that $[x, y] \in \delta(X_{k,n})$. Hence we cannot derive further information about the commutator in general if $k = m$ and $l = n$.

For (iii), assume $k \neq m$ and $l = n$; thus we look at $[X_{k,n}, X_{m,n}]$.

We first deal with the subcase $\min\{k, m\} = m$. If $k = n$, then we have $[X_{k,n}, X_{m,n}] = [X_n, X_{m,n}] \leq X_{m+1,n-1}\delta(X_n)$ by the observation of Remark 3.16. If $k \neq n$, and thus $k < n$ by $k \leqslant l = n$, then $m < k + i < n$ for all $0 \leqslant i \leqslant l-k-1$ and we infer $[x_{k+i}, y]^{x_{k+i+1}\cdot\ldots\cdot x_n} \in X_{m+1,n-1}^{x_{k+i+1}\cdot\ldots\cdot x_n} = X_{m+1,n-1}$ for all $0 \leqslant i \leqslant l-k-1$ by Lemma 3.21 while $[x_n, y_n] \in \delta(X_n)$, so that $[x, y] \in X_{m+1,n-1}\delta(X_n)$ and $[X_{k,n}, X_{m,n}] \leq X_{m+1,n-1}\delta(X_n)$.
Now, let $\min\{k, m\} = k$. If $m = n$, then $[X_{k,n}, X_{m,n}] = [X_{k,n}, X_n] \leq X_{k+1,n-1}\delta(X_n)$ as above. If $m < n$, then $[X_{k,n}, X_{m,n}] \leq X_{k+1,n-1}\delta(X_n)$ follows also like the second case of the first subcase.

For the second statement, let $k = m$ and $l \neq n$, i.e. $\max\{l, n\} = n$ by our pending assumption that $l \leqslant n$; thus we look at $[X_{k,l}, X_{k,n}]$ with $l < n$.

If $l = m$, then $[X_{k,l}, X_{k,n}] = [X_k, X_{k,n}] \leq \delta(X_k)X_{k+1,n-1}$ by the previous remark. Else $l > m$ and $m < k + i < n$ for all $1 \leqslant i \leqslant l - k$. Hence

$$[x_{k+i}, y]^{x_{k+i+1}\cdot\ldots\cdot x_l} \in X_{m+1,n-1}^{x_{k+i+1}\cdot\ldots\cdot x_l} = X_{m+1,n-1} = X_{k+1,n-1}$$

for all $1 \leqslant i \leqslant l - k$, again by Lemma 3.21. Since $[x_k, y] \in \delta(X_k)X_{k+1,n-1}$ as seen a few lines ago, we have $[x, y] \in \delta(X_k)X_{k+1,n-1}$ and $[X_{k,l}, X_{k,n}] \leq \delta(X_k)X_{k+1,n-1}$.

At last, we tend to (i) where we assume that $k \neq m$ and $l \neq n$ what is assumed to be equivalent to $k \neq m$ and $l < n$. There are eight more cases to look at in which always $\max\{l, n\} = n$ holds.

If $l = m$, then $k \neq l$ as well as $m \neq n$ and there is only the case $[X_{k,l}, X_{l,n}]$ with $k < l < n$ to consider. As $k + i < l$ for all $0 \leqslant i \leqslant l - k - 1$, we infer $[x_{k+i}, y]^{x_{k+i+1}\cdot\ldots\cdot x_n} \in X_{k+1,n-1}^{x_{k+i+1}\cdot\ldots\cdot x_n} = X_{k+1,n-1}$ for all $0 \leqslant i \leqslant l-k-1$ by Lemma 3.18. Further, we have $[x_l, y] \in X_{l,n-1}$ by Remark 3.19. Together, this implies $[x, y] \in X_{k+1,n-1}$ and $[X_{k,l}, X_{l,n}] \leq X_{k+1,n-1} = X_{\min\{k,m\}+1,\max\{l,n\}-1}$.

For the remainder of this proof, let $l \neq m$.

The first subcase we tend to is $\min\{l, m\} = l$, i.e. $k \leqslant l < m \leqslant n$.

If $k = l$, then there are the cases (a) $m = n$ and the case (b) $m < n$. In case (a), property (M3) of the definition of a $\mathbb{Z}$-system directly implies $[X_{k,l}, X_{m,n}] = [X_k, X_n] \leqslant X_{k+1,n-1}$. Since $k < m$, we have $[X_{k,l}, X_{m,n}] = [X_k, X_{m,n}] \leq X_{k+1,n-1}$ in case (b) by Lemma 3.18.
If $k < l$, we again have to consider the cases (a') $m = n$ and (b') $m < n$. In the former case, we apply Lemma 3.15 to see $[X_{k,l}, X_{m,n}] = [X_{k,l}, X_n] \leqslant X_{k+1,n-1}$. Concerning case (b'), we still have $k + i < m$ for all $0 \leqslant i \leqslant l - k$, so that each commutator in the decomposition of $[x, y]$ above is contained in $X_{k+i+1,n-1}$ by Lemma 3.18. This is also true for each factor, so that $[x, y] \in X_{k+1,n-1}$. In summary, we have $[X_{k,l}, X_{m,n}] \leq X_{k+1,n-1} = X_{\min\{k,m\}+1,\max\{l,n\}-1}$ in this subcase.

We turn now to the second subcase $\min\{l, m\} = m$, i.e. $m < l < n$.

We consider three last cases (c) $m < k = l$, (d) $m < k < l$, and (e) $k < m$. Under (c), we immediately deduce $[X_{k,l}, X_{m,n}] = [X_k, X_{m,n}] \leqslant X_{m+1,n-1}$ by Lemma 3.21. In the second case, we again have $m < k + i < n$ for all $0 \leqslant i \leqslant l - k$. The same lemma implies therefore that the factors of $[x, y]$ as well as the commutator itself are contained in $X_{m+1,n-1}$. For (e) we need to look closer on the individual factors of the decomposition. For $0 \leqslant i \leqslant m - k - 1$ we have $k \leqslant k + i < m$ and the factor $[x_{k+i}, y]^{x_{k+i+1} \cdot \ldots \cdot x_l}$ is an element of $X_{k+i+1,n-1} \subseteq X_{k+1,n-1}$ by Lemma 3.18. If $i = m - k$, then $[x_{k+i}, y] = [x_m, y] \in X_{m,n-1}$ by Remark 3.19; hence, $[x_{k+i}, y]^{x_{k+i+1} \cdot \ldots \cdot x_l} \in X_{m,n-1} \subseteq X_{k+1,n-1}$. For $m - k + 1 \leqslant i \leqslant l - k$, the indices $k + i$ are greater than m (and smaller than n), so that Lemma 3.21 is applicable and $[x_{k+i}, y]^{x_{k+i+1} \cdot \ldots \cdot x_l}$ is an element of $X_{m+1,n-1} \subseteq X_{k+1,n-1}$. Together we infer $[x, y] \in X_{k+1,n-1}$.

Altogether, we have shown that $[X_{k,l}, X_{m,n}] \leqslant X_{\min\{k,m\}+1,\max\{l,n\}-1}$ if $k \neq m$ and $l \neq n$. □

Remark 3.23. Using both remarks above or this proposition directly, we observe $X_{m,n} \trianglelefteq X_{m-1,n+1}$ for all integers $m \leqslant n$.

Combining this result with the observations of the last two corollaries concerning the abelian case, we infer the following

Corollary 3.24. *Let $k \leqslant l$ and $m \leqslant n$ be integers. If the rooted groups are abelian, then*

$$[X_{k,l}, X_{m,n}] \leq X_{\min\{k,m\}+1,\max\{l,n\}-1}.$$

3.3 T-equivariant maps

In the first half of this section we introduce T-equivariant homomorphisms and state some useful implications regarding T-generated groups.

Definition 3.25. A homomorphism $f : G \to H$ between two T-groups is called a *T-homomorphism* if it commutes with the action of T, i.e. $f(t.g) = t.f(g)$ for all $t \in T$ and $g \in G$ (see, for example, [14, p. 29]). We also say that f is *T-equivariant* or that f centralizes T.

Clearly, the composition of T-homomorphisms $f : G \to H$ and $g : H \to K$ is a T-homomorphism, again.

If $f : G \to H$ is an T-equivariant isomorphism, then its inverse $f^{-1} : H \to G$ is also T-invariant. For a quick proof, let $t \in T$, $h \in H$, and $g \in G$ such that $f(g) = h$. Furthermore, let $h' \in H$ with $f(t.g) = h'$. Then

$$t.f^{-1}(h) = t.g = f^{-1}(h') = f^{-1}(f(t.g)) = f^{-1}(t.f(g)) = f^{-1}(t.h).$$

A T-equivariant isomorphism between two T-groups G and H is called a T-isomorphism. We then say that G and H are *isomorphic as T-groups* or short *T-isomorphic* and, in analogue to the T-subgroup symbol, expand the symbol $\cong$ by an index T to $\cong_T$.

Note that the identity map $id_G : G \to G : g \mapsto g$ is T-equivariant. Thus the set of T-automorphisms of G form a subgroup of $\mathrm{Aut}(G)$.

Another easy example of a T-isomorphism is $f : G \to G/\{1\} : g \mapsto \{g\}$. Indeed, this map is an isomorphism and we have $t.f(g) = t.\{g\} = \{t.g\} = f(t.g)$ for all $t \in T$ and $g \in G$.

We state two general facts about T-homomorphisms. The following proposition extends Theorem 1.8 to T-groups and T-isomorphisms and will be used several times in the remainder:

Proposition 3.26.

(i) If $f : G \to H$ is a T-homomorphism, then the map

$$F : G/\ker(f) \to f(G) : g\ker(f) \mapsto f(g)$$

is a T-isomorphism.

(ii) Let $H \leq_T G$ and $N \trianglelefteq_T G$. Then $N \cap H \trianglelefteq_T H$ and

$$\varphi : H \to HN/N : h \mapsto hN$$

is a T-epimorphism with kernel $N \cap H$. Thus, $H/(N \cap H) \cong_T HN/N$.

(iii) Let G be a T-group and $N \leq H$ be normal T-subgroups of G, then

$$\psi : G/N \to G/H : gN \mapsto gH$$

is a T-epimorphism with kernel H/N. Hence, $(G/N)/(H/N) \cong_T G/H$.

Proof. We prove part (i) first and start by showing that $\ker(f) \trianglelefteq G$ as well as $f(G) \leq H$ are T-invariant. Therefore, let $k \in \ker(f)$, $h \in f(G)$ and $g \in G$ with $f(g) = h$, and $t \in T$. Then

$$f(t.k) = t.f(k) = t.1 = 1,$$

as f centralizes T, so that $t.k \in \ker(f)$, the kernel is a T-subgroup of G, and $G/\ker(f)$ is a well-defined T-group, again. Similarly, we have

$$t.h = t.f(g) = f(t.g) \in f(G)$$

by the T-invariance of G and T-equivariance of f. Hence $f(G) \leq_T H$.

The map F is an isomorphism by (I1) of Theorem 1.8. For $g \in G$ and $t \in T$ we compute

$$t.F(g\ker(f)) = t.f(g) = f(t.g) = F((t.g)\ker(f)) = F(t.g\ker(f)).$$

Hence, F is T-equivariant.

Clearly, $N \cap H \trianglelefteq_T H$ as an intersection of T-groups normalized by H. Note that HN is a T-group and that the quotients are well-defined T-groups. By Theorem 1.8 (I2) the epimorphism φ has kernel $N \cap H$. It remains to show that φ is T-equivariant. Therefore, let $h \in H$ and $t \in T$. Then

$$t.\varphi(h(N \cap H)) = t.hN = (t.h)N = \varphi((t.h)(N \cap H)) = \varphi(t.h(N \cap H)).$$

This proves assertion (ii) by applying the first one.

For the last part, we first note that $N \trianglelefteq_T H$. Thus G/N, G/H, and H/N are T-groups. The map $\psi : G/N \to G/H : gN \mapsto gH$ is an epimorphism with kernel H/N by (I3) of Theorem 1.8, so that $(G/N)/(H/N) \cong G/H$. Furthermore, the map ψ centralizes T, i.e. $\psi \circ t = t \circ \psi$ for all $t \in T$. Indeed, we compute

$$\psi(t.gN) = \psi((t.g)N) = (t.g)H = t.gH = t.\psi(gN)$$

for $g \in G$ and $t \in T$. Now, the first part of this proposition implies that the T-epimorphism ψ induces a T-isomorphism from $(G/N)/(H/N)$ to G/H. $\square$

If we choose $N = \{1\}$ in (iii), then we immediately see by $G \cong_T G/\{1\}$ that the canonical projection $\rho : G \to G/H : g \mapsto gH$ is a T-epimorphism.

The lemma hereinafter and its corollaries below will come in handy, especially in Chapter 6.

Lemma 3.27. *Let $f : G \to H$ be a T-epimorphism. If G is T-generated by a subset M, then H is T-generated by $f(M)$. In particular, if G is finitely T-generated, then so is H.*

Proof. Let $h \in H$. By surjectivity there is $g \in G$ with $h = f(g)$. Since $G = \langle M \rangle_T$, the element g can be written as a product $\prod_{i=1}^{k} m_i$ with $k \in \mathbb{N}$ and $m_i \in T(M) \cup T(M)^{-1}$ for all $1 \leqslant i \leqslant k$ by Lemma 3.7. Now the T-equivariance of f implies

$$f\left(T(M) \cup T(M)^{-1}\right) = f\left(T(M)\right) \cup f\left(T(M)^{-1}\right) = T(f(M)) \cup T(f(M))^{-1},$$

so that $f(m_i) \in T(f(M)) \cup T(f(M))^{-1}$ for all i and

$$\begin{aligned} h = f(g) = \prod_{i=1}^{k} f(m_i) &\in \langle T(f(M)) \cup T(f(M))^{-1} \rangle \\ &= \langle T(f(M)) \rangle = \langle f(M) \rangle_T. \end{aligned}$$

Since H is a T-group containing $f(M)$, the inclusion $\langle f(M) \rangle_T \subseteq H$ is obvious. We infer $H = \langle f(M) \rangle_T$. $\square$

In particular, this implies that quotients of finitely T-generated groups are finitely T-generated, again.

Naturally, this lemma can especially be applied to T-isomorphisms. Since the inverse of a T-isomorphism is a T-invariant isomorphism, we obtain the following resulting equivalence:

Corollary 3.28. *Let G and H be T-isomorphic groups. Then G is finitely T-generated if and only if H is finitely T-generated. In particular, both the set of T-generators of G and the set of T-generators of H can be chosen such that they have equal cardinality in the finitely T-generated case.*

Combining Proposition 3.26 (iii) and the lemma above we observe:

Corollary 3.29. *Let G be a T-group, $N \leq H$ two normal T-subgroups of G, and $\psi : G/N \to G/H : gN \mapsto gH$. If G/N is (finitely) T-generated by a subset M, then G/H is (finitely) T-generated by $\psi(M)$.*
Conversely, if H is of infinite T-index, then so is N.

We continue this section with some results about an arbitrary $\mathbb{Z}$-system $\Xi = (X, (X_k)_{k\in\mathbb{Z}}, \varsigma, T)$. Since every $x \in X^*$ possesses a normal form by Lemma 3.12, the maps introduced hereinafter are well-defined. The definition is greatly inspired by [19, Definition 2.4].

Definition 3.30. Let $i \in \mathbb{Z}$ and let $x \in X^*$ with normal form $x_{\mu(x)} \cdot \ldots \cdot x_{\nu(x)}$. We define the *projection map*

$$\pi_i : X \to X_i$$

via $\pi_i(1) := 1$ as well as $\pi_i(x) := 1$ if either $i < \mu(x)$ or $i > \nu(x)$ and $\pi_i(x) := x_i$ if $\mu(x) \leqslant i \leqslant \nu(x)$. Note that these maps are surjective.

Lemma 3.31. *([19, Lemma 2.5]) Let $i \in \mathbb{Z}$ and $m \leqslant n$ be integers. The projection map π_i restricted to $X_{m,n}$ is a homomorphism if $i \leqslant m$ or $i \geqslant n$.*

Proof. Let $i \in \mathbb{Z}$ and $x, y \in X_{m,n}$. If $x = 1$, then $\pi_i(x) = 1$ and $\pi_i(xy) = \pi_i(y) = \pi_i(x)\pi_i(y)$. So let $x \neq 1 \neq y$ in the following.
For normal forms $x = x_{\mu(x)} \cdot \ldots \cdot x_{\nu(x)}$ and $y = y_{\mu(y)} \cdot \ldots \cdot y_{\nu(y)}$ we set $x_k := 1$ for all $m \leqslant k < \mu(x)$ and $\nu(x) < k \leqslant n$ as well as $y_l := 1$ for all $m \leqslant l < \mu(y)$ and $\nu(y) < l \leqslant n$, so that $x = x_m \cdot \ldots \cdot x_n$ and $y = y_m \cdot \ldots \cdot y_n$.
If either $i < m$ or $i > n$, then $\pi_i(x) = 1 = \pi_i(y)$ and $\pi_i(xy) = 1 = \pi_i(x)\pi_i(y)$, since $xy \in X_{m,n}$, i.e. π_i is the trivial homomorphism on $X_{m,n}$.
If $i \in \{m, n\}$, then we observe

$$xy = x_m \cdot \ldots \cdot x_n y_m \cdot \ldots \cdot y_n = x_m y_m x' y_{m+1} \cdot \ldots \cdot y_n$$

for some $x' \in X_{m+1,n}$, since X_m normalizes $X_{m+1,n}$ by Remark 3.19, and

$$xy = x_m \cdot \ldots \cdot x_{n-1} y' x_n y_n$$

for some $y' \in X_{m,n-1}$ with Remark 3.16. Comparing the normal form of xy with these decompositions, we infer $\pi_i(xy) = x_i y_i = \pi_i(x)\pi_i(y)$. ☐

Similarly we see that π_i restricted to $X_{i,\infty}$ resp. $X_{-\infty,i}$ is a homomorphism.

Whenever we say, by abuse of language, that a projection map is a homomorphism, we will always mean a suitable restriction.

We can state some direct consequences for the restricted projection maps. They roughly resemble the ideas of the first two parts of [19, Lemma 3.8].

Corollary 3.32. *Let $U \leq_T X$, $m \leqslant n$ integers, and $i \in \{m, n\}$.*

(i) The projection map $\pi_i : X \to X_i$ restricted to $U \cap X_{m,n}$ is T-equivariant and its image $\pi_i(U \cap X_{m,n})$ is a T-invariant subgroup of X_i.

(ii) If T acts irreducibly on the rooted groups and $\pi_i(U \cap X_{m,n})$ is non-trivial, then $\pi_i(U \cap X_{m,n}) = X_i$ and the restriction of π_i to $U \cap X_{m,n}$ is a T-epimorphisms.

Proof. First note that the intersections $U \cap X_{m,n}$ as well as $U \cap X_i$ are T-subgroups of X. Let $u := u_m \cdot ... \cdot u_n \in U \cap X_{m,n}$ and $t \in T$. Since T acts via automorphisms, we compute

$$t.\pi_i(u) = t.u_i = \pi_i(t.u_m \cdot ... \cdot t.u_n) = \pi_i(t.u) \in \pi_i(U \cap X_{m,n}).$$

Hence, the restriction of the projection map π_i is T-equivariant and the image $\pi_i(U \cap X_{m,n}) \leq X_i$ is a T-invariant subgroup by the lemma above.

Since T acts irreducibly on X_i in (ii), we have $\pi_i(U \cap X_{m,n}) = X_i$ by the first part of this corollary if $\pi_i(U \cap X_{m,n}) \neq \{1\}$, i.e. π_i restricted to $U \cap X_{m,n}$ is surjective and thus an epimorphism. □

Note that these statements can also be extended to T-subgroups of the form $U \cap X_{-\infty,n}$ for π_n and $U \cap X_{m,\infty}$ for π_m, respectively.

Let us collect a few more consequences. The first three are similar to the assertions of [19, Corollary 2.6] and the last two are analogous to [9, Lemma 4.4(ii) and (iii)]:

Lemma 3.33. *Let $x, y \in X$. Then the following statements hold:*

(i) $\mu(x^{-1}) = \mu(x)$ and $\nu(x^{-1}) = \nu(x)$.

(ii) If $\mu(x) \neq \mu(y)$, then $\mu(xy) = \min\{\mu(x), \mu(y)\} = \mu(yx)$.

(iii) If $\nu(x) \neq \nu(y)$, then $\nu(xy) = \max\{\nu(x), \nu(y)\} = \nu(yx)$.

Let $x \neq 1$ and let $1 \neq y \in U \leq_T X$. Then we further have:

(iv) If $\mu(x) = \mu(y)$ and $\pi_{\mu(y)}\left(U \cap X_{\mu(y),\nu(y)}\right)$ contains the first letter $x_{\mu(x)}$ of x, then there exists $y' \in U \cap X_{\mu(y),\nu(y)}$ such that $\mu(x) < \mu(y'x)$ and $\omega(y'x) < \max\{\omega(x), \omega(y)\}$. Moreover, the same holds for $\mu(xy')$ and $\omega(xy')$.

(v) If $\nu(x) = \nu(y)$ and $\pi_{\nu(y)}\left(U \cap X_{\mu(y),\nu(y)}\right)$ contains the last letter $x_{\nu(x)}$ of x, then there exists $y' \in U \cap X_{\mu(y),\nu(y)}$ such that $\nu(y'x) < \nu(x)$ and $\omega(y'x) < \max\{\omega(x), \omega(y)\}$. Moreover, the same holds for $\nu(xy')$ and $\omega(xy')$.

Proof. If $x = 1 = x^{-1}$, then $\mu(x) = \infty$ and $\nu(x) = -\infty$ and the first three statements hold as $\mu(xy) = \mu(yx) = \mu(y) = \min\{\mu(x), \mu(y)\}$ and $\nu(xy) = \mu(yx) = \nu(y) = \max\{\nu(x), \nu(y)\}$.

So, for the rest of this proof, we take a look at the case $x \neq 1$.

For $i \in \{\mu(x), \nu(x)\}$ we have $\pi_i\left(x^{-1}\right) = \pi_i(x)^{-1} = x_i^{-1} \in X_i^*$, since the projection map π_i restricted to $X_{\mu(x),\nu(x)}$ is a homomorphism; hence (i) follows.

It suffices to show (ii) to complete the proof of the next two assertions, because the third assertion can be proven in a similar way by using the projection π_n.
Hence, let $\mu(x) \neq \mu(y)$. Up to an interchange of x and y, we may assume that $\mu(x) < \mu(y)$. Setting $k := \max\{\nu(x), \nu(y)\}$, we have $x, y \in X_{\mu(x),k}$. With Lemma 3.31 we compute

$$\begin{aligned} \pi_{\mu(x)}(xy) &= \pi_{\mu(x)}(x)\pi_{\mu(x)}(y) = \pi_{\mu(x)}(x) \cdot 1 = x_{\mu(x)} \neq 1 \quad \text{and} \\ \pi_m(xy) &= \pi_m(x)\pi_m(y) = 1 \cdot 1 = 1 \end{aligned}$$

for all $m < \mu(x)$. Similarly, we get $\pi_{\mu(x)}(yx) = x_{\mu(x)}$ and $\pi_m(yx) = 1$ for all $m < \mu(x)$. Thus, we have $\mu(xy) = \mu(yx) = \mu(x) = \min\{\mu(x), \mu(y)\}$.

Again, we only proof (v), since (iv) can be seen analogously by using the projection map on the first factor of the normal forms of x and y.
Let $x \in X^*$ and $y \in U^*$ with $\nu(x) = \nu(y)$ such that $\pi_{\nu(y)}\left(U \cap X_{\mu(y),\nu(y)}\right)$ contains $x_{\nu(y)}$. Since the image is a subgroup, it also contains $x_{\nu(y)}^{-1}$; hence there is $y' \in U \cap X_{\mu(y),\nu(y)}$ with $\pi_{\nu(y)}(y') = x_{\nu(y)}^{-1}$. Using Lemma 3.31 again, we have

$$\pi_{\nu(x)}(xy') = \pi_{\nu(x)}(x)\pi_{\nu(x)}(y') = x_{\nu(y)}x_{\nu(y)}^{-1} = 1 = \pi_{\nu(x)}(y'x).$$

Thus, we infer $y'x, xy' \in X_{-\infty,l}$ for some $l \leqslant \nu(y) - 1$ and $\nu(y'x), \nu(xy') < \nu(x)$. Since $\mu(y'x), \mu(xy') \geqslant \min\{\mu(x), \mu(y)\}$, we also observe

$$\begin{aligned} \omega(y'x) &= \nu(y'x) - \mu(y'x) + 1 \\ &\qquad < \nu(y) - \mu(y'x) + 1 \leqslant \nu(y) - \min\{\mu(x), \mu(y)\} + 1, \\ \omega(xy') &< \nu(y) - \min\{\mu(x), \mu(y)\} + 1, \end{aligned}$$

which are equivalent to $\omega(y'x), \omega(xy') < \max\{\omega(x), \omega(y)\}$. □

If we set $U = X$ in the last two assertions, then $\pi_i\left(U \cap X_{\mu(y),\nu(y)}\right) = X_i$ for $i = \mu(y), \nu(y)$, so that they particularly hold for two arbitrary, non-trivial elements x and y of X.

If the $\mathbb{Z}$-system is irreducible, then also $\pi_i\left(U \cap X_{\mu(y),\nu(y)}\right) = X_i$.

3.4 Sets of even and odd words

As usual, let $\Xi = (X, (X_k)_{k\in\mathbb{Z}}, \varsigma, T)$ denote a $\mathbb{Z}$-system. Starting in this section, we distinguish two sets of words in a non-trivial, shift-invariant T-subgroup Y of X. These sets depend on the parity of the index of the first letter of an element in Y and are closely related to the T-index of Y in X.

Definition 3.34. ([9, Definition 6.1]) Let $Y \leq X$ be shift-invariant. We set

$$\begin{aligned} Y_{even} &:= \{y \in Y^* \mid \mu(y) \in 2\mathbb{Z}\} \text{ as well as} \\ Y_{odd} &:= \{y \in Y^* \mid \mu(y) \in 1 + 2\mathbb{Z}\}, \end{aligned}$$

and define (as in [19, p. 4]), if the respective sets are non-empty,

$$\omega_0 := \min\{\omega(y) \mid y \in Y_{even}\} \text{ as well as } \omega_1 := \min\{\omega(y) \mid y \in Y_{odd}\}.$$

We call an element $y \in Y_{even}$ (resp. Y_{odd}) *even* (resp. *odd*).

Remark 3.35. ([19, Lemma 3.8(iii)]) Let $Y \leq_T X$ be shift-invariant. If $\pi_0(Y_{0,\omega_0-1}) = \{1\}$, then $Y_{even} = \varnothing$, and if $\pi_1(Y_{1,\omega_1}) = \{1\}$, then $Y_{odd} = \varnothing$. Indeed, if one of the images $\pi_0(Y_{0,\omega_0-1})$ or $\pi_1(Y_{1,\omega_1})$ is trivial, then the first letters of any even or odd word of minimal width in Y_{0,ω_0-1} or Y_{1,ω_1} is 1, assuming that the corresponding set Y_{even} or Y_{odd} is non-empty. But then those even or odd words are actually even shorter, contradiction the minimality of ω_0 or ω_1, respectively; therefore $Y_{even} = \varnothing$ or $Y_{odd} = \varnothing$.

Next we take a closer look on the relationship between the T-index of a non-trivial and shift-invariant T-subgroup and its sets consisting of even or odd words. Therefore we prove an adapted version of [9, Lemma 6.3]. We separately prove the easier implication first, because it is nearly the same reasoning and we do not need any additional assumptions on the projection maps or on a certain subgroup of X:

Lemma 3.36. *([9, Lemma 6.3]) Let $\{1\} \neq Y \trianglelefteq_T X$ be shift-invariant. If Y is of finite T-index in X, then both sets Y_{even} and Y_{odd} are non-empty.*

Proof. We do the proof by contradiction. Without loss of generality we suppose that $Y_{odd} = \varnothing$, i.e. $\mu(y) \in 2\mathbb{N}$ for all $y \in Y^*$.
Since Y is shift-invariant, we have $\mu(Y^*) = \{\mu(y) \mid y \in Y^*\} = 2\mathbb{Z}$. Let k be an odd integer, $y \in Y$ be arbitrary and $x_k \in X_k^*$, which is obviously not contained in Y. By Lemma 3.33 (ii) we have $\mu(x_k y) = \min\{k, \mu(y)\}$. Thus, we get

$$\mu(x_k Y) = \{k\} \cup \{\mu(y) \mid \mu(y) < k, y \in Y^*\} = \{k\} \cup \{n \in 2\mathbb{Z} \mid n < k\}.$$

This means that for different odd integers k, k' and elements $x_k \in X_k^*$ resp. $x_{k'} \in X_{k'}^*$ we observe $\mu(x_k Y) \neq \mu(x_{k'} Y)$. But this implies $x_k Y \neq x_{k'} Y$. As $\langle x_k Y \rangle_T \subseteq X_k Y$ and Y contains no odd elements, the quotient X/Y can not be finitely T-generated, which contradicts our premise that Y is of finite T-index. $\square$

We need the following observation to prove Lemma 3.40, the almost converse of the previous lemma. It is inspired by the proof of [9, Lemma 6.3] and is a consequence of Lemma 3.33 (v).

Lemma 3.37. *Let $Y \leq_T X$ be shift-invariant and let both sets Y_{odd} and Y_{even} be non-empty. If $\pi_{\omega_0 - 1}\left(Y_{0,\omega_0-1}\right) = X_{\omega_0-1}$ or $\pi_{\omega_1}\left(Y_{1,\omega_1}\right) = X_{\omega_1}$, then $\omega_0 - 1$ and ω_1 have different parity.*

Proof. Without loss of generality, let $a \in Y^*_{0,\omega_0-1}$ and $b \in Y^*_{1,\omega_1}$ be even and odd words of shortest width, respectively. Assume that $\nu(a) = \omega_0 - 1$ and $\nu(b) = \omega_1$ have same parity. Then there exists $k \in \mathbb{Z}$ such that $\nu\left(\varsigma^k(a)\right) = \nu(b)$ with $\varsigma^k(a) \in Y$. We observe $\pi_{\omega_1}\left(\varsigma^k(a)\right) \in X_{\omega_1} = \pi_{\omega_1}\left(Y_{1,\omega_1}\right)$; thus there is $b' \in Y_{1,\omega_1}$ such that

$$\begin{aligned} &\nu\left(b'\varsigma^k(a)\right) < \nu\left(\varsigma^k(a)\right) \quad \text{and} \\ &\omega\left(b'\varsigma^k(a)\right) < \max\left\{\omega\left(\varsigma^k(a)\right), \omega(b')\right\} = \max\{\omega_0, \omega_1\} \end{aligned}$$

by Lemma 3.33 (v). If $\omega_1 < \omega_0$ or equivalently $\omega_1 \leqslant \omega_0 - 1$, then $k \leqslant 0$, $b'\varsigma^k(a)$ is even by assertion (ii) of the stated lemma, and $\omega(b'\varsigma^k(a)) < \omega_0$. If $\omega_0 < \omega_1$, then $k > 0$, $b'\varsigma^k(a)$ is odd, and $\omega\left(b'\varsigma^k(a)\right) < \omega_1$. (The case $\omega_0 = \omega_1$ can not occur since they have different parity by our assumption.) In either case we observe a contradiction to the minimality of ω_0 resp. ω_1. $\square$

Note that an irreducible Z-system always meets the If-criterion.

Lemma 3.38. *Let $k \in 2\mathbb{Z}$, $m, n \in \mathbb{Z}$, and $Y \leq_T X$ be shift-invariant.*

(i) If $\pi_0(Y_{0,\infty}) = X_0$, then $\pi_k(Y_{k,\infty}) = X_k$, and if $\pi_1(Y_{1,\infty}) = X_1$, then $\pi_{1+k}(Y_{1+k,\infty}) = X_{k+1}$. In particular, for $x \in X^*$ with $\mu(x) = m$ exists $y \in Y^*_{m,\infty}$ such that $\mu(x) = \mu(y)$ and $\pi_m(x) = \pi_m(y) \in \pi_m(Y_{m,\nu(y)})$.

(ii) If $\pi_{\omega_0-1}(Y_{-\infty,\omega_0-1}) = X_{\omega_0-1}$, then $\pi_{\omega_0-1+k}(Y_{-\infty,\omega_0-1+k}) = X_{\omega_0-1+k}$, and if $\pi_{\omega_1}(Y_{-\infty,\omega_1}) = X_{\omega_1}$, then $\pi_{\omega_1+k}(Y_{-\infty,\omega_1+k}) = X_{\omega_1+k}$. Furthermore, if $\omega_0 - 1$ and ω_1 are of different parity, then for $x \in X^*$ with $\nu(x) = n$ there exists $y \in Y^*_{-\infty,n}$ such that $\nu(x) = \nu(y)$ and $\pi_n(x) = \pi_n(y) \in \pi_n(Y_{\mu(y),n})$.

Proof. We show that $\pi_k(Y_{k,\infty}) = X_k$ holds for all $k \in 2\mathbb{Z}$ if $\pi_0(Y_{0,\infty}) = X_0$. The other three equalities then follow by the same reasoning.
Note that $\varsigma^{\frac{k}{2}}(\pi_0(Y_{0,\infty})) = \varsigma^{\frac{k}{2}}(X_0) = X_k$ for all $k \in 2\mathbb{Z}$ by (M4). Therefore it suffices to prove $\pi_k(Y_{k,\infty}) = \varsigma^{\frac{k}{2}}(\pi_0(Y_{0,\infty}))$ for all even k. We perform an induction on k.
The statement is true for $k = 0$ by our assumption. For $k \in \{-2, 2\}$, let $l = \frac{k}{2}$ and $y \in Y_{0,\infty}$. Then $\varsigma^l(y) \in Y_{k,\infty}$ and, since ς is an automorphism, also $\varsigma^l(\pi_0(y)) = \pi_k(\varsigma^l(y)) \in \pi_k(Y_{k,\infty})$; thus $\varsigma^l(\pi_0(Y_{0,\infty})) \subseteq \pi_k(Y_{k,\infty})$. If $y \in Y_{k,\infty}$, then $\varsigma^{-l}(y) \in Y_{0,\infty}$ and

$$\pi_0(\varsigma^{-l}(y)) = \varsigma^{-l}(\pi_k(y)) \Leftrightarrow \pi_k(y) = \varsigma^l(\pi_0(\varsigma^{-l}(y))) \in \varsigma^l(\pi_0(Y_{0,\infty})).$$

Hence, we have equality and $\pi_k(Y_{k,\infty}) = \varsigma^l(\pi_0(Y_{0,\infty})) = X_k$.
For the inductive step, let $|k| > |2|$ and $\varsigma^{l\mp1}(\pi_0(Y_{0,\infty})) = \pi_{k\mp2}(Y_{k\mp2,\infty})$ be true, where $k = 2l$. Using the arguments of the base case $k \in \{-2, 2\}$ for the last equality below, we infer

$$X_k = \varsigma^l(\pi_0(Y_{0,\infty})) = \varsigma^{\pm1}(\varsigma^{l\mp1}(\pi_0(Y_{0,\infty}))) = \varsigma^{\pm1}(\pi_{k\mp2}(Y_{k\mp2,\infty})) = \pi_k(Y_{k,\infty}).$$

The latter parts of both assertions are clear now. □

Particularly, if the $\mathbb{Z}$-system is irreducible and if Y_{even} or Y_{odd} are non-empty, then the restrictions of the corresponding projection maps meet the required conditions. Thus an application of the last two assertions of Lemma 3.33 is always possible if $\omega_0 - 1$ and ω_1 exist and have different parity.

Remark 3.39. We can replace ∞ by $\omega_0 - 1$ resp. ω_1 in the condition of assertion (i) above as well as $-\infty$ by 0 resp. 1 in the condition of assertion (ii) to obtain similar statements for shortest words in Y, i.e.

$$\pi_k(Y_{k,\omega_0-1+k}) = X_k \quad \text{and} \quad \pi_{1+k}(Y_{1+k,\omega_1+k}) = X_{k+1} \quad \text{as well as}$$
$$\pi_{\omega_0-1+k}(Y_{k,\omega_0-1+k}) = X_{\omega_0-1+k} \quad \text{and} \quad \pi_{\omega_1+k}(Y_{1+k,\omega_1+k}) = X_{\omega_1+k}$$

for all $k \in 2\mathbb{Z}$. In particular, $\omega_0 - 1$ and ω_1 are of different parity by Lemma 3.37 if $\pi_{\omega_0-1}(Y_{0,\omega_0-1}) = X_{\omega_0-1}$ or $\pi_{\omega_1}(Y_{1,\omega_1}) = X_{\omega_1}$.

The following result corresponds to the remaining implication of [9, Lemma 6.3] and is up to some assumptions the converse of Lemma 3.36 above. Our version is also inspired by [19, Lemma 3.3].

Lemma 3.40. *Let $Y \trianglelefteq_T X$ be shift-invariant such that both sets Y_{odd} and Y_{even} are non-empty and such that $\pi_{\omega_0-1}(Y_{0,\omega_0-1}) = X_{\omega_0-1}$ and $\pi_{\omega_1}(Y_{1,\omega_1}) = X_{\omega_1}$. Set $w := \max\{\omega_0 - 1, \omega_1\}$.*
If $\pi_0(Y_{0,\omega_0-1}) = X_0$ and $\pi_1(Y_{1,\omega_1}) = X_1$, then $X = X_{0,w}Y$. Additionally, if $X_{0,w}$ is finitely T-generated, then Y has finite T-index in X.

Proof. By (one of) the first two assumptions on the images of the projection maps the minimal widths $\omega_0 - 1$ and ω_1 are of distinct parity by Lemma 3.37. Furthermore, all four suppositions on the images together with the remark above secure the existence of even and odd words and of words with odd or even ending in Y (either way of minimal width) to apply Lemma 3.33 (iv) and (v) at will to any $x \in X^*$.

We claim that $X = X_{0,w}Y$, where $w = \max\{\omega_0 - 1, \omega_1\}$. Therefore, we prove $X_{-\infty,w}Y \subseteq X_{0,w}Y$ and $X_{0,\infty}Y \subseteq X_{0,w}Y$, which leads, since $X = X_{-\infty,-1}X_{0,\infty} = X_{-\infty,w}X_{0,\infty}$ holds, to the desired equation.

Note that the inclusions $X_{l,w}Y, X_{0,l}Y \subseteq X_{0,w}Y$ hold for all $0 < l < w$.

First, we prove that $X_{-k,w}Y \subseteq X_{0,w}Y$ holds for all $k \in \mathbb{N}$. We perform the proof by induction on k. If $k = 0$, then there is nothing to show.
Now, let $k \geqslant 1$ and suppose that our claim is correct for $-k+1$. Let x denote an element $x_{-k} \cdot \ldots \cdot x_w \in X_{-k,w}$ with $x_i \in X_i$ for all $-k \leqslant i \leqslant w$, and let $y \in Y$. If $\mu(x) > -k$, i.e. $x_{-k} = 1$, then we can directly apply our induction hypothesis to see that $xy \in X_{-k+1,w}Y \subseteq X_{0,w}Y$, so we may assume that this is not the case. Since $\mu(x) = -k$, there exists $y' \in Y^*_{-k,\infty}$ of minimal width such that $\mu(xy') > \mu(x)$ by Lemma 3.38 (i) and its remark together with Lemma 3.33 (iv). It follows that $xy'y \in X_{-(k-1),w}Y \subseteq X_{0,w}Y$ by our induction hypothesis. By normality of Y in X, we have $xy(y')^y = xy'y \in X_{0,w}Y \Leftrightarrow xy \in X_{0,w}Y$. We infer $X_{-\infty,w}Y = \bigcup_{l \leqslant w}(X_{l,w}Y) \subseteq X_{0,w}Y$.

Next, we show that $X_{0,k}Y \subseteq X_{0,w}Y$ holds for all $k \geqslant w$. We perform an induction on k, again. If $k = w$, then there is nothing to show.
Now, let $k > w$ and suppose that our statement holds for $k-1$. Let $x := x_0 \cdot \ldots \cdot x_k \in X_{0,k}$ with $x_i \in X_i$ for all $0 \leqslant i \leqslant k$, and $y \in Y$. As above we may assume $x_k \neq 1$. Using Lemma 3.38 (ii) and its remark combined with Lemma 3.33 (v), there exists $y'' \in Y^*_{-\infty,k}$ of minimal width such that $\nu(xy'') < \nu(x)$. Hence, we get $xy''y \in X_{0,k-1}Y \subseteq X_{0,w}Y$ by our induction hypothesis. Like before, we infer $xy(y'')^y = xy''y \in X_{0,w}Y$, and consequently $xy \in X_{0,w}Y$. We conclude $X_{0,\infty}Y = \bigcup_{l \in \mathbb{N}}(X_{0,l}Y) \subseteq X_{0,w}Y$.

Finally, we have

$$X = X_{-\infty,w}X_{0,\infty} \subseteq X_{-\infty,w}YX_{0,\infty}Y \subseteq X_{0,w}YX_{0,w}Y = X_{0,w}Y,$$

as $X_{0,w}Y \leq X$, what proves the equality and the first part of the lemma.

The second part is now easy. With the adapted Second Isomorphism Theorem 3.26 (ii), we observe

$$X/Y = X_{0,w}Y/Y \cong_T X_{0,w}/Y_{0,w}.$$

If $X_{0,w}$ is finitely T-generated, then the quotient $X_{0,w}/Y_{0,w}$ is also finitely T-generated by Lemma 3.27. Applying Corollary 3.28, we further infer that X/Y is finitely T-generated and Y is of finite T-index in X. □

Referring to [9, Lemma 6.3], we summarize this section for an irreducible $\mathbb{Z}$-system via the following

Proposition 3.41. *Let $\Xi = (X, (X_k)_{k\in\mathbb{Z}}, \varsigma, T)$ be an irreducible $\mathbb{Z}$-system and $\{1\} \neq Y \trianglelefteq_T X$ be shift-invariant. Then the following are equivalent:*

(i) The T-subgroup Y is of finite T-index in X.

(ii) Both subsets Y_{even} and Y_{odd} are non-empty.

Proof. The implication (i) ⇒ (ii) holds by Lemma 3.36. For the reverse implication note first that, under Corollary 3.32 (ii), all requirements on the projection maps of Lemma 3.37 and Lemma 3.40 are satisfied by irreducibility. Since T acts irreducibly on X_k for all $k \in \mathbb{Z}$, we have $\langle x_k\rangle_T = X_k$ for any element $x_k \in X_k^*$ and any $k \in \mathbb{Z}$. Therefore, the group $X_{0,w}$ is finitely T-generated by $w + 1$ elements. The second implication now follows by the previous lemma. □

3.5 Generators of a shift-invariant subgroup

In this section we study the set of (T-)generators of a shift-invariant T-subgroups Y of X for an irreducible $\mathbb{Z}$-system $\Xi = (X, (X_k)_{k\in\mathbb{Z}}, \varsigma, T)$.

For the first result we follow [9, Proposition 6.4(i)] if $Y_{even} = \varnothing$ or $Y_{odd} = \varnothing$. In the case $Y_{even} \neq \varnothing \neq Y_{odd}$ we utilize [9, Lemma 6.5(i)].

Note that we can substitute $Y_{even} = \varnothing$ or $Y_{odd} = \varnothing$ with Y being of infinite T-index and $Y_{even} \neq \varnothing \neq Y_{odd}$ with Y being of finite T-index by Proposition 3.41, but then we also require $Y \trianglelefteq X$.

Proposition 3.42. *Let $Y \leq_T X$ be non-trivial and shift-invariant.*

(i) If either $Y_{even} = \varnothing$ or $Y_{odd} = \varnothing$, then there is $s \in Y^*$ of minimal width such that $Y = \left\langle \varsigma^k(s) \,\middle|\, k \in \mathbb{Z}\right\rangle_T$.

(ii) If $Y_{even} \neq \varnothing \neq Y_{odd}$, then there are $a \in Y_{even}^*$ and $b \in Y_{odd}^*$ of their respective minimal width such that $Y = \left\langle \varsigma^k(a), \varsigma^l(b) \,\middle|\, k, l \in \mathbb{Z}\right\rangle_T$.

Proof. First, let either Y_{even} or Y_{odd} be empty, i.e. each two non-trivial words have same parity. Let $s \in Y^*$ be of minimal width in Y and set $U := \left\langle \varsigma^k(s) \,\middle|\, k \in \mathbb{Z}\right\rangle_T$. We have $U \subseteq Y$ by the shift- and T-invariance of Y.

For the reverse inclusion we use induction on $\omega(y)$ for $y \in Y$. If $\omega(y) = 0$, or more general $\omega(y) < \omega(s)$, then $y = 1$ is contained in U. So, let $\omega(y) \geqslant \omega(s)$ and assume that $z \in U$ for all $z \in Y$ with $\omega(z) < \omega(y)$. Since s and y have same parity, there is $l \in \mathbb{Z}$ such that $\mu(y) = \mu(s) + 2l = \mu\left(\varsigma^l(s)\right)$, where $\varsigma^l(s) \in U^*$ by definition. Since Ξ is irreducible and

$$1 \neq \pi_{\mu\left(\varsigma^l(s)\right)}\left(\varsigma^l(s)\right) \in \pi_{\mu\left(\varsigma^l(s)\right)}\left(U \cap X_{\mu\left(\varsigma^l(s)\right),\nu\left(\varsigma^l(s)\right)}\right),$$

we have

$$y_{\mu(y)} \in X_{\mu\left(\varsigma^l(s)\right)} = \pi_{\mu\left(\varsigma^l(s)\right)}\left(U \cap X_{\mu\left(\varsigma^l(s)\right),\nu\left(\varsigma^l(s)\right)}\right)$$

by Corollary 3.32 (ii). Therefore, Lemma 3.33 (iv) yields the existence of an element $y' \in U \cap X_{\mu\left(\varsigma^l(s)\right),\nu\left(\varsigma^l(s)\right)}$ such that

$$\omega(y'y) < \max\left\{\omega(y), \omega\left(\varsigma^l(s)\right)\right\} = \omega(y),$$

as $\omega(y') = \omega\left(\varsigma^l(s)\right) = \omega(s) \leqslant \omega(y)$. Hence, we have $y'y \in U$ by our induction hypothesis. Since $y' \in U$, we further infer $y \in U$ and the equality $U = Y$.

The proof of assertion (ii) is done analogously. The inclusion of the T-subgroup $\left\langle \varsigma^k(a), \varsigma^l(b) \,\middle|\, k, l \in \mathbb{Z}\right\rangle_T$ in Y is clear. For the other inclusion one uses a similar induction on the width of an element $y \in Y$ and replaces s, depending on the parity of y, either by a or b. $\square$

We combine both parts to obtain a general statement (see [9, Lemma 6.6]):

Corollary 3.43. *Let $Y \leq_T X$ be shift-invariant. Then there exist $a, b \in Y$ such that*

$$Y = \left\langle \varsigma^k(a), \varsigma^l(b) \,\middle|\, k, l \in \mathbb{Z}\right\rangle_T.$$

Proof. If $Y = \{1\}$, then we choose $a = 1 = b$. Otherwise, we either use $a = 1$ and $b = s$ with s as in Proposition 3.42 (i) or a and b as in assertion (ii) of the same proposition. $\square$

There is another, more handy description of Y in terms of generators when we use the fact that ς normalizes T :

Corollary 3.44. *Let $Y \leq_T X$ be shift-invariant. Then there exist $a, b \in Y$ such that*

$$Y = \left\langle \varsigma^k(a'), \varsigma^l(b') \,\middle|\, k, l \in \mathbb{Z},\ a' \in \langle a\rangle_T, b' \in \langle b\rangle_T\right\rangle.$$

Proof. As a consequence of the preceding corollary, there are $a, b \in Y$ such that $Y = \left\langle \varsigma^k(a), \varsigma^l(b) \,\middle|\, k, l \in \mathbb{Z} \right\rangle_T$. Let G denote the group on the right-hand side in the statement above. Since Y is a T-invariant subgroup, it contains $\langle a \rangle_T$ and $\langle b \rangle_T$. By its shift-invariance, we further infer $\varsigma^k(a'), \varsigma^l(b') \in Y$ for all $k, l \in \mathbb{Z}$, $a' \in \langle a \rangle_T$, and $b' \in \langle b \rangle_T$. Hence, the inclusion $G \subseteq Y$ follows.
The group G clearly contains all shifts of a and b. It remains to show that G is T-invariant for the reverse inclusion. Let $t \in T$, $k \in \mathbb{Z}$, and $a' \in \langle a \rangle_T$. If we set $t' := t^{\varsigma^k}$, which is an element of T by (M5), then we get $t.\varsigma^k(a') = \varsigma^k(t'.a')$ with $t'.a' \in \langle a \rangle_T$. Thus, the element $t.\varsigma^k(a')$ lies in G. Analogously, we see that $t.\varsigma^l(b') \in G$ for all $l \in \mathbb{Z}$ and $b' \in \langle b \rangle_T$. As every element $g \in G$ is a finite product $\varsigma^{j_1}(c_1) \cdot \ldots \cdot \varsigma^{j_r}(c_r)$ with $r \in \mathbb{N}$, integers $j_1, \ldots, j_r$ and elements $c_1, \ldots, c_r \in \langle a \rangle_T \cup \langle b \rangle_T$, we infer that G is T-invariant and $Y \subseteq G$. $\square$

Remark 3.45. Let m denote the minimum of ω_0 and ω_1 (provided that both exist). If we assume, by shift-invariance, that the generating elements a and b of minimal width are contained in Y_{0,ω_0-1} and in Y_{1,ω_1}, respectively, then $\langle a \rangle_T = Y_{0,\omega_0-1}$ if $m = \omega_0$, and $\langle b \rangle_T = Y_{1,\omega_1}$ if $m = \omega_1$.
Indeed, since Y_{0,ω_0-1} is a T-group, it contains the T-span of a. If $\langle a \rangle_T \neq Y_{0,\omega_0-1}$, then there exists $y \in Y_{0,\omega_0-1} \setminus \langle a \rangle_T$. Since $\mu(y) = \mu(a)$ and

$$\pi_0\left(\langle a \rangle_T\right) = \pi_0\left(\left(\langle a \rangle_T \cap Y\right) \cap X_{0,\omega_0-1}\right) = X_0$$

by irreducibility, Lemma 3.33 (iv) implies the existence of $y' \in \langle a \rangle_T$ such that $\omega(y'y) < \omega_0$. If $y'y = 1$, then we get $y = (y')^{-1} \in \langle a \rangle_T$. If $y'y$ is even or odd, then it is of width smaller than the minimum of both widths. Hence, we derive a contradiction either way. The claim for b follows analogously.

This remark especially holds if either $Y_{even} = \varnothing$ or $Y_{odd} = \varnothing$, i.e. the subgroup $\langle s \rangle_T$ equals $Y_{\mu(s),\nu(s)}$, where $s \in Y^*$ is a word of minimal width as in Proposition 3.42 (i).

Lemma 3.46. *Let $Y \leq_T X$ be shift-invariant with either $Y_{even} = \varnothing$ or $Y_{odd} = \varnothing$ and $s \in Y^*$ be an element of minimal width T-generating Y. If $1 \neq s' \in \langle s \rangle_T$, then also $s \in \langle s' \rangle_T$.*

Proof. By the preceding remark we have $\langle s \rangle_T = Y_{\mu(s),\nu(s)}$. Hence, an element $s' \in \langle s \rangle_T^*$ has minimal width, again. Since Ξ is irreducible and $\mu(s) = \mu(s')$, there is $s'' \in \langle s' \rangle_T \cap X_{\mu(s),\nu(s)}$ such that $\omega(s''s) < \omega(s)$ by Corollary 3.32 (ii) and Lemma 3.33 (iv), i.e. we have $s''s = 1$ and infer $s = (s'')^{-1} \in \langle s' \rangle_T$. In particular, we observe $\langle s' \rangle_T = Y_{\mu(s),\nu(s)}$. $\square$

Lemma 3.47. *Let $m \leqslant n \in \mathbb{Z}$ and $Y \leq_T X$ be a shift-invariant subgroup such that either $Y_{even} = \varnothing$ or $Y_{odd} = \varnothing$. If the shifts of $s \in Y^*$ generate Y as a T-group, then*

$$Y_{m,n} = \left\langle \varsigma^k(c) \,\middle|\, k \in \mathbb{Z}, c \in \langle s \rangle_T : m \leqslant \mu\left(\varsigma^k(c)\right),\ \nu\left(\varsigma^k(c)\right) \leqslant n \right\rangle.$$

In particular, we also have

$$Y_{-\infty,n} = \left\langle \varsigma^k(c) \,\middle|\, k \in \mathbb{Z}, c \in \langle s \rangle_T : \nu\left(\varsigma^k(c)\right) \leqslant n \right\rangle,$$
$$Y_{m,\infty} = \left\langle \varsigma^k(c) \,\middle|\, k \in \mathbb{Z}, c \in \langle s \rangle_T : m \leqslant \mu\left(\varsigma^k(c)\right) \right\rangle.$$

Proof. Since any y in one of the latter groups is contained in some $Y_{m,n}$, we only prove the first statement. Let S denote the span on the right-hand side of that first claim. Then $S \leq Y \cap X_{m,n} = Y_{m,n}$ by the fact that the set generating S is a subset of the generating set of Y and by the defining property of S.
For the reverse inclusion, let $y \in Y_{m,n}$. If $y = 1$, then it is the shift of 1 with $\mu(1) = \infty$ and $\nu(1) = -\infty$. Hence, let $y \neq 1$ in the following, i.e. $m \leqslant \mu(y) \leqslant \nu(y) \leqslant n$. If y is of minimal width, then the preceding remark implies that $\varsigma^k(y) \in \langle s \rangle_T$ for some $k \in \mathbb{Z}$. Hence, $y = \varsigma^{-k}(c)$ for a non-trivial $c \in \langle s \rangle_T$ is a single shift with $m \leqslant \mu\left(\varsigma^{-k}(c)\right) \leqslant \nu\left(\varsigma^{-k}(c)\right) \leqslant n$.
Now, let $\omega(s) < \omega(y)$ and assume that elements $z \in Y_{m,n}$ with $\omega(z) < \omega(y)$ have a decomposition consisting of shifts whose starting resp. ending letters have indices in between m and n. Since $\mu(y)$ is of same parity as $\mu(s)$, there exist $1 \neq c \in \langle s \rangle_T$ and $l \in \mathbb{Z}$ such that $\mu\left(\varsigma^l(c)\right) = \mu(y)$. By irreducibility, Lemma 3.33 (iv) implies that there is $y' \in Y_{\mu(\varsigma^l(c)),\nu(\varsigma^l(c))} \subseteq Y_{m,n}$ with $\omega(s) = \omega(y') < \omega(y)$ such that $y'y \in Y_{m+1,n}$ and $\omega(y'y) < \omega(y)$. By our induction hypothesis, we infer $y', y'y \in S$ and thus $y \in S$. $\square$

The descriptions of the subgroups of Y in the preceding lemma is also true if both Y_{even} and Y_{odd} are non-empty, but this fact is not explicitly needed in the later chapters. Nevertheless we give a proof.

Lemma 3.48. *Let $m \leqslant n \in \mathbb{Z}$ and $Y \leq_T X$ be a shift-invariant subgroup such that both Y_{even} and Y_{odd} are non-empty. If Y is T-generated by the shifts of words $a \in Y^*_{even}$ and $b \in Y^*_{odd}$ of their respective minimal width, then*

$$Y_{m,n} = \left\langle \varsigma^k(c) \,\middle|\, k \in \mathbb{Z}, c \in \langle a \rangle_T \cup \langle b \rangle_T : m \leqslant \mu\left(\varsigma^k(c)\right),\ \nu\left(\varsigma^k(c)\right) \leqslant n \right\rangle.$$

Proof. As seen above, the span, let us call it S again, is contained in $Y_{m,n}$. For the reverse inclusion, we may assume that $\omega_0 \leqslant \omega_1$. Note that in case of $\omega_0 = \omega_1$, Remark 3.45 implies that both even and odd words of minimal width are single shifts of elements in $\langle a\rangle_T$ and $\langle b\rangle_T$, respectively. An induction similar to the one in the lemma above yields that both even and odd words of $Y_{m,n}$ are contained in S. Therefore, we may further suppose that $\omega_0 < \omega_1$.

We perform an induction on $l := n - m + 1$. If $l < \omega_0$, then $Y_{m,n} = \{1\}$ is a subset of S. If $\omega_0 \leqslant l < \omega_1$, then $Y_{m,n}$ does not contain any odd words and the same induction used in the previous proof gives $Y_{m,n} \subseteq S$.

Now, let $l \geqslant \omega_1$, so that $Y_{m,n}$ contains even and odd words. Furthermore, let $y \in Y_{m,n}$ with $\omega(y) = l$ and let $Y_{m',n'}$ with indices $m \leqslant m' \leqslant n' \leqslant n$ such that $n' - m' + 1 < l$ be contained in S. If y is even, we can shorten it with a shift y' of an element in $\langle a\rangle_T$ with $\mu(y) = \mu(y')$ by Lemma 3.33 (iv) such that $y'y \in Y^*_{m+1,n}$. Hence, we infer $y \in S$ by our induction hypothesis applied to $y'y$ and y'. In case of y being odd, we can shorten it again with a shift y'' of an element $\langle b\rangle_T$ by the same part of the stated lemma. Note that $y'' \in S$ as a single shift with $\mu(y'') = m$ and $\nu(y'') = \omega_1 + m - 1 \leqslant n$. If y is of width ω_1, then $y''y$ is either trivial or even. In the first case, we have $y = (y'')^{-1} \in S$. We use the induction hypothesis for $Y_{m+1,n}$ to infer $y = (y'')^{-1}y''y \in S$ in the second case. If y has width greater that ω_1, then $y''y$ is either even or odd and we analogously infer $y \in S$ by our induction hypothesis in either case. The claim is proven. □

If $Y_{even} = \varnothing$ or $Y_{odd} = \varnothing$, then we can rearrange the shifts that generate Y. In fact, these shifts can be sorted by exponents as seen in the following

Lemma 3.49. *Let $Y \leq_T X$ be shift-invariant with either $Y_{even} = \varnothing$ or $Y_{odd} = \varnothing$. Let s denote an element of minimal width w whose shifts generate Y. Let $k \leqslant l$ be integers and $s', s'' \in \langle s\rangle_T$. Then the shifts $\varsigma^k(s')$ and $\varsigma^l(s'')$ can be ordered by exponent up to some further sorted shifts $y_1, \ldots, y_r$ with exponents $k < e_1 < \ldots < e_r < l$ such that*

$$\varsigma^l(s'')\varsigma^k(s') = \varsigma^k(s')y_1^{e_1} \cdot \ldots \cdot y_r^{e_r}\varsigma^l(s'').$$

Proof. We perform an induction on $N := l - k \in \mathbb{N}$. For $N = 0$, we simply combine the shifts and write $\varsigma^k(s')\varsigma^l(s'') = \varsigma^k(s's'')$, where $s's'' \in \langle s\rangle_T$. If $N = 1$, then $\left[\varsigma^l(s''), \varsigma^k(s')\right] \in Y_{m+1,n-1}$ by Proposition 3.22 (i) with $m = \mu\left(\varsigma^k(s')\right)$ and $n = \nu\left(\varsigma^l(s')\right)$. In fact, by the parity of words, the commutator is contained in $Y_{m+2,n-2}$ which is trivial as words within have width $n - 2 - (m + 2) + 1 < n - (m + 2) + 1 = w$. So they do commute in this case and we interchange them as desired. For $N = 2$ we similarly infer that $y = \left[\varsigma^l(s''), \varsigma^k(s')\right] \in Y_{m+2,n-2}$. But then y is either trivial and both shifts commute or it is of minimal width and for $z = k + 1 = l - 1$ we have $\varsigma^{-z}(y) \in Y_{\mu(s),\nu(s)}$, as $m + 2 = \mu(s) + 2k + 2 = \mu(s) + 2(k + 1)$ and

$n-2=\nu(s)+2l-2=\nu(s)+2(l-1)$. By Remark 3.45, $\varsigma^{-z}(y)$ is then an element s''' of $\langle s\rangle_T$, i.e. $y=\varsigma^z(s''')$. Involving the idea of Lemma 3.14 and the case $N=1$, we get

$$\varsigma^l(s'')\varsigma^k(s')=\varsigma^k(s')\varsigma^l(s'')y=\varsigma^k(s')y\varsigma^l(s'').$$

Now, let $N>0$ and assume that shifts z and z' such that the difference of their exponents is less than N can be ordered by exponents up to a product of already sorted shifts with exponents in between the respective exponents of z and z'. As before we have $y'=\left[\varsigma^l(s'),\varsigma^k(s'')\right]\in Y_{m+2,n-2}$ and $\varsigma^l(s'')\varsigma^k(s')=\varsigma^k(s')\varsigma^l(s'')y'$. Since y' is a product of shifts with exponents strictly between k and l by the lemma above, the difference of l and any of these exponents of y' is smaller than N. The differences among the exponents of the shifts in the decomposition of y' are also smaller than N. Hence, we can recursively apply our induction hypothesis to all those shifts while using the operations of summarizing, interchanging or inserting of a specific shift as seen for $N=0,1$, and 2, respectively. After finitely many steps we have ordered the exponents in a product with smallest exponent k and biggest exponent l. □

We obtain the following proposition on the *sorted form* of an element:

Proposition 3.50. *Let $\{1\}\neq Y\leq_T X$ be shift-invariant with $Y_{even}=\varnothing$ or $Y_{odd}=\varnothing$. Then there exists $s\in Y$ such that any $y\in Y$ can be written as*

$$y=\varsigma^{j_1}(s_1)\cdot\ldots\cdot\varsigma^{j_r}(s_r),$$

where $r\in\mathbb{N}$, $j_1<j_2<\ldots<j_r$ are integers, and $s_1,s_2,\ldots,s_r$ are non-trivial elements in $\langle s\rangle_T$.

Proof. Let s be as in Proposition 3.42 (i). If y has width smaller than or equal to w, then it is either trivial and has the empty product as its decomposition, or it is of minimal width, so that $y=\varsigma^j(s')$, for some $j\in\mathbb{Z}$ and $s'\in\langle s\rangle_T$, is the product of one single shift by Remark 3.45 as seen in the proof above.
In the remaining cases, and if y is not of sorted form, we apply Corollary 3.44 and the procedure of the preceding lemma. □

We have analogues for both observations if Y contains even and odd words:

Lemma 3.51. *Let $Y\leq_T X$ be shift-invariant with both Y_{even} and Y_{odd} non-empty. Let $a\in Y_{even}$ and $b\in Y_{odd}$ denote elements of minimal width whose shifts generate Y. Let $k\leqslant l$ be integers and $c,c'\in\langle a\rangle_T\cup\langle b\rangle_T$. Then the shifts $\varsigma^k(c)$ and $\varsigma^l(c')$ can be ordered by exponent up to some further* *sorted shifts with exponents between k and l (with respect to their beginning in case of equal exponents) in between $\varsigma^k(c)$ and $\varsigma^l(c')$.*

Proof. We may assume that $a \in Y_{0,\omega_1-1}$ and $b \in Y_{1,\omega_1}$ by the shift-invariance of Y, and that $\omega_0 \leqslant \omega_1$. We only look into the cases, where the shifts $\varsigma^k(c)$ and $\varsigma^l(c')$ are not sorted by exponent yet.

First, we deal with the case $\omega_0 = \omega_1 =: w$. We perform an induction on $N := l - k \in \mathbb{N}$, again. Let $N = 0$. If $c, c' \in \langle a \rangle_T$ or $c, c' \in \langle b \rangle_T$, then we combine the shifts. Else they commute, since their commutator is contained in $Y_{2l+1,w+2l-1}$ by Proposition 3.22 (i), which is trivial by $w + 2l - 1 - (2l + 1) + 1 < w$ and the minimality of w.
For $N = 1$, we observe in the cases $c, c' \in \langle a \rangle_T$ or $c, c' \in \langle b \rangle_T$ that the shifts either commute and we are done or only commute up to a commutator $z \in Y_{2l-1,w+2(l-1)}$ or $z \in Y_{2l,w+2l-1}$ of minimal width by the fact that $yx = xy[y, x]$ holds for arbitrary group elements together with Proposition 3.22 (i). By Remark 3.45 this commutator is a single shift and of different parity than c and c'. The commutators $\left[z, \varsigma^k(c)\right]$ and $\left[z, \varsigma^l(c')\right]$ are trivial by the stated proposition and the minimality of w, so that z commutes with both shifts and we can sort them by exponent. If $c' \in \langle a \rangle_T$ and $c \in \langle b \rangle_T$, then $\left[\varsigma^l(c'), \varsigma^k(c)\right] \in Y_{2l,w+2(l-1)} = \{1\}$ by $w+2(l-1)-2l+1 < w$ and they commute. If $c \in \langle a \rangle_T$ and $c' \in \langle b \rangle_T$, then $\left[\varsigma^l(c'), \varsigma^k(c)\right] \in Y_{2l-1,w+2l-1}$ has width $w + 1$ and is a product of shifts $z_1, \ldots, z_n$ of elements in $\langle a \rangle_T \cup \langle b \rangle_T$ with $\mu(z_i) \in \{2l - 1, 2l\}$ for all $1 \leqslant i \leqslant n$ by Lemma 3.48. If $\mu(z_1) = 2l$, then we are in the case $N = 0$ and z_1 commutes with $\varsigma^l(c')$. If $\mu(z_1) = 2l-1$, then $\left[z_1, \varsigma^l(c')\right] \in Y_{2l,w+2l-1}$ is either trivial and we can switch z_1 with $\varsigma^l(c')$ or it is an even word of minimal width (and thus a single shift y_1 by Remark 3.45) with the same exponent l and they also commute. Hence, we have

$$\varsigma^l(c')\varsigma^k(c) = \varsigma^k(c) z_1 y_1 \varsigma^l(c') z_2 \cdot \ldots \cdot z_n$$

with either $y_1 = 1$ or y_1 of minimal width and $\mu(y_1) = 2l$. Inductively, we can pull $\varsigma^l(c')$ through to the end up to a product $z_1 y_1 z_2 y_2 \cdot \ldots \cdot z_n y_n$ in between $\varsigma^k(c)$ and $\varsigma^l(c')$ with y_i trivial or $\mu(y_i) = 2l$ for $1 \leqslant i \leqslant n$. Each two adjacent factors of this product can either be summarized to one shift if one of them is trivial or if both of them are even resp. odd, or one of them is odd with exponent one smaller than the even word and they commute as seen in the second case for $N = 1$. Algorithmically, we get a product $\varsigma^k(c) z y \varsigma^l(c')$ with $z = \varsigma^k(b')$ and $y = \varsigma^l(a')$ for $a' \in \langle a \rangle_T$ and $b' \in \langle b \rangle_T$ that is sorted by exponents (and beginnings in case of same exponents).

Now, let $N > 0$ and assume that shifts x and x' such that the difference of their exponents is less than N can be ordered by exponents up to a product of already sorted shifts with exponents (not necessarily strictly) in between the respective exponents of x and x' and so that shifts with same exponent are sorted by their beginning. We distinguish four cases depending on the parity of c and c'. If $c, c' \in \langle a \rangle_T$, then $y = \left[\varsigma^l(c'), \varsigma^k(c)\right] \in Y_{2k+1,w+2l-2}$ is a product of shifts with exponents between k and $l-1$ by Lemma 3.48 and as $w+2l-2-w+1 = 2l-1$ is the largest possible beginning of a non-trivial shift. This product is sorted by our induction hypothesis, i.e. we have

$$\varsigma^k(c)\varsigma^l(c')y = \varsigma^k(c)\varsigma^l(c')d_1 \cdot \ldots \cdot d_r$$

for $r \in \mathbb{N}$, shifts d_j with exponents $e_j \leqslant e_{j+1}$ for $1 \leqslant j \leqslant r-1$ such that either d_j is even with $\mu(d_{j+1}) = \mu(d_j)+1$ and $e_j = e_{j+1}$ or $e_j < e_{j+1}$. If $\mu(d_1) \geqslant 2k+2$, then its exponent is strictly greater than k and we can algorithmically sort every shift in the right-hand side product by our hypothesis. Otherwise we commute $\varsigma^l(c')$ and d_1 up to $d = \left[\varsigma^l(c'), d_1\right] \in Y_{2k+2,w+2l-2}$ which is a product of shifts with exponents in between $k+1$ and $l-1$. We apply the induction hypothesis to order all unsorted shifts. In the case $c, c' \in \langle b \rangle_T$ we have $y \in Y_{2k+2,w+2l-1}$ which is a product of shifts with exponents between $k+1$ and l, and we use our induction hypothesis to sort all shifts. If $c' \in \langle a \rangle_T$ and $c \in \langle b \rangle_T$, then $y \in Y_{2k+2,w+2l-2}$ is a product of shifts with exponents between $k+1$ and $l-1$, and we use our induction hypothesis, again. In the last case $c \in \langle a \rangle_T$ and $c' \in \langle b \rangle_T$ we observe $y \in Y_{2k+1,w+2l-1}$. It is a product of shift with exponents between k and l. Following the arguments of the first case, we may have to commute $\varsigma^l(c')$ with a shift of exponent k up to a commutator that is a product of shifts with exponents greater or equal $k+1$ before applying the induction hypothesis.

Now, let $\omega_0 < \omega_1$, $m = \mu\left(\varsigma^k(c)\right)$, $n = \max\left\{\nu\left(\varsigma^k(c)\right), \nu\left(\varsigma^l(c')\right)\right\}$, and set $M := n - m \geqslant 0$. As before, we observe

$$\varsigma^l(c')\varsigma^k(c) = \varsigma^k(c)\varsigma^l(c')\left[\varsigma^l(c'), \varsigma^k(c)\right] \in Y_{m,n}$$

with the commutator, denoted by z, contained in $Y_{m+1,n}$ and of width at most $n-(m+1)+1 = M$ by Proposition 3.22. If $k = l$ we can either summarize two even resp. odd shifts or commute shifts of distinct parity up to $z \in Y_{m+1,n} = Y_{2l+1,\omega_1+2l-1}$ with width $M < \omega_1$. If $z = 1$, we are done; else it is a product of even shifts with exponents greater or equal to $l+1$ and we can assume that it is sorted by Lemma 3.49, so that the right-hand side is sorted as desired. Hence, let $k < l$ in the following. We perform an induction on M.

If $M < \omega_0$, then $z = 1$ and there is nothing to prove. If $M = \omega_0$, then either $z = 1$ or z is of minimal width ω_0 with $\mu(z) = m+1$ even. In the second case we have m odd, $c \in \langle b \rangle_T$, and $c' \in \langle a \rangle_T$, since no odd shift with exponent $l > k$ is contained in $Y_{m+1,n}$ by $M = \omega_0$. We summarize z and $\varsigma^l(c') \in Y_{m+1,n}$ of minimal width.

Now, let $M > \omega_0$ and assume that we can sort shifts contained in $Y_{m',n'}$ with $n' - m' < M$ by exponent as claimed. Since $x = \varsigma^l(c')\left[\varsigma^l(c'), \varsigma^k(c)\right] \in Y_{m+1,n}$ by $k < l$, the element x is a product of shifts of elements in $\langle a \rangle_T \cup \langle b \rangle_T$ that are contained in $Y_{m+1,n}$ by Lemma 3.48. As $n - m - 1 < M$, we apply the induction hypothesis to all those shifts (and the ones appearing in the process) to obtain a desired decomposition of x of shifts with sorted exponents of at least k. If all exponents are greater than k, then, as $\varsigma^k(c)$ has exponent k, the claim follows. If at least one of the first two factors has also exponent k, then we either summarize $\varsigma^k(c)$ and the first one, do nothing, or apply the hypothesis to sort all shifts, again. □

Proposition 3.52. *Let $Y \leq_T X$ be shift-invariant with both Y_{even} and Y_{odd} non-empty. Then there exist $a \in Y_{even}$ of minimal width ω_0 and $b \in Y_{odd}$ of minimal width ω_1 such that any $y \in Y$ can be written as*

$$y = \varsigma^{j_1}(c_1) \cdot \ldots \cdot \varsigma^{j_r}(c_r),$$

where $r \in \mathbb{N}$, $j_1 \leqslant j_2 \leqslant \ldots \leqslant j_r$ are integers, and $c_1, c_2, \ldots, c_r$ are non-trivial elements in $\langle a \rangle_T \cup \langle b \rangle_T$ such that $j_i = j_{i+1}$ for $1 \leqslant i \leqslant r-1$ only if $\varsigma^{j_i}(c_i)$ is even and $\varsigma^{j_{i+1}}(c_{i+1})$ is odd with $\mu\left(\varsigma^{j_i}(c_i)\right) + 1 = \mu\left(\varsigma^{j_{i+1}}(c_{i+1})\right)$.

Proof. Let a and b be as in Proposition 3.42 (ii). If y has width smaller than or equal to $w = \min\{\omega_0, \omega_1\}$, then it is either trivial or it is of minimal width and $y = \varsigma^j(c)$, for some $j \in \mathbb{Z}$ and $c \in \langle a \rangle_T \cup \langle b \rangle_T$, is the product of one single shift by Remark 3.45.
In the remaining cases, and if y is not of sorted form, we apply Corollary 3.44 and the preceding lemma. □

We have the following

Corollary 3.53. *([9, Lemma 6.8]) Let $Y \leq_T X$ be shift-invariant. Then for each $k \in \mathbb{Z}$ there exists $l \in \mathbb{Z}$ such that $Y = Y_{-\infty,l} Y_{k,\infty}$.*

Proof. Since Y is T-generated by shifts of elements a and b as in Corollary 3.43, we have to choose l large enough such that $Y_{-\infty,l}$ contains all shifts of a and b that are not contained in $Y_{k,\infty}$. By shift-invariance we may assume $a \in Y_{0,\omega_0-1}$ and $b \in Y_{1,\omega_1}$. If $\omega_0 - 1 \geqslant \omega_1$, then shifts z of a resp. b with $\mu(z) \leqslant k$ are contained in $Y_{-\infty,k+\omega_0-1}$ by

$$\nu(z) = \omega(z) + \mu(z) - 1 \leqslant \omega_0 + k - 1.$$

If $\omega_1 > \omega_0 - 1$, then shifts z of a resp. b with $\mu(z) \leqslant k$ are contained in $Y_{-\infty,k+\omega_1-1}$. Hence, it suffices to set $l := \max\{k+\omega_0-1, k+\omega_1-1\}$ for $Y_{-\infty,l}$ to contain all shifts not contained in $Y_{k,\infty}$.
If $y \in Y$, then it has a decomposition sorted by exponents as proven in the preceding propositions. Each shift in this decomposition is contained in $Y_{k,\infty}$ or in $Y_{-\infty,l}$, since l is large enough. Therefore, we infer $Y \subseteq Y_{-\infty,l}Y_{k,\infty}$ which implies the desired equality. $\square$

In fact, we can choose l smaller in the case of either $Y_{even} = \varnothing$ or $Y_{odd} = \varnothing$ if k is of same parity as the element $s \in Y$ of minimal width w whose shifts T-generate Y. Indeed, if we assume, by shift-invariance of Y, that $\mu(s) = k$, then on one hand $\varsigma^j(s') \in Y_{k,\infty}$ for all $j \in \mathbb{N}$ and $s' \in \langle s \rangle_T$, and on the other hand $\varsigma^j(s') \in Y_{-\infty,k+w-1+2j}$ for all $j < 0$ and $s' \in \langle s \rangle_T$, where

$$Y_{-\infty,k+w-1+2j} \subseteq Y_{-\infty,k+w-1-2} = Y_{-\infty,k+w-3}$$

for all $j < 0$. Thus we can set $l = k + w - 3 = \nu(s) - 2$.

Chapter 4

T-local nilpotency

In this chapter we assume that $\Xi = (X, (X_k)_{k\in\mathbb{Z}}, \varsigma, T)$ is a nilpotent $\mathbb{Z}$-system, i.e. its rooted groups X_k are nilpotent for all $k \in \mathbb{Z}$. Our goal is to show that X is T-locally nilpotent (see Definition 3.4). This assumption plays a crucial role in the upcoming chapters.

First, we state a general observation:

Lemma 4.1. *Let G be a group and $A, U, V \leq G$ subgroups of G such that $U \trianglelefteq V \leq N_G(A)$. Then $AU \trianglelefteq AV$.*

Proof. Note that $AU \leq AV \leq G$ as both U and V normalize A, and thus permute with A. Let $x \in AU$ and $y \in AV$ with $x = au$ and $y = bv$ for some $a, b \in A$, $u \in U$, and $v \in V$. Since $A, V \subseteq N_G(A)$ and $U \trianglelefteq V \subseteq N_G(A)$, we compute

$$a^y = v^{-1}\underbrace{b^{-1}ab}_{=:\,a'\,\in\,A}v = v^{-1}a'v \in A \quad \text{as well as}$$
$$u^y = v^{-1}b^{-1}ubv = v^{-1}\underbrace{b^{-1}ubu^{-1}}_{=:\,b'\,\in\,A}uv = \underbrace{v^{-1}b'v}_{=:\,b''\,\in\,A}\;\underbrace{v^{-1}uv}_{=:\,u'\,\in\,U} = b''u' \in AU.$$

We infer $x^y = a^y u^y \in AU$ and $AU \trianglelefteq AV$. $\square$

The proof of the following lemma uses ideas of Schreiers Refinement Theorem [14, 3.1.2] and is due to Scholz (cf. [19, Lemma 4.1]).

Lemma 4.2. *The subgroup $X_{m,n}$ is nilpotent for all $m \leqslant n \in \mathbb{Z}$.*

Proof. We proof the statement for $X_{0,n}$ by induction on $n \in \mathbb{N}$. As $X_{m,n}$ is a subgroup of $X_{0,n}$ for $0 \leqslant m \leqslant n$ and subgroups of nilpotent groups are also nilpotent (see, for example, [14, 5.1.4]), the claim follows then also for $m = 1$ and by shifting for all choices of $m < n \in \mathbb{Z}$.

If $n = 0$, then $X_{0,n} = X_0$ is nilpotent by our assumption.

Now, let $X_{0,n}$ be nilpotent for some $n \in \mathbb{N}$ with nilpotent class $s \in \mathbb{N}$. Using the definition of nilpotency, we obtain a lower central series

$$X_{0,n} = \gamma_0(X_{0,n}) =: L_0 \trianglerighteq \gamma_1(X_{0,n}) =: L_1 \trianglerighteq \ldots \trianglerighteq \{1\} = \gamma_s(X_{0,n}) =: L_s.$$

By Remark 3.16 we know that X_{n+1} normalizes $X_{0,n} = L_0$. Since L_i is a characteristic subgroup of L_0 for all $0 \leqslant i \leqslant s$, the group X_{n+1} is contained in the normalizer of L_i for all such i. This fact will come in handy later in this proof.

For $0 \leqslant i \leqslant s-1$ and $0 \leqslant j \leqslant n$ we define

$$K_{i,j} := L_{i+1}(L_i \cap X_{j,n}).$$

First, we observe $K_{i,0} = L_i$ for each $0 \leqslant i \leqslant s-1$. Since $L_{i+1} \trianglelefteq L_i$ for $0 \leqslant i \leqslant s-1$, the set $K_{i,j}$ is indeed a subgroup of L_i for every j.

Using Lemma 3.15, we see $X_{j+1,n} \trianglelefteq X_{j,n}$ for all $0 \leqslant j \leqslant n-1$. Indeed, since $[X_{j+1,n}, X_{j,n}] \leq X_{j+1,n}$, we infer $X_{j+1,n}^y \subseteq X_{j+1,n}X_{j+1,n} = X_{j+1,n}$ for all $y \in X_{j,n}$.

Consequently, we have $L_i \cap X_{j+1,n} \trianglelefteq L_i \cap X_{j,n}$ for all i and j as above. Furthermore, this leads to $K_{i,j+1} \trianglelefteq K_{i,j}$ by applying the preceding lemma with $A = L_{i+1}, U = L_i \cap X_{j+1,n}$, and $V = L_i \cap X_{j,n}$.

For $1 \leqslant i \leqslant s-1$ we have, since $L_i \trianglelefteq L_{i-1}$, at last

$$K_{i-1,n} = L_i(L_{i-1} \cap X_n) \trianglerighteq L_i = K_{i,0}.$$

Therefore, we obtain a refined series

$$\begin{aligned} K_{0,0} &\trianglerighteq K_{0,1} \trianglerighteq \ldots \trianglerighteq K_{0,n} \trianglerighteq K_{1,0} \trianglerighteq \ldots \trianglerighteq K_{1,n} \trianglerighteq \ldots \\ &\trianglerighteq K_{s-2,0} \trianglerighteq \ldots \trianglerighteq K_{s-2,n} \trianglerighteq K_{s-1,0} \trianglerighteq \ldots \trianglerighteq K_{s-1,n} \trianglerighteq \{1\}. \end{aligned}$$

For convenience we set $K_{s,0} := \{1\}$. Renaming this series, we get a series

$$\begin{aligned} C_1 := K_{0,0} \trianglerighteq C_2 := K_{0,1} \trianglerighteq \ldots \trianglerighteq C_{n+1} := K_{0,n} \trianglerighteq C_{1+n+1} := K_{1,0} \trianglerighteq \ldots \\ \trianglerighteq C_{2(n+1)} := K_{1,n} \trianglerighteq \ldots \trianglerighteq C_{(s-1)(n+1)} := K_{s-2,n} \trianglerighteq \ldots \\ \trianglerighteq C_{s(n+1)} := K_{s-1,n} \trianglerighteq C_{s(n+1)+1} := \{1\}. \end{aligned}$$

Note that each group C_l, $1 \leqslant l \leqslant s(n+1)$, is normalized by X_{n+1}, since L_i and $X_{j,n}$ are normalized by X_{n+1} for all $0 \leqslant i \leqslant s-1$ and $0 \leqslant j \leqslant n$ (see the first part of this proof).

We will prove in the following that this series, let us call it S, is a terminating descending central series for $C_1 = X_{0,n}$, i.e. for each $1 \leqslant l < s(n+1)$

we have $C_{l+1} \trianglelefteq C_1$ and $C_l/C_{l+1} \leq Z(C_1/C_{l+1})$. In the case $l = s(n+1)$, where $C_l = K_{s-1,n} \leq L_{s-1}$ and $C_{l+1} = \{1\}$, there is nothing to show since by the choice of s we have $C_l \leq Z(C_1)$.

Since $[C_1, C_l] \leq C_{l+1}$ for all $1 \leqslant l < s(n+1)$ does not only imply that all C_{l+1} are normal in C_1, but also that the center condition for nilpotency is satisfied, it suffices to prove this condition.

First, we deal with the case where $C_l = K_{i-1,n}$ for some $1 \leqslant i \leqslant s-1$. Using Lemma 1.12 and the fact $L_i = [X_{0,n}, L_{i-1}] \trianglelefteq L_{i-1}$, we have

$$\begin{aligned}
[C_1, C_l] &= [X_{0,n}, K_{i-1,n}] = [X_{0,n}, L_i(L_{i-1} \cap X_n)] \\
&= [X_{0,n}, L_{i-1} \cap X_n][X_{0,n}, L_i]^{L_{i-1} \cap X_n} \\
&\leq [X_{0,n}, L_{i-1}][X_{0,n}, L_i]^{L_{i-1}} \\
&\leq L_i L_i^{L_{i-1}} = L_i = K_{i,0} = C_{l+1}.
\end{aligned}$$

For $C_l = K_{i,j}$ with $0 \leqslant i \leqslant s-1$ and $0 \leqslant j < n$, and by using the observation that $K_{i+1,0} \leq K_{i,j+1}$ holds, we compute

$$\begin{aligned}
[C_1, C_l] &= [X_{0,n}, K_{i,j}] = [X_{0,n}, L_{i+1}(L_i \cap X_{j,n})] \\
&= [X_{0,n}, L_i \cap X_{j,n}][X_{0,n}, L_{i+1}]^{L_i \cap X_{j,n}} \\
&\leq [X_{0,n}, L_i][X_{0,n}, L_{i+1}]^{L_i} \\
&\leq L_{i+1} L_{i+1}^{L_i} = L_{i+1} = K_{i+1,0} \leq K_{i,j+1} = C_{l+1}.
\end{aligned}$$

If we distinguish both cases again, we see that S is, in some sense, stabilized by X_{n+1}, i.e. even $[X_{n+1}, C_l] \leq C_{l+1}$ holds for all $1 \leqslant l < s(n+1)$. The remark at the beginning of this proof, i.e. X_{n+1} normalizes L_i for all $1 \leqslant i \leqslant s$, as well as Lemma 3.14 are utilized in the reasoning below:

$$\begin{aligned}
[X_{n+1}, C_l] &= [X_{n+1}, K_{i-1,n}] = [X_{n+1}, L_i(L_{i-1} \cap X_n)] \\
&= [X_{n+1}, L_{i-1} \cap X_n][X_{n+1}, L_i]^{L_{i-1} \cap X_n} \\
&\leq [X_{n+1}, X_n][X_{n+1}, L_i]^{L_{i-1}} \\
&\leq \{1\} L_i^{L_{i-1}} = L_i = K_{i,0} = C_{l+1}.
\end{aligned}$$

Analogously, by applying Lemma 3.15 to see that $[X_{n+1}, X_{j,n}] \leq X_{j+1,n}$, we have

$$\begin{aligned}
[X_{n+1}, C_l] &= [X_{n+1}, K_{i,j}] = [X_{n+1}, L_{i+1}(L_i \cap X_{j,n})] \\
&= [X_{n+1}, L_i \cap X_{j,n}][X_{n+1}, L_{i+1}]^{L_i \cap X_{j,n}} \\
&\leq ([X_{n+1}, L_i] \cap [X_{n+1}, X_{j,n}])[X_{n+1}, L_{i+1}]^{L_i} \\
&\leq (L_i \cap X_{j+1,n}) L_{i+1}^{L_i} = (L_i \cap X_{j+1,n}) L_{i+1} = K_{i,j+1} = C_{l+1}.
\end{aligned}$$

Adding $C_0 := X_{0,n+1} \trianglerighteq X_{0,n}$ to the descending series S, we will show that

$$C_0 \trianglerighteq C_1 \trianglerighteq \ldots \trianglerighteq C_{s(n+1)} \trianglerighteq \{1\}$$

is a central series for $X_{0,n+1}$. Since S is a central series for $X_{0,n} = C_1$ stabilized by X_{n+1} and since X_{n+1} normalizes C_l for $1 \leqslant l \leqslant s(n+1)$, we conclude

$$\begin{aligned}[C_0, C_t] = [X_{0,n+1}, C_t] = [C_1 X_{n+1}, C_t] = [C_1, C_t]^{X_{n+1}}[X_{n+1}, C_t] \\ \leq C_{t+1}^{X_{n+1}} C_{t+1} = C_{t+1}\end{aligned}$$

for all $0 \leqslant t \leqslant s(n+1)$ by the commutator relations showed above. This implies that $X_{0,n+1}$ is nilpotent, which completes the proof. □

We adjust [19, Proposition 4.3] to the T-locally case.

Proposition 4.3. *The group X is T-locally nilpotent.*

Proof. Let $U := \langle x_1, \ldots, x_k \rangle_T$ be a finitely T-generated subgroup of X. Without loss of generality we may assume that x_i is non-trivial for all $0 \leqslant i \leqslant k$.
Hence, for each $0 \leqslant i \leqslant k$ there exist $m_i \leqslant n_i \in \mathbb{Z}$ such that $x_i \in X^*_{m_i,n_i}$ by Lemma 3.12. We set

$$m := \min\{m_1, \ldots, m_k\} \text{ and } n := \max\{n_1, \ldots, n_k\}.$$

Then $U \leq X_{m,n}$ is nilpotent by the previous lemma, since subgroups of nilpotent groups are nilpotent again. □

If $T = \{\mathrm{id}_X\}$, then it readily follows that the group X is locally nilpotent.

The following observations are used in Chapter 6.

Lemma 4.4. *Let G be a T-locally nilpotent group, $U \leq_T G$, and $N \trianglelefteq_T G$. Then U and G/N are T-locally nilpotent.*

Proof. If W is a finitely T-generated subgroup of U, then W is a finitely T-generated T-subgroup of G, and thus nilpotent.

Let $H = \langle g_1 N, \ldots, g_n N \rangle_T$ be a finitely T-generated subgroup of G/N and set $K = \langle g_1, \ldots, g_n \rangle_T$, i.e. $H = KN/N = \rho(KN)$, where $\rho : G \to G/N$ is the canonical projection. By $\rho(L) = \rho(LN)$ for all $L \leq G$ together with Lemma 1.18 (ii) and (i), we get

$$\gamma_d(\rho(KN)) = \rho(\gamma_d(KN)) \leq \rho(\gamma_d(K)N) = \rho(\gamma_d(K)) \leq \rho(\gamma_d(KN))$$

for all $d \geqslant 0$, and infer

$$\gamma_d(H) = \gamma_d(\rho(KN)) = \rho(\gamma_d(K)N) = (\gamma_d(K)N)/N.$$

Since G is T-locally nilpotent, the finitely T-generated subgroup K is nilpotent. Hence there is $s \in \mathbb{N}$ such that $\gamma_s(K)$ eventually vanishes, and so does $\gamma_s(H) = (\gamma_s(K)N)/N = N/N = \{N\}$. Therefore, H is nilpotent and G/N is T-locally nilpotent. $\square$

Chapter 5

One-sided normal T-subgroups

We begin this section with two general observations regarding commutator groups.

Lemma 5.1. *([9, Lemma 7.3]) Let G be a nilpotent group and $H \leq G$ with $H \leq [H, G]$. Then $H = \{1\}$.*

Proof. Since G is nilpotent, its lower central series

$$G \trianglerighteq [G, G] \trianglerighteq [G, G, G] \trianglerighteq \dots$$

reaches $\{1\}$ after finitely many steps. Thus $[H, G, \dots, G]$ eventually vanishes. By successively using $H \leq [H, G]$, we infer that

$$H \leq [H, G] \leq [H, G, G] \leq [H, G, \dots, G] = \{1\},$$

i.e. $H = \{1\}$. □

We adapt [9, Lemma 7.4] to our setting of T-groups. Note that $[H, G] \trianglelefteq G$ by Corollary 1.14, so that $\langle [H, G] \rangle^G = [H, G]$ and we can omit the normal closure notation seen in our main source.

Lemma 5.2. *Let G be a T-locally nilpotent group and $H \leq_T G$ finitely T-generated. Then $H \leq [H, G] \Leftrightarrow H = \{1\}$.*

Proof. If H is trivial, then $H \leq [H, G]$ holds with equality.

For the other direction, let $H \leq [H, G]$ and let $\{h_1, \dots, h_n\} \subseteq G$ be a T-generating set of H, i.e. $H = \langle h_1, \dots, h_n \rangle_T$ for some $n \in \mathbb{N}$. Then for each $1 \leqslant i \leqslant n$ there exist $l_i \in \mathbb{N}$, elements $h_{i,j} \in H$ and $g_{i,j} \in G$ as well as $\epsilon_{i,j} \in \{-1, 1\}$, where $1 \leqslant j \leqslant l_i$, such that

$$h_i = [h_{i,1}, g_{i,1}]^{\epsilon_{i,1}} \cdot \ldots \cdot [h_{i,l_i}, g_{i,l_i}]^{\epsilon_{i,l_i}}.$$

We define

$$U := \langle g_{i,j} \mid 1 \leqslant i \leqslant n, 1 \leqslant j \leqslant l_i \rangle_T \quad \text{as well as} \quad V := \langle H, U \rangle_T$$

and observe $H \leq [H, U]$ and $H, U \leq_T V$. Since U and H are finitely T-generated, the subgroup V of the T-locally nilpotent group G also has this property, and thus is nilpotent.

Applying the lemma above to $H \leq_T V$ with $H \leq [H, U] \leq [H, V]$, we infer $H = \{1\}$. □

From here on we assume that $\Xi = (X, (X_k)_{k \in \mathbb{Z}}, \varsigma, T)$ is an irreducible $\mathbb{Z}$-system. Given a shift-invariant T-subgroup Y, we successively build up results for subgroups of the form $Y_{-\infty,k}$, $Y_{m,n}$, and $Y_{k,\infty}$, respectively, until we arrive at conditions for the subgroups $Y_{-\infty,k}$ and $Y_{k,\infty}$ to be normal in X; hence the term one-sided normal subgroup for such Y. In particular, it turns out that Y is abelian in those cases.

We start with an adjusted version of [9, Lemma 7.5].

Lemma 5.3. *Let $Y \leq_T X$ be shift-invariant. The following hold:*

(i) For each $k \in \mathbb{Z}$ there exists an element $y_{k-1} \in Y_{k-1,\infty}$ such that

$$Y_{k-1,\infty} = \langle y_{k-1}, Y_{k,\infty} \rangle_T.$$

Furthermore, the equality

$$Y_{k-1,k+N} = \langle y_{k-1}, Y_{k,k+N} \rangle_T$$

holds for all $N \geqslant \omega(y_{k-1}) - 2$.

(ii) For each $k \in \mathbb{Z}$ there exists an element $y_{k+1} \in Y_{-\infty,k+1}$ such that

$$Y_{-\infty,k+1} = \langle y_{k+1}, Y_{-\infty,k} \rangle_T.$$

Furthermore, we have

$$Y_{k-N,k+1} = \langle y_{k+1}, Y_{k-N,k} \rangle_T$$

for all $N \geqslant \omega(y_{k+1}) - 2$.

Proof. We only prove (i) in detail, since part (ii) follows in a symmetric way.

Let $k \in \mathbb{Z}$. We distinguish two cases. If $Y_{k-1,\infty} = Y_{k,\infty}$, then choose $y_{k-1} = 1$ and $Y_{k-1,\infty} = \langle y_{k-1}, Y_{k,\infty} \rangle_T$ holds.

If $Y_{k-1,\infty} \supsetneq Y_{k,\infty}$, then there is $y_{k-1} \in Y^*_{k-1,\infty}$ with $\mu(y_{k-1}) = k-1$. Using the restricted projection map $\pi_{k-1} : Y_{k-1,\infty} \to X_{k-1}$, we observe

$$\pi_{k-1}(Y_{k-1,\infty}) = X_{k-1} = \pi_{k-1}\left(\langle y_{k-1}, Y_{k,\infty}\rangle_T\right)$$

by the irreducible action of T, where $\langle y_{k-1}, Y_{k,\infty}\rangle_T \leq_T Y_{k-1,\infty}$. It follows by Lemma 3.31 and Lemma 1.7 that

$$Y_{k-1,\infty} = \ker(\pi_{k-1})\langle y_{k-1}, Y_{k,\infty}\rangle_T = Y_{k,\infty}\langle y_{k-1}, Y_{k,\infty}\rangle_T = \langle y_{k-1}, Y_{k,\infty}\rangle_T$$

Regarding the second part of (i), we observe in the first case above that $Y_{k-1,k+N} = Y_{k,k+N}$ and $Y_{k-1,k+N} = \langle y_{k-1}, Y_{k,k+N}\rangle_T$ with $y_{k-1} = 1$.
In the second case there is $y_{k-1} \neq 1$ with $\mu(y_{k-1}) = k-1$ and, since

$$\begin{aligned} \nu(y_{k-1}) &= \omega(y_{k-1}) + \mu(y_{k-1}) - 1 \\ &= \omega(y_{k-1}) + k - 2 \leqslant k + N \;\Leftrightarrow\; \omega(y_{k-1}) - 2 \leqslant N, \end{aligned}$$

$y_{k-1} \in Y_{k-1,k+N}$ for $N \geqslant \omega(y_{k-1}) - 2$. We infer $Y_{k-1,k+N} = \langle y_{k-1}, Y_{k,k+N}\rangle_T$ by replacing ∞ by $k+N$ in the reasoning above.

The proof of (ii) is analogously done by looking at $\nu(y_{k+1})$ for $y_{k+1} \in Y_{-\infty,k+1}$ and using the projection map π_{k+1}. For the second assertion of (ii) we only note in case of $y_{k+1} \neq 1$ that $\omega(y_{k+1}) - 2 \leqslant N$ is equivalent to

$$k - N \leqslant k - \omega(y_{k+1}) + 2 \leqslant k - \nu(y_{k+1}) + \mu(y_{k+1}) - 1 + 2 = \mu(y_{k+1})$$

and that $y_{k+1} \in Y_{k-N,k+1}$ holds in exactly these cases. Equality can be shown by using the projection map, again. □

The elements y_{k-1} and y_{k+1} in the statement above can be chosen, if non-trivial, to be of minimal width ω_0 or ω_1, since we only required $\mu(y_{k-1}) = k-1$ and $\nu(y_{k+1}) = k+1$.

A nice and direct consequence of this lemma is the following observation:

Lemma 5.4. *Let $Y \leq_T X$ be shift-invariant. Then $Y_{m,n}$ is finitely T-generated for all $m \leqslant n \in \mathbb{Z}$.*

Proof. If Y is trivial, i.e. finitely T-generated by zero elements, then so is $Y_{m,n}$ for all $m \leqslant n$. Thus, we only look into the case where Y is non-trivial in the following.

Set $w := \min\{\omega_0, \omega_1\}$ and let $m, n \in \mathbb{Z}$ with $m \leqslant n$. We perform an induction on $l := n - m + 1$ to show that $Y_{m,n}$ is finitely T-generated for all $m, n \in \mathbb{Z}$.

If $l < w$, then $Y_{m,n} = \{1\}$ by the minimality of w. In case of $l = w$ we distinguish two cases. If m is even and $w \neq \omega_0$ or m is odd and

$w \neq \omega_1$, then $Y_{m,n} = \{1\}$. If m is even and $w = \omega_0$ or m is odd and $w = \omega_1$, then there exits $y_m \in Y^*_{m,n}$ of width w with $\mu(y_m) = m$. Since $N = n-m-1 = w(y_m)-2$ and $n-(m+1)+1 < w$, the proof of assertion (i) of the preceding lemma implies (for $k = m+1$) that

$$Y_{m,n} = Y_{m,m+1+N} = \langle y_m, Y_{m+1,m+1+N}\rangle_T = \langle y_m, Y_{m+1,n}\rangle_T = \langle y_m\rangle_T.$$

The subgroup $Y_{m,n}$ is finitely T-generated in either case.
Now, let $l > w$ and suppose that $Y_{m',n'}$ with $m \leqslant m' \leqslant n' \leqslant n$ and $n' - m' + 1 < l$ is finitely T-generated. If $Y_{m,n} = Y_{m+1,n}$, then we are done, since $n-(m+1)+1 < l$. Otherwise there is $y_m \in Y_{m,n}$ such that $\mu(y_m) = m$. Since $l \geqslant \omega(y_m)$, we have $N = n-m-1 \geqslant \omega(y_m) - 2$ and we infer $Y_{m,n} = \langle y_m, Y_{m+1,n}\rangle_T$, again. Using our induction hypothesis, we conclude that $Y_{m,n}$ is finitely T-generated. □

For the following, we recall that the normal closure of a T-invariant subgroup A in X is T-invariant by Remark 3.9 again, i.e. $\langle\langle A\rangle^X\rangle_T = \langle A\rangle^X = \langle\langle A\rangle_T\rangle^X$. So the notion of T-invariance can be omitted when working with the normal closure of a T-invariant subgroup to ensure better legibility.

Lemma 5.5. *([9, Lemma 7.6]) Let $k \in \mathbb{Z}$ and let $Y \leq_T X$ be shift-invariant.*

(i) If $Y_{k-1,\infty} \subseteq \langle Y_{k,\infty}\rangle^X$, then there exists $M \in \mathbb{N}$ such that

$$Y_{k-1,k+N} \subseteq \langle Y_{k,k+N}\rangle^X$$

for all $N \geqslant M$.

(ii) If $Y_{-\infty,k+1} \subseteq \langle Y_{-\infty,k}\rangle^X$, then there exists $M \in \mathbb{N}$ such that

$$Y_{k-N,k+1} \subseteq \langle Y_{k-N,k}\rangle^X$$

for all $N \geqslant M$.

Proof. We only prove the first part in detail, since part (ii) follows in a symmetric way by using Lemma 5.3 (ii).

If $Y_{k-1,\infty} = Y_{k,\infty}$, then we also have $Y_{k-1,k+N} = Y_{k,k+N}$ (otherwise this would imply $Y_{k-1,\infty} \supsetneq Y_{k,\infty}$) and $Y_{k-1,k+N} \subseteq \langle Y_{k,k+N}\rangle^X$ for all $N \in \mathbb{N}$.
Hence, we suppose that $Y_{k,\infty} \subsetneq Y_{k-1,\infty} \subseteq \langle Y_{k,\infty}\rangle^X$. Let $y_{k-1} \in Y^*_{k-1,\infty}$ with $\mu(y_{k-1}) = k-1$. Then there exist $l \in \mathbb{N}$, elements $z_1, \dots, z_l \in Y^*_{k,\infty}$ and $x_1, \dots, x_l \in X$ such that

$$y_{k-1} = z_1^{x_1} \cdot \ldots \cdot z_l^{x_l},$$

as $Y_{k-1,\infty} \subseteq \langle Y_{k,\infty}\rangle^X$ and the latter is generated by conjugates of elements of $Y_{k,\infty}$. Define

$$M := \max\{\nu(z_j), \nu(y_{k-1}) \mid j = 1, \ldots, l\} - k.$$

Since for all $j = 1, \ldots, l$ we have $\nu(z_j) - k \geqslant k - k = 0$ by Lemma 3.12, we observe $M \in \mathbb{N}$. By the choice of M we deduce $z_j \in Y_{k,k+M}$ for all $j = 1, \ldots, l$; and thus $y_{k-1} \in \langle Y_{k,k+M}\rangle^X$. As

$$M \geqslant \nu(y_{k-1}) - k = \nu(y_{k-1}) - (\mu(y_{k-1}) + 1) = \omega(y_{k-1}) - 2,$$

we infer

$$Y_{k-1,k+M} = \langle y_{k-1}, Y_{k,k+M}\rangle_T \subseteq \langle Y_{k,k+M}\rangle^X$$

by Lemma 5.3 (i). If we replace M with $N \geqslant M$, then this statement remains true, and (i) has been proven.

For the second part of the lemma, we only need to set

$$M := k - \min\{\mu(z_j), \mu(y_{k+1}) \mid j = 1, \ldots, l\} \in \mathbb{N},$$

in case of $Y_{-\infty,k+1} \neq Y_{-\infty,k}$ if we have chosen an analogous expression $y_{k+1} = z_1^{x_1} \cdot \ldots \cdot z_l^{x_l}$ for $y_{k+1} \in Y_{-\infty,k+1}$ with $\mu(y_{k+1}) = k + 1$, where $l \in \mathbb{N}$, $z_1, \ldots, z_l \in Y^*_{-\infty,k}$ and $x_1, \ldots, x_l \in X$. Similarly, we observe $y_{k+1} \in \langle Y_{k-M,k}\rangle^X$ as well as $M \geqslant \omega(y_{k+1}) - 2$, and infer

$$Y_{k-M,k+1} = \langle y_{k+1}, Y_{k-M,k}\rangle_T \subseteq \langle Y_{k-M,k}\rangle^X$$

by Lemma 5.3 (ii) which remains true for $N \geqslant M$. □

Lemma 5.6. *([9, Lemma 7.7]) Let $Y \leq_T X$ be shift-invariant.*

(i) If $Y_{k-1,\infty} \subseteq \langle Y_{k,\infty}\rangle^X$ for all $k \in \mathbb{Z}$, then there exists $M \in \mathbb{N}$ such that

$$Y_{-\infty,M} \subseteq \langle Y_{0,M}\rangle^X.$$

(ii) If $Y_{-\infty,k+1} \subseteq \langle Y_{-\infty,k}\rangle^X$ for all $k \in \mathbb{Z}$, then there exists $M \in \mathbb{N}$ such that

$$Y_{0,\infty} \subseteq \langle Y_{0,M}\rangle^X.$$

Proof. We first proof (i). Let $k \in \mathbb{Z}$ be fixed for the following observation: By induction on $l \in \mathbb{N}$ we observe $Y_{k-l,\infty} \subseteq \langle Y_{k,\infty}\rangle^X$, since by our premise and the induction hypothesis we obtain

$$Y_{k-(l+1),\infty} = Y_{(k-l)-1,\infty} \subseteq \langle Y_{k-l,\infty}\rangle^X \subseteq \langle Y_{k,\infty}\rangle^X.$$

This particularly implies $Y = \bigcup_{l\in\mathbb{N}} Y_{k-l,\infty} \subseteq \langle Y_{k,\infty}\rangle^X$.

On the other hand we have elements $a, b \in Y$ such that

$$Y = \left\langle \varsigma^s(a), \varsigma^t(b) \,\middle|\, s, t \in \mathbb{Z} \right\rangle_T$$

by Corollary 3.43. Either both a and b are of different parity and of minimal width ω_0 and ω_1, respectively, or at least one of them is trivial. In either case, we may assume by the shift-invariance of Y that $a \in Y_{-2,\omega_0-3}$ and $b \in Y_{-1,\omega_1-2}$.

Setting $k = 0$ and $l = -2$ in the statement we have proven at the beginning, we see

$$a, b \in Y_{-2,\omega_0-3} \cup Y_{-1,\omega_1-2} \subseteq Y_{-2,\infty} \subseteq \langle Y_{0,\infty} \rangle^X.$$

If we apply Lemma 5.5 (i) twice with $k_1 = -1$ and $k_2 = 0$, respectively, then there exist $M_1, M_2 \in \mathbb{N}$ with $M_1 \geqslant \omega_0 - 2$ and $M_2 \geqslant \omega_1 - 2$ (see the previous proof) such that

$$\begin{aligned} &a \in Y_{-2,\omega_0-3} \subseteq Y_{-2,-1+M_1} \subseteq \langle Y_{-1,-1+M_1} \rangle^X \quad \text{and} \\ &b \in Y_{-1,\omega_1-2} \subseteq Y_{-1,M_2} \subseteq \langle Y_{0,M_2} \rangle^X. \end{aligned}$$

Set $M := \max\{M_1, M_2\}$. Then both inclusions lead to

$$Y_{-2,-1+M} \subseteq \langle Y_{-1,-1+M} \rangle^X \subseteq \langle Y_{-1,M} \rangle^X \subseteq \langle Y_{0,M} \rangle^X,$$

so that $a, b \in \langle Y_{0,M} \rangle^X$. Now, using the irreducible action of T and argue as in the proof of Lemma 5.3 (i) with $k = 0$, $N = M$, and $y_{-1} = b$ resp. $k = -1$, $N = M + 1$, and $y_{-2} = a$, we have

$$Y_{-1,M} = \langle b, Y_{0,M} \rangle_T \quad \text{and} \quad Y_{-2,M} = \langle a, Y_{-1,M} \rangle_T, \ \text{respectively;}$$

hence $Y_{-2,M} = \langle a, b, Y_{0,M} \rangle_T \subseteq \langle Y_{0,M} \rangle^X$. This remains true if we replace M by some $N \geqslant M$. By the shift-invariance of Y and X we therefore conclude

$$Y_{-4,M} = \varsigma^{-1}(Y_{-2,M+2}) \subseteq \varsigma^{-1}\left(\langle Y_{0,M+2} \rangle^X\right) = \langle Y_{-2,M} \rangle^X \subseteq \langle Y_{0,M} \rangle^X.$$

Inductively, it follows that $Y_{-r,M} \subseteq \langle Y_{0,M} \rangle^X$ for all $r \in 2\mathbb{N}$, and hence $Y_{-\infty,M} \subseteq \langle Y_{0,M} \rangle^X$.

Using the second parts of the previous lemmata, the same reasoning leads to $Y_{-M,\infty} \subseteq \langle Y_{-M,0} \rangle^X$ for some $M \in \mathbb{N}$. Up to replacing M by $M + 1$ (so that $M \in 2\mathbb{N}$), and shifting by $\varsigma^{\frac{M}{2}}$, we infer

$$Y_{0,\infty} = \varsigma^{\frac{M}{2}}(Y_{-M,\infty}) \subseteq \varsigma^{\frac{M}{2}}\left(\langle Y_{-M,0} \rangle^X\right) = \langle Y_{0,M} \rangle^X. \qquad \square$$

Corollary 5.7. *([9, Lemma 7.8]) Let $Y \leq_T X$ be shift-invariant. If $Y_{k-1,\infty} \subseteq \langle Y_{k,\infty}\rangle^X$ and $Y_{-\infty,k+1} \subseteq \langle Y_{-\infty,k}\rangle^X$ for all $k \in \mathbb{Z}$, then there exists $M \in \mathbb{N}$ such that $Y \leq \langle Y_{0,M}\rangle^X$.*

Proof. This follows immediately by the preceding lemma and Corollary 3.53, since there exists $M \in \mathbb{N}$ large enough such that

$$Y = Y_{-\infty,M} Y_{0,\infty} \leq \langle Y_{0,M}\rangle^X. \qquad \square$$

To most of the remaining results of this section we add the premise that X is T-locally nilpotent.

Lemma 5.8. *([9, Lemma 7.9]) Let X be T-locally nilpotent, $Y \leq_T X$ be non-trivial and shift-invariant. If $[Y, X] = Y$, then there is $k \in \mathbb{Z}$ such that at least one of the following holds:*

(i) $Y_{k-1,\infty} \nsubseteq \langle Y_{k,\infty}\rangle^X$,

(ii) $Y_{-\infty,k+1} \nsubseteq \langle Y_{-\infty,k}\rangle^X$.

Proof. We do the proof by contradiction. Hence, we assume that $Y_{k-1,\infty} \subseteq \langle Y_{k,\infty}\rangle^X$ and $Y_{-\infty,k+1} \subseteq \langle Y_{-\infty,k}\rangle^X$ for all $k \in \mathbb{Z}$. Then there exists $M \in \mathbb{Z}$ such that $Y \leq \langle Y_{0,M}\rangle^X$ by the foregoing corollary.

Since $Y_{0,M}$ is finitely T-generated by Lemma 5.4, we can write $Y_{0,M} = \langle y_1, \ldots, y_N\rangle_T$ for some $N \in \mathbb{N}$ and $y_i \in Y_{0,M} \subseteq Y = [Y, X]$ for $1 \leqslant i \leqslant N$. Hence, for each $1 \leqslant i \leqslant N$, there exist non-negative integers $l_i \in \mathbb{N}$, elements $z_{i,j} \in Y$ and $x_{i,j} \in X$ as well as exponents $\epsilon_{i,j} \in \{-1, 1\}$, where $1 \leqslant j \leqslant l_i$, such that

$$y_i = [z_{i,1}, x_{i,1}]^{\epsilon_{i,1}} \cdot \ldots \cdot [z_{i,l_i}, x_{i,l_i}]^{\epsilon_{i,l_i}} .$$

We set

$$\begin{aligned} m &:= \min\{\mu(z_{i,j}) \mid 1 \leqslant i \leqslant N, 1 \leqslant j \leqslant l_i\} \text{ as well as} \\ n &:= \max\{\nu(z_{i,j}) \mid 1 \leqslant i \leqslant N, 1 \leqslant j \leqslant l_i\}, \end{aligned}$$

and observe $y_i \in [Y_{m,n}, X]$ for all $1 \leqslant i \leqslant N$; thus $Y_{0,M} \subseteq [Y_{m,n}, X]$.

Finally, we have

$$Y_{m,n} \leq Y \leq \langle Y_{0,M}\rangle^X \subseteq \langle [Y_{m,n}, X]\rangle^X = [Y_{m,n}, X],$$

where the last equality follows by Corollary 1.14. Since $Y_{m,n}$ is finitely T-generated by Lemma 5.4, we infer $Y_{m,n} = \{1\}$ by Lemma 5.2. But then the same is true for Y contradicting that Y is non-trivial. $\square$

We call a T-subgroup Y *one-sided normal*, if it satisfies the following

Lemma 5.9. *([9, Lemma 7.10]) Let $\{1\} \neq Y \trianglelefteq_T X$ be shift-invariant and of infinite T-index in X.*

(i) If there exists $k \in \mathbb{Z}$ such that $Y_{k,\infty} \nsubseteq \langle Y_{k+1,\infty}\rangle^X$, then $Y_{k,\infty} \trianglelefteq X$.

(ii) If there exists $k \in \mathbb{Z}$ such that $Y_{-\infty,k} \nsubseteq \langle Y_{-\infty,k-1}\rangle^X$, then $Y_{-\infty,k} \trianglelefteq X$.

Proof. As in other results in this section, we only outline the proof of (i), since the other claim follows by symmetric arguments.

By Proposition 3.42 and its Corollary 3.44, there is some odd or even word $s \in Y^*$ of minimal width such that $Y = \langle \varsigma^r(s') \mid r \in \mathbb{Z}, s' \in \langle s\rangle_T\rangle$.

Assume that $Y_{k,\infty} \nsubseteq \langle Y_{k+1,\infty}\rangle^X$ for some $k \in \mathbb{Z}$, so that $\pi_k(Y_{k,\infty})$ is non-trivial. By shift-invariance of Y we may assume further that $\mu(s) = k$, so that $s \in Y_{k,\infty}$ but $s \notin \langle Y_{k+1,\infty}\rangle^X$.

Suppose that there is $h < k$ and $x_h \in X_h^*$ such that $1 \neq [x_h, s]$. This commutator is contained in $Y_{h+1,\nu(s)-1}$ by Lemma 3.18. Since $Y \trianglelefteq X$ and thus $[Y, X] \leq Y$, there exist $1 \leqslant r \in \mathbb{N}$, negative integers $j_1 < j_2 < \ldots < j_r$, and elements $s_1, \ldots, s_r \in \langle s\rangle_T^*$ such that

$$[x_h, s] = \varsigma^{j_1}(s_1) \cdot \ldots \cdot \varsigma^{j_r}(s_r)$$

by Proposition 3.50. Applying ς^{-j_1} to both sides, we get

$$[x_{h-2j_1}, \varsigma^{-j_1}(s)] = s_1 \cdot \varsigma^{j_2-j_1}(s_2) \cdot \ldots \cdot \varsigma^{j_r-j_1}(s_r).$$

Since $\mu(\varsigma^{-j_1}(s)) \geqslant k+2$, we have $(\varsigma^{-j_1}(s)^{-1})^{x_{h-2j_1}} \in \langle Y_{k+1,\infty}\rangle^X$, and infer $[x_{h-2j_1}, \varsigma^{-j_1}(s)] \in \langle Y_{k+1,\infty}\rangle^X$. Moreover, all integers $j_2 - j_1, \ldots, j_r - j_1$ are positive, so that the product

$$s_1 = [x_{h-2j_1}, \varsigma^{-j_1}(s)] \left(\varsigma^{j_2-j_1}(s_2) \cdot \ldots \cdot \varsigma^{j_r-j_1}(s_r)\right)^{-1}$$

is contained in $\langle Y_{k+1,\infty}\rangle^X$. But $\mu(s_1) = \mu(s) = k$ and the same reasoning as seen in the proof of Lemma 5.3 (i) yields

$$Y_{k,\infty} = \langle s_1, Y_{k+1,\infty}\rangle_T \leq \langle Y_{k+1,\infty}\rangle^X,$$

a contradiction to our premise $Y_{k,\infty} \nsubseteq \langle Y_{k+1,\infty}\rangle^X$. Therefore, we have $[x_h, s] = 1$ for all $x_h \in X_h^*$ and for all $h < k$.

For all $h < k$ it follows from the T-invariance of X_h that every element of the T-orbit of s commutes with each element of X_h. Since the commutator between a product of elements from $\langle s\rangle_T$ and an element x_h equals the product of conjugated commutators between the factors of the product with x_h by Lemma 1.12, we further observe $X_{-\infty,k-1} \leq C_X(\langle s\rangle_T)$. Since $Y_{k,\infty}$ is generated by the shifts $\varsigma^j(s')$ for $j \in \mathbb{N}$ and $s' \in \langle s\rangle_T$ by Lemma 3.47 as well as the fact that

$$[x_h, \varsigma^j(s')] = \varsigma^j\left([\varsigma^{-j}(x_h), s']\right) \in \varsigma^j\left([X_{h-2j}, \langle s\rangle_T]\right) = \{1\},$$

we also infer that $X_{-\infty,k-1} \leq C_X(Y_{k,\infty})$. Using $Y \trianglelefteq X$, we conclude

$$[Y_{k,\infty}, X] = [Y_{k,\infty}, X_{k,\infty}] \leq [Y, X_{k,\infty}] \cap X_{k,\infty} \leq Y \cap X_{k,\infty} = Y_{k,\infty},$$

which implies that $Y_{k,\infty} \trianglelefteq X$. □

Finally, we can state the analogue of [9, Proposition 7.11] which is the main result of Section 7 in our main source.

Proposition 5.10. *Let X be T-locally nilpotent, let $Y \trianglelefteq_T X$ be non-trivial and shift-invariant. If Y is of infinite T-index and $[Y, X] = Y$, then there exists $k \in \mathbb{Z}$ such that $Y_{k,\infty} \trianglelefteq X$ or $Y_{-\infty,k} \trianglelefteq X$.*

Proof. We first use Lemma 5.8 to get $l \in \mathbb{Z}$ such that $Y_{l-1,\infty} \nsubseteq \langle Y_{l,\infty}\rangle^X$ or $Y_{-\infty,l+1} \nsubseteq \langle Y_{-\infty,l}\rangle^X$ holds. Then we apply the previous result for $k = l-1$ resp. $k = l+1$ to see that $Y_{k,\infty} \trianglelefteq X$ or $Y_{-\infty,k} \trianglelefteq X$. □

Remark 5.11. We could replace the assumption $[Y, X] = Y$ in the preceding proposition by the premise of Lemma 5.9, i.e. the existence of $k \in \mathbb{Z}$ such that $Y_{k,\infty} \nsubseteq \langle Y_{k+1,\infty}\rangle^X$ or $Y_{-\infty,k} \nsubseteq \langle Y_{-\infty,k-1}\rangle^X$ holds. This is also true for the corollary below.

The proposition above will be applied in the proof of the Main Theorem 7.7. In particular, the premise that X is T-locally nilpotent holds if our $\mathbb{Z}$-system is nilpotent by Proposition 4.3. In fact, under our assumed irreducibility of Ξ, the T-local nilpotency of X is equivalent to the nilpotency of the rooted groups.

We have a nice consequence thereof that can be viewed as an enhancement of Lemma 3.49. It is due to the observation that the rooted groups, being T-irreducible and nilpotent, are abelian (see Lemma 3.10).

Corollary 5.12. *Let X be T-locally nilpotent, let $Y \trianglelefteq_T X$ be non-trivial and shift-invariant. If Y is of infinite T-index and $[Y, X] = Y$, then Y is abelian.*

Proof. By the proposition above there exists $m \in \mathbb{Z}$ such that $Y_{m,\infty} \trianglelefteq X$ or $Y_{-\infty,m} \trianglelefteq X$. Without loss of generality we assume the first case. Let $s \in Y^*$ be of shortest width with $\mu(s) = m$ and $\nu(s) = n$ such that it's shifts T-generate Y. If $s', s'' \in \langle s\rangle_T \subseteq Y_{m,n}$, then these elements commute as $[s', s''] \in Y_{m+1,n-1} = \{1\}$ by Corollary 3.24 and the minimality width of such elements. By $Y_{m,\infty} \trianglelefteq X$ we further infer

$$\left[\varsigma^k(s'), s''\right] \in Y_{m+2k+1,n-1} \cap Y_{m,\infty} = Y_{m,n-1} = \{1\}$$

for all $k < 0$ using the same reasoning. Therefore, arbitrary elements in $\langle s\rangle_T$ commute with negative shifts of elements in $\langle s\rangle_T$. Applying ς^{-k} to $[\varsigma^k(s'), s''] = 1$, we see that arbitrary elements in $\langle s\rangle_T$ also commute with positive shifts of elements in $\langle s\rangle_T$. It follows that each two shifts of such elements commute, since

$$\left[\varsigma^k(s'), \varsigma^l(s'')\right] = \varsigma^l\left(\left[\varsigma^{k-l}(s'), s''\right]\right) = 1$$

for all $k, l \in \mathbb{Z}$. As these shifts generate Y by Corollary 3.44, the claim follows. $\square$

Chapter 6

Infinite abelianization

Again, let the quadruple $\Xi = (X, (X_k)_{k\in\mathbb{Z}}, \varsigma, T)$ denote an irreducible $\mathbb{Z}$-system. In addition to the previous chapter, we assume that Ξ is nilpotent, so that X is T-locally nilpotent by Proposition 4.3.

The goal of this chapter is to prove that the abelianization $X_{ab} = X/\delta(X)$ of X is not finitely T-generated. Along the way we state some helpful observations regarding nilpotent T-groups that we utilize in the proofs.

It is notable that the proof of [9, Lemma 8.2] displays a possibly mistaken assumption of minimality on a certain natural number. We fix it by adding another case to our proof of an adjusted version.

Proposition 6.1. *([9, Lemma 8.2]) Let $\{1\} \neq Y \trianglelefteq_T X$ be shift-invariant. Then $\delta(Y) < Y$.*

Proof. We have $[Y, X] \leq Y$ by Corollary 1.13 as $Y \trianglelefteq X$. If Y is of infinite T-index, then either $[Y, X] < Y$ and $\delta(Y) \leq [Y, X] < Y$ or $[Y, X] = Y$ and Y is abelian with $\{1\} = \delta(Y) < Y$ by Corollary 5.12. Therefore, let Y be of finite T-index for the remainder of this proof.

Suppose that $\delta(Y) = Y$. Since $Y = \delta(Y) \leq [Y, X] \leq Y$, we also have $Y = [Y, X]$. Now, at least one of the statements of Lemma 5.8 is valid. We assume that (i) holds and prove the claimed implication only for that case, as the proof of (ii) can be shown in a similar fashion.

By Lemma 5.8 we have $Y_{k-1,\infty} \not\subseteq \langle Y_{k,\infty}\rangle^X$ for some $k \in \mathbb{Z}$. For legibility, we set $Z := \langle Y_{k,\infty}\rangle^X \trianglelefteq X$. We know that $Y_{k,\infty} \leq Z < Y$ and that Z is normal in Y and also T-invariant by Remark 3.9. Using Corollary 3.53, we also get $l \in \mathbb{Z}$ such that $Y = Y_{-\infty,l} Y_{k,\infty}$. Thus, we have

$$Y = Y_{-\infty,l} Y_{k,\infty} \leq Y_{-\infty,l} Z \leq Y,$$

and infer $Y_{-\infty,l}Z = Y$.

If $l \in \mathbb{Z}$ can be chosen minimal with the property that $Y_{-\infty,l}Z = Y$, we do so for this passage. By Corollary 1.13 and $Z \trianglelefteq X$ we infer $\delta(Y_{-\infty,l}Z) \leq \delta(Y_{-\infty,l})Z$. Moreover, we observe $\delta(Y_{-\infty,l}) \leq Y_{-\infty,l-1}$ by Corollary 3.24, so that

$$\delta(Y) = \delta(Y_{-\infty,l}Z) \leq \delta(Y_{-\infty,l})Z \leq Y_{-\infty,l-1}Z \lneq Y,$$

as $l \in \mathbb{Z}$ was chosen minimal such that equality holds. But this is a contradiction to our supposition. Hence, $\delta(Y) < Y$.

In the further course we deal with the case that l cannot be chosen minimal such that $Y_{-\infty,l}Z = Y$. This implies that $Y_{-\infty,l}Z = Y$ holds for all $l \in \mathbb{Z}$. We only deal with the case k even in detail and may assume that $k = 0$ for legibility. The case for odd k follows then by replacing Ξ by Ξ^{pc} as defined in Remark 2.34.

Since Y is shift-invariant, we observe

$$\varsigma(Z) = \varsigma\left(\langle Y_{0,\infty}\rangle^X\right) = \left(\langle Y_{2,\infty}\rangle^X\right) \subseteq Z,$$

and hence that $xZ = yZ$ implies $\varsigma(x)Z = \varsigma(y)Z$ for all $x, y \in Y$. The shift operator therefore induces a well-defined endomorphism $\bar{\varsigma} : Y/Z \to Y/Z : yZ \mapsto \varsigma(y)Z$ as

$$\begin{aligned}\bar{\varsigma}(xZyZ) &= \bar{\varsigma}((xy)Z) = \varsigma(xy)Z = (\varsigma(x)\varsigma(y))Z\\ &= \varsigma(x)Z\varsigma(y)Z = \bar{\varsigma}(xZ)\bar{\varsigma}(yZ)\end{aligned}$$

for all $x, y \in Y$. Clearly, the element $\varsigma^{-1}(y)Z$ is a pre-image of $yZ \in Y/Z$ for all $y \in Y$. Thus, the endomorphism $\bar{\varsigma}$ is surjective.

We claim that an element $z \in Z$ with $\mu(z) < 0$ is even. Indeed, if we assume that $\mu(z) = -2n-1$ for some $n \in \mathbb{N}$, then

$$z \in \langle Y_{0,\infty}\rangle^X \cap Y_{-2n-1,\infty} \leq \langle Y_{-2n,\infty}\rangle^X \cap Y_{-2n-1,\infty}.$$

Applying ς^n yields $\varsigma^n(z) \in \langle Y_{0,\infty}\rangle^X \cap Y_{-1,\infty} = Z \cap Y_{-1,\infty}$ and $\mu\left(\varsigma^n(z)\right) = -1$. As Z is chosen via Lemma 5.8, this contradicts our premise that $Y_{-1,\infty}$ is not contained in Z by the irreducible action of T and the reasoning used in the proof of Lemma 5.3.

Moreover, we claim that for all $n \in \mathbb{N}$ there exists $z \in Z^*$ such that $\mu(z) = -2n$. Let $n \in \mathbb{N}$ and $x \in Y^*_{-2n-1,\infty}$ with $\mu(x) = -2n-1$ which is possible as Y is of finite T-index in X. Since $Y = Y_{-\infty,-2n-2}Z$, there is $y \in Y_{-\infty,-2n-2}$ and $z \in Z$ such that $x = yz$. In particular, we have $z \neq 1$ by $\mu(y) \leqslant \nu(y) \leqslant -2n-2 < \mu(x)$, Lemma 3.33 (ii) and Lemma 3.12.

Suppose that $\mu(z) > -2n$. Then $\mu(z) > -2n-2 \geqslant \mu(y)$ and the first mentioned lemma imply

$$-2n-1 = \mu(x) = \mu(yz) = \min\{\mu(z), \mu(y)\} = \mu(y) \leqslant -2n-2,$$

a contradiction. Therefore, we have $\mu(z) \leqslant -2n$. Note that z is even by the previous passage, so there is $m \geqslant n$ such that $\mu(z) = -2m$. If we apply ς^{m-n}, then $\varsigma^{m-n}(z) \in Z^*$ with $\mu\left(\varsigma^{m-n}(z)\right) = -2n$.

Next, we show that for all $x \in Y$ there exist $y \in Y$ and $z \in Z$ such that $x = yz$ and y is either trivial or odd.
If $\mu(x) \geqslant 0$, then $x \in Y_{0,\infty} \leq Z$ and we choose $z = x$ and $y = 1$. Else we assume that $\mu(x) = -n$ for some $n \geqslant 1$ and perform an induction on n. If $n = 1$ is odd, then we choose $y = x$ and $z = 1$.
Now, let $\mu(x) = -n-1$ and assume that the statement holds for n. If n is even, then $\mu(x)$ is odd and we choose $y = x$ and $z = 1$, again. If n is odd, then there exists $z \in Z^*$ with $\mu(z) = -n-1$ by the preceding passage. By T-invariance of Z and the irreducibility of Ξ we may choose z such that $z_{\mu(x)} = x_{\mu(x)}^{-1}$, and thus $\mu(xz) > \mu(x)$ by Corollary 3.32 (ii) and Lemma 3.33 (iv). We infer the existence of $y \in Y$ and $u \in Z$ with $xz = yu$ and $y = 1$ or $\mu(y)$ odd by our induction hypothesis. Therefore, we have $x = yuz^{-1}$ with $uz^{-1} \in Z$, and the claim follows.

Let $n \geqslant 1$ and let K_n denote the kernel of the surjective endomorphism $\bar{\varsigma}^n : Y/Z \to Y/Z : yZ \mapsto \varsigma^n(y)Z$. We claim that $K_n = Y_{1-2n,\infty}Z/Z$. For $x \in Y_{1-2n,\infty}$ we observe $\varsigma^n(x) \in Y_{1,\infty} \leq Z$, and thus $Y_{1-2n,\infty}Z/Z \subseteq K_n$. For the reverse inclusion let $xZ \in K_n$ for $x \in Y$. Then there are $y \in Y$ and $z \in Z$ with $x = yz$ and y trivial or odd by the recent observation. As $\varsigma^n(Z) \subseteq Z$, we have

$$Z = \bar{\varsigma}^n(xZ) = \varsigma^n(x)Z = (\varsigma^n(y)\varsigma^n(z))Z = \varsigma^n(y)Z,$$

and therefore $\varsigma^n(y) \in Z$. If $y \neq 1$, then $\mu(\varsigma^n(y)) = \mu(y) + 2n$ is odd; hence $\mu(\varsigma^n(y)) \geqslant 1$ by the fact that words in $z \in Z^*$ with $\mu(z) < 0$ are even. We infer $\mu(y) \geqslant 1-2n$, i.e. $y \in Y_{1-2n,\infty}$, and $xZ = yzZ \in Y_{1-2n,\infty}Z/Z$. This proves equality. In particular, we have shown that $Y_{1-2n,\infty}Z/Z \trianglelefteq Y/Z$ for all $n \geqslant 1$, especially $Y_{-1,\infty}Z/Z \trianglelefteq Y/Z$.

Note that the projection map $\pi_{-1} : Y_{-1,\infty} \to X_{-1}$ is an epimorphism by Corollary 3.32 with kernel $Y_{0,\infty} = Y_{-1,\infty} \cap Z$. Proposition 3.26 (iii) and (i) imply that

$$Y_{-1,\infty}Z/Z \cong_T Y_{-1,\infty}/(Y_{-1,\infty} \cap Z) \cong_T X_{-1}.$$

As T acts irreducibly on X_{-1}, the rooted group X_{-1} is finitely T-generated by one non-trivial element. The same holds for $Y_{-1,\infty}Z/Z$ by Corollary 3.28 which is therefore non-trivial.

We claim that $Y_{-1,\infty}Z/Z \leq Z(Y/Z)$. Let $x \in Y \setminus Z$ and let W denote the T-subgroup $\langle xZ, Y_{-1,\infty}Z/Z\rangle_T$ of Y/Z. Since X is T-locally nilpotent, $Y \leq_T X$ and thus Y/Z share the same property by Lemma 4.4. As $Y_{-1,\infty}Z/Z$ is finitely T-generated by any non-trivial element by our previous observation, the group W is finitely T-generated, again. Hence, it is nilpotent and it has a non-trivial center (see, for instance, [14, 5.2.1]).
The group $Y_{-1,\infty}Z/Z$ is normalized by W as a normal subgroup of Y/Z, so that $Y_{-1,\infty}Z/Z \trianglelefteq W$. Therefore $Z(W) \cap Y_{-1,\infty}Z/Z \neq 1$. Since the center is characteristic in W, the intersection $Z(W) \cap Y_{-1,\infty}Z/Z$ is a T-subgroup of $Y_{-1,\infty}Z/Z$. We infer $Y_{-1,\infty}Z/Z \leq Z(W)$ by irreducibility, and thus $xZ \in C_{Y/Z}(Y_{-1,\infty}Z/Z)$. As x was arbitrary, we have $Y/Z \subseteq C_{Y/Z}(Y_{-1,\infty}Z/Z)$, and the claim follows.

More general, we have $Y_{1-2n,\infty}Z/Z \leq \zeta_n(Y/Z)$ for all $n \geqslant 1$. We argue via an induction on n. The previous claim shows that this statement is true for $n = 1$.
Let $n > 1$ and assume that the inclusion holds for n. The surjective endomorphism $\bar{\varsigma}^n$ has kernel $Y_{1-2n,\infty}Z/Z$ and maps $Y_{1-2(n+1),\infty}Z/Z = Y_{-1-2n,\infty}Z/Z$ onto $Y_{-1,\infty}Z/Z \leq Z(Y/Z)$, i.e. in shorter notion as shown before $\bar{\varsigma}^n(K_n) = \{Z\}$ and $\bar{\varsigma}^n(K_{n+1}) = K_1 \leq Z(Y/Z)$. Using Proposition 3.26 (i) and $K_n \trianglelefteq_T K_{n+1}$, we observe

$$(Y/Z)/K_n \cong_T Y/Z \quad \text{and} \quad K_{n+1}/K_n \cong_T K_1 \leq_T Z(Y/Z) \cong_T Z((Y/Z)/K_n)$$

and infer $[xK_n, yK_n] = [x,y]K_n = K_n \Leftrightarrow [x,y] \in K_n$ for all $x \in K_{n+1}$ and $y \in Y/Z$. Since $K_n \leq \zeta_n(Y/Z)$ by our induction hypothesis, we conclude $K_{n+1} \leq \zeta_{n+1}(Y/Z)$.

Now, observe the equality $Y = Y_{-\infty,\infty} = \bigcup_{n\in\mathbb{N}} Y_{1-2n,\infty}$ which implies that Y/Z is the union of all $Y_{1-2n,\infty}Z/Z$ or of all $\zeta_n(Y/Z)$.
If $Y_{1-2n,\infty}Z/Z \leq Z(Y/Z)$ for all $n \in \mathbb{N}$, then Y/Z is contained in its center and abelian. This yields $\delta(Y) \leq Z < Y$ by Lemma 1.11 via the canonical projection $Y \to Y/Z$ with kernel Z, a contradiction.
Otherwise there is $m \geqslant 2$ such that $Y_{1-2m,\infty}Z/Z \not\leq Z(Y/Z)$ and $\zeta_m(Y/Z) \not\leq Z(Y/Z)$ by the preceding passage. Suppose that $\zeta_2(Y/Z) = Z(Y/Z)$. Then

$$\begin{aligned}\zeta_3(Y/Z) &= \{x \in Y/Z \mid \forall y \in Y/Z : [x,y] \in \zeta_2(Y/Z)\}\\ &= \{x \in Y/Z \mid \forall y \in Y/Z : [x,y] \in Z(Y/Z)\} = \zeta_2(Y/Z) = Z(Y/Z)\end{aligned}$$

and, inductively,

$$\begin{aligned}\zeta_{n+1}(Y/Z) &= \{x \in Y/Z \mid \forall y \in Y/Z : [x,y] \in \zeta_n(Y/Z)\}\\ &= \{x \in Y/Z \mid \forall y \in Y/Z : [x,y] \in Z(Y/Z)\} = Z(Y/Z)\end{aligned}$$

for all $n \geqslant 2$. But this is a contradiction for all $n \geqslant m$. We infer the existence of $a \in \zeta_2(Y/Z) \setminus Z(Y/Z)$.

The map $\varphi_a : Y/Z \to Z(Y/Z) : y \mapsto [y, a]$ is well-defined per definition of $\zeta_2(Y/Z)$ and the fact that $xZ = x'Z$ implies $[x, a']Z = [x', a']Z$ for $a = a'Z$ and for all $x, x' \in Y$. Using Lemma 1.12 as well as the abelianess of the center to obtain

$$\varphi_a(yy') = [yy', a] = [y, a]^{y'}[y', a] = [y, a][y', a] = \varphi_a(y)\varphi_a(y')$$

for all $y, y' \in Y/Z$, we see that φ_a is a homomorphism. There is $b \in Y/Z$ such that $[b, a] \neq Z$ by $a \notin Z(Y/Z)$; hence $\ker(\varphi_a) \neq Y/Z$ and φ_a is non-trivial. On the one hand we therefore have $\delta(Y/Z) \leq \ker(\varphi_a) < Y/Z$ by Lemma 1.11, since $Z(Y/Z)$ is abelian. On the other hand we observe $\delta(Y/Z) = Y/Z$ by $\delta(Y/Z) = (\delta(Y)Z)/Z$, which follows by Lemma 1.16, and our assumption $\delta(Y) = Y$, a contradiction. Thus, we infer $\delta(Y) < Y$. This closes the proof. □

By applying the proposition above to $Y = X$, we see $\delta(X) < X$.

Lemma 6.2. *Let G be a T-group and $N \trianglelefteq G$ be T-invariant. If N and G/N are finitely T-generated by m resp. n elements, then G is finitely T-generated by $m + n$ elements. Particularly, G is then the T-span of the T-generators of N together with a set of representatives of the T-generators of G/N.*

Proof. Since G and N are T-invariant, the quotient G/N is a well-defined T-group. Let $n_1, \ldots, n_m$ resp. $g_1, \ldots, g_n$ be elements of G such that $N = \langle n_1, \ldots, n_m\rangle_T$ and $G/N = \langle g_1N, \ldots, g_nN\rangle_T$. Let M and H denote the sets containing the T-generators of N and G/N, respectively. Let $g \in G$. Then $g \in gN$ and the coset can be written as a product $\prod_{i=1}^k h_iN$ for some $k \in \mathbb{N}$ and $h_i \in T(H) \cup T(H)^{-1}$ for all $1 \leqslant i \leqslant k$ by Lemma 3.7. We observe

$$gN = \prod_{i=1}^{k} h_iN = \left(\prod_{i=1}^{k} h_i\right) N \;\Leftrightarrow\; g^{-1} \cdot \prod_{i=1}^{k} h_i \in N.$$

Thus, there exist $l \in \mathbb{N}$ and $m_j \in T(M) \cup T(M)^{-1}$ for all $1 \leqslant j \leqslant l$ such that

$$g^{-1} \cdot \prod_{i=1}^{k} h_i = \prod_{j=1}^{l} m_j \;\Leftrightarrow\; g = \prod_{i=1}^{k} h_i \cdot \left(\prod_{j=1}^{l} m_j\right)^{-1} = \prod_{i=1}^{k} h_i \cdot \prod_{j=1}^{l} m_{l-j+1}^{-1}.$$

The factors on the right-hand side are contained in $T(M\cup H)\cup T(M\cup H)^{-1}$; so that g lies in $\langle n_1, \ldots, n_m, g_1, \ldots, g_n\rangle_T \subseteq G$. This proves the claim. □

The following fact is a combination of adapted versions of Lemma 8.4 from [9] and 5.2.6 of [14]:

Proposition 6.3. *Let G be a nilpotent T-group. If $G_{ab} = G/\delta(G)$ is finitely T-generated, then so is G. In particular, if $M \subseteq G$ is finite such that $G_{ab} = \langle m\delta(G) \mid m \in M\rangle_T$, then $G = \langle M\rangle_T$.*

Proof. We do the proof by induction on the nilpotent class $nc(G)$ of G.

If $nc(G) = 0$, then G is trivial. If $nc(G) = 1$, then $\{1\} = \gamma_1(G) = \delta(G)$ and $G \cong_T G/\delta(G)$. In either case the group G is finitely T-generated.

Now, let G be a nilpotent T-group of nilpotent class $n+1$ and let G_{ab} be finitely T-generated, say by the image of a finite set $M \subseteq G$ under the canonical projection $\rho : G \to G_{ab}$.
Let $N := \gamma_n(G)$ denote the last non-trivial term of the lower central series for G. Then N is characteristic in G, thus a T-invariant and normal subgroup, and G/N is a T-group. Furthermore, we observe $N \trianglelefteq \delta(G)$ as term in the lower central series and $N \leq Z(G)$ as $\gamma_{n+1}(G) = \{1\}$.
On the one hand we see

$$\gamma_n(G/N) = \gamma_n(G)N/N = NN/N = \{N\}$$

by Lemma 1.18 (ii) as the canonical projection $G \to G/N$ is an epimorphism. On the other hand, since G is nilpotent, we have $\gamma_n(G) < \gamma_{n-1}(G)$, such that

$$\gamma_n(G/N) = \gamma_n(G)N/N < \gamma_{n-1}(G)N/N = \gamma_{n-1}(G/N).$$

Thus, the quotient G/N is nilpotent of class n. Furthermore, we observe $\delta(G/N) = \delta(G)N/N = \delta(G)/N$ by Lemma 1.16 (ii), so that Proposition 3.26 (iii) yields

$$(G/N)_{ab} = (G/N)/\delta(G/N) = (G/N)/(\delta(G)/N) \cong_T G_{ab}.$$

As G_{ab} is finitely T-generated by the set $\{m\delta(G) \mid m \in M\}$, an application of Corollary 3.28 yields that $(G/N)_{ab}$ is finitely T-generated by the set $\{mN\delta(G/N) \mid m \in M\}$ which is the pre-image of the T-generating set of G_{ab} under the T-isomorphism given in the stated proposition. Our induction hypothesis now implies that G/N is also finitely T-generated by the set $\{mN \mid m \in M\}$.
Let $H := \langle T(M)\rangle$ be the T-subgroup of G generated by the T-orbit of M. Then we have $G = HN$. Since $N \leq Z(G)$, we infer $H \trianglelefteq G$ and that the quotient G/H is an abelian T-group. On the one hand we have the natural equality $H = \ker(\rho')$ for the canonical projection $\rho' : G \to G/H$. On the other hand we see $N \leq \ker(\rho') = H$ which follows by $N \leq \delta(G)$, the abelianess of G/H, and Lemma 1.11. We conclude $G \leq \ker(\rho') = H$ and that $G = H$ is finitely T-generated by the set M. □

With the following lemma we come closer to our sectional goal.

Lemma 6.4. *([9, Lemma 8.5]) There exists $k \in \mathbb{N}$ such that $X/\delta_k(X)$ is not finitely T-generated.*

Proof. If G is a group, $N \trianglelefteq G$, and $\pi : G \to G/N$ the canonical projection, then an application of Lemma 1.16 (ii) as above yields

$$\delta_l(G/N) = \delta_l(\pi(G)) = \pi(\delta_l(G)) = (\delta_l(G)N)/N$$

for all $l \in \mathbb{N}$, so that the derived length of a quotient of a solvable group G is bounded above by the one of G (see also [14, 5.1.1]).

We assume now the converse of the claim, i.e. that $G_k = X/\delta_k(X)$ is finitely T-generated for all $k \in \mathbb{N}$. Then there exists a finite set $M \subseteq X$ such that $G_1 = X_{ab} = \langle m\delta(X) \mid m \in M\rangle_T$.

Let k be arbitrary in the following. The quotient G_k is T-locally nilpotent by Lemma 4.4 which implies that G_k is nilpotent under our assumption, and thus solvable. Moreover, with the equality

$$\delta(G_k) = \delta(X)\delta_k(X)/\delta_k(X) = \delta(X)/\delta_k(X)$$

we infer similar to the proof above that

$$(G_k)_{ab} = G_k/\delta(G_k) = G_k/\left(\delta(X)/\delta_k(X)\right) \cong_T G_1$$

by Proposition 3.26 (iii). An application of Corollary 3.28 yields $(G_k)_{ab} = \langle m'\delta(G_k) \mid m \in M\rangle_T$, where $m' := m\delta_k(X)$ is the pre-image of $m\delta(X)$ under the T-isomorphism given in the corresponding isomorphism theorem. By the nilpotency of G_k we infer $G_k = \langle m' \mid m \in M\rangle_T$ by Proposition 6.3, so that $X = H\delta_k(X)$ for $H := \langle M\rangle_T$. As X is T-locally nilpotent, H is nilpotent, and thus solvable with a finite derived length, say $d \in \mathbb{N}$. We have

$$G_k = (H\delta_k(X))/\delta_k(X) \cong H/(H \cap \delta_k(X))$$

by Theorem 1.8 (I2). Since the quotient on the right-hand side has derived length at most d by the initial observation, the same is true for G_k.

Since X is not finitely T-generated by Remark 3.6, its shift-invariant and normal T-subgroup $\delta_k(X)$ is not finitely T-generated by Lemma 6.2. In particular, it is non-trivial. Therefore, Proposition 6.1 implies that $\delta_{k+1}(X) = \delta(\delta_k(X)) < \delta_k(X)$. Hence, we have

$$\begin{aligned}\delta_k(G_{k+1}) &= (\delta_k(X)\delta_{k+1}(X))/\delta_{k+1}(X) = \delta_k(X)/\delta_{k+1}(X)\\ &\neq \{\delta_{k+1}(X)\} = \delta_{k+1}(X)/\delta_{k+1}(X) = \delta_{k+1}(G_{k+1}).\end{aligned}$$

Thus, the quotient G_{k+1} has derived length $k+1$.

As k was arbitrary, this is a contradiction to the boundedness by d for $k > d-1$. $\square$

We need three more observations for the main result of this section.

Lemma 6.5. *Let G be a group, $N \trianglelefteq G$ abelian, $H \leq G$ with $G = NH$, and $a \in N$. If there exists $m \in \mathbb{N}$ with $[a, h_1, \ldots, h_m] = 1$ for all $h_1, \ldots, h_m \in H$, then $[a, g_1, \ldots, g_m] = 1$ for all $g_1, \ldots, g_m \in G$.*

Proof. Let $a \in N$ and $g_1, \ldots, g_m \in G$. Since $G = NH$, there exist $n_i \in N$ and $h_i \in H$ such that $g_i = n_i h_i$ for all $1 \leqslant i \leqslant m$. We perform an induction on $1 \leqslant m \in \mathbb{N}$ to show that $[a, g_1, \ldots, g_m] = [a, h_1, \ldots, h_m]$.
For $m = 1$ we infer

$$[a, g_1] = [a, n_1 h_1] = [a, h_1][a, n_1]^{h_1} = [a, h_1]$$

by Lemma 1.12, since N is abelian.
For the inductive step, let $m > 1$ and $[a, g_1, \ldots, g_{m-1}] = [a, h_1, \ldots, h_{m-1}]$. As $N \trianglelefteq G$, successive commutators of a with h_i are contained in N, again. Using the same lemma together with the induction hypothesis and the abelianess of N, we conclude

$$\begin{aligned}[a, g_1, \ldots, g_m] &= [[a, g_1, \ldots, g_{m-1}], g_m] = [[a, h_1, \ldots, h_{m-1}], n_m h_m] \\ &= [[a, h_1, \ldots, h_{m-1}], h_m][[a, h_1, \ldots, h_{m-1}], n_m]^{h_m} \\ &= [a, h_1, \ldots, h_m].\end{aligned}$$

The claim directly follows. $\square$

Lemma 6.6. *Let G be a group and $a \in G$ with $[a, g_1, \ldots, g_m] = 1$ for all $g_1, \ldots, g_m \in G$ for some $m \in \mathbb{N}$. Then $a \in \zeta_m(G)$.*

Proof. We do the proof by induction on $1 \leqslant m \in \mathbb{N}$. For $m = 1$ we have $[a, g_1] = 1$ for all $g_1 \in G$, i.e. $a \in Z(G) = \zeta_1(G)$.

Now, let $m > 1$ and assume that any element $g \in G$ with $[g, g_1, \ldots, g_m] = 1$ for all $g_1, \ldots, g_m \in G$ is contained in $\zeta_m(G)$. Let $a \in G$ such that $[a, g_1, \ldots, g_{m+1}] = 1$ for all $g_1, \ldots, g_{m+1} \in G$. Then $[a, g_1] \in \zeta_m(G)$ for all $g_1 \in G$ by our induction hypothesis.
Let $\rho : G \to G/\zeta_m(G)$ denote the canonical projection. Since $[a, g_1] \in \zeta_m(G)$ for all $g_1 \in G$, we have

$$\zeta_m(G) = \rho\left([a, g_1]\right) = [\rho(a), \rho(g_1)] = [a\zeta_m(G), g_1\zeta_m(G)]$$

for all $g_1 \in G$. This implies $a\zeta_m(G) \in Z\left(G/\zeta_m(G)\right)$, and thus

$$a = \rho^{-1}\left(a\zeta_m(G)\right) \in \rho^{-1}\left(Z\left(G/\zeta_m(G)\right)\right) = \zeta_{m+1}(G). \qquad \square$$

The lemma below can, for example, be found in [2] as Theorem 2.5.

Lemma 6.7. *Let G be a group, $N \trianglelefteq G$ with $N \leqslant \zeta_m(G)$ for some $1 \leqslant m \in \mathbb{N}$, and G/N be nilpotent. Then G is nilpotent.*

Proof. Since $N \trianglelefteq \zeta_m(G)$, the Third Isomorphism Theorem 1.8 (I3) gives

$$G/\zeta_m(G) \cong (G/N) \,/\, (\zeta_m(G)/N)\,,$$

so that $G/\zeta_m(G)$ is a quotient of the nilpotent group G/N, and thus nilpotent again (cf. [14, 5.1.4]). Now, we take a look at the (a priori possibly infinite) upper central series of G, i.e.

$$\zeta_0(G) \trianglelefteq \zeta_1(G) \trianglelefteq \zeta_2(G) \trianglelefteq \ldots \trianglelefteq \zeta_m(G) \trianglelefteq \zeta_{m+1}(G) \trianglelefteq \ldots \trianglelefteq G.$$

We factor out $\zeta_m(G)$ beginning with $\zeta_m(G)$ itself up to G to obtain a series

$$\{\zeta_m(G)\} = \zeta_m(G)/\zeta_m(G) \trianglelefteq \zeta_{m+1}(G)/\zeta_m(G) \trianglelefteq \ldots \trianglelefteq G/\zeta_m(G).$$

This is indeed the upper central series for $G/\zeta_m(G)$, as the mentioned part of Theorem 1.8 does not only imply

$$\begin{aligned}(\zeta_{m+i+1}(G)/\zeta_m(G)) \,/\, (\zeta_{m+i}(G)/\zeta_m(G)) &\cong \zeta_{m+i+1}(G)/\zeta_{m+i}(G)\\ &= Z\left(G/\zeta_{m+i}(G)\right)\end{aligned}$$

for all $i \in \mathbb{N}$, but also

$$Z\left(G/\zeta_{m+i}(G)\right) \cong Z\left((G/\zeta_m(G)) \,/\, (\zeta_{m+i}(G)/\zeta_m(G))\right)$$

for all $i \in \mathbb{N}$. The nilpotency of $G/\zeta_m(G)$ implies the existence of some $n \geqslant m$ such that $G/\zeta_m(G) = \zeta_n(G)/\zeta_m(G)$. Hence, $G = \zeta_n(G)$, so that the upper central series of G is of finite length and G is nilpotent. $\square$

Now we can prove our main result of this chapter which is essentially Theorem 8.9 of our main source ([9]). To emphasize its importance, we include our assumptions on the $\mathbb{Z}$-system in its formulation.

Theorem 6.8. *Let $\Xi = (X, (X_k)_{k\in\mathbb{Z}}, \varsigma, T)$ be an irreducible and nilpotent $\mathbb{Z}$-system. Then $X_{ab} = X/\delta(X)$ is not finitely T-generated.*

Proof. For $k \in \mathbb{N}$ let

$$G_k := X/\delta_k(X) \quad\text{and}\quad H_{k+1} := (\delta_k(X))_{ab} = \delta_k(X)/\delta_{k+1}(X).$$

Then H_{k+1} is an abelian and normal subgroup of G_{k+1}. Since $G_0 = \{X\}$ is finitely T-generated, there exists $l \in \mathbb{N}$ such that G_l is finitely and G_{l+1} is not finitely T-generated by Lemma 6.4. In particular, G_l is again T-locally nilpotent, and thus nilpotent by Lemma 4.4. We further observe

$$G_{l+1}/H_{l+1} = (X/\delta_{l+1}(X)) \,/\, (\delta_l(X)/\delta_{l+1}(X)) \cong_T X/\delta_l(X) = G_l$$

by Proposition 3.26 (iii) via the T-isomorphism F induced by the map

$$f : G_{l+1} \to G_l : x\delta_{l+1}(X) \mapsto x\delta_l(X).$$

Thus, the subgroup H_{l+1} can not be finitely T-generated by Corollary 3.28 and Lemma 6.2.

Since $\delta_l(X)$ is non-trivial, shift-invariant, and of finite T-index in X, Proposition 3.42 (ii) implies that $\delta_l(X)$ is T-generated by shifts of suitable chosen elements $a, b \in \delta_l(X)^*$, i.e.

$$H_{l+1} = \left\langle \varsigma^s(a)\delta_{l+1}(X), \varsigma^t(b)\delta_{l+1}(X) \,\middle|\, s, t \in \mathbb{Z} \right\rangle_T .$$

We want to show that G_{l+1} is nilpotent in the following. Knowing that its quotient G_{l+1}/H_{l+1} is T-isomorphic to G_l, and thus also nilpotent, it suffices to show that $H_{l+1} \leqslant \zeta_m(G_{l+1})$ for some $1 \leqslant m \in \mathbb{N}$ by the previous lemma. Since $\zeta_m(G_{l+1})$ is shift- and T-invariant, it is sufficient to prove that the T-generators $a' = a\delta_{l+1}(X)$ and $b' = \delta_{l+1}(X)$ are contained in $\zeta_m(G_{l+1})$.

Since G_l is finitely T-generated, there exists a finite set $M \subseteq X$ disjoint to $\delta_l(X)$ such that its image under the canonical projection $\rho : X \to G_l : x \mapsto x\delta_l(X)$ is T-generating G_l. The set $M \cup \{a, b\}$ is finite, so its T-span, let us denote it by U, is a finitely T-generated subgroup of X, and thus nilpotent by our premise. Hence, there is $n \in \mathbb{N}$ such that $\gamma_n(U) = \{1\}$, i.e. for all $x_1, \ldots, x_n \in U$ we have

$$[a, x_1, \ldots, x_n] = 1 = [b, x_1, \ldots, x_n].$$

In particular, this is true for all $x_1, \ldots, x_n \in \langle M \rangle_T$.

As $G_l = \langle \rho(M) \rangle_T$ holds, the quotient G_{l+1}/H_{l+1} is (finitely) T-generated by the set $F^{-1}(\rho(M))$. Let

$$N := \langle m\delta_{l+1}(X) \mid m \in M \rangle_T = (\langle M \rangle_T \delta_{l+1}(X))/\delta_{l+1}(X)$$

denote the T-subgroup of G_{l+1} generated by the representatives of the generators in $F^{-1}(\rho(M))$, where the equality follows by the T-invariance of $\delta_{l+1}(X)$. Then we have $G_{l+1} = H_{l+1}N$. By the commutator relations above, we infer

$$[a', x_1\delta_{l+1}(X), \ldots, x_n\delta_{l+1}(X)] = [a, x_1, \ldots, x_n]\delta_{l+1}(X) = \delta_{l+1}(X)$$

and, similarly, $[b', x_1\delta_{l+1}(X), \ldots, x_n\delta_{l+1}(X)] = \delta_{l+1}(X)$ for all $x_1, \ldots, x_n \in \langle M\rangle_T$. But then

$$[a', x'_1, \ldots, x'_n] = \delta_{l+1}(X) = [b', x'_1, \ldots, x'_n]$$

for all $x'_1, \ldots, x'_n \in N$. Since H_{l+1} is abelian, Lemma 6.5 implies

$$[a', g_1, \ldots, g_n] = \delta_{l+1}(X) = [b', g_1, \ldots, g_n]$$

for all $g_1, \ldots, g_n \in G_{l+1}$. Hence we have $a', b' \in \zeta_n(G_{l+1})$ by Lemma 6.6. We infer that G_{l+1} is nilpotent.

As G_{l+1} is not finitely T-generated, the abelianization $(G_{l+1})_{ab}$ is not finitely T-generated by Proposition 6.3. We conclude, since $\delta(G_{l+1}) = (\delta(X)\delta_{l+1}(X))/\delta_{l+1}(X) = \delta(X)/\delta_{l+1}(X)$ by Lemma 1.16 (i), that

$$X_{ab} \cong_T G_{l+1}/\left(\delta(X)/\delta_{l+1}(X)\right) = (G_{l+1})_{ab}$$

is not finitely T-generated by a reapplication of Corollary 3.28. $\square$

We may say in short that X has *infinite T-abelianization*.

Example 6.9. Let $\Xi = (X, (X_k)_{k\in\mathbb{Z}}, \varsigma, T)$ be an irreducible and nilpotent $\mathbb{Z}$-system. Its characteristic T-subgroup $\delta(X)$ has infinite T-index by the preceding theorem, so that at least one of the sets $\delta(X)_{even}$ and $\delta(X)_{odd}$ is empty by Proposition 3.41.
If there is $f \in \mathrm{Aut}(X)$ such that f normalizes T, $f^2 = \varsigma$, and $f(X_k) = X_{k+1}$ for all $k \in \mathbb{Z}$, then any non-trivial, normal T-subgroup Y of X invariant under f (and thus also shift-invariant) contains even and odd words. In particular, the minimal widths ω_0 and ω_1 then coincide and Y is generated by all $f^l(a')$ with $l \in \mathbb{Z}$ and $a' \in \langle a\rangle_T$ for an element $a \in Y^*$ of minimal width by Corollary 3.44 and Lemma 3.46.
Hence, if $\delta(X)$ were non-trivial, it would contain even and odd words, contradicting its infinite T-index. So $\delta(X) = \{1\}$ and X is abelian.

Chapter 7

Main result

In this final chapter, let $\Xi = (X, (X_k)_{k\in\mathbb{Z}}, \varsigma, T)$ denote an irreducible and nilpotent $\mathbb{Z}$-system. So we can utilize all previous results without restrictions. Some lemmata and the single proposition we formulate along the way towards the main result of this thesis are, up to some exceptions, based on the central paper by Grüninger, Horn, and Mühlherr ([9]).

First, let us take a closer look at a normal subgroup N of an arbitrary group G for the upcoming commutator computations.

Lemma 7.1. *([9, Lemma 9.1]) Let $N \trianglelefteq G$ and $g, g' \in G$ and $n, n' \in N$. Then*

$$[nn', g] \in [n, g][n', g][N, G, G] \quad \text{and} \quad [n, gg'] \in [n, g'][n, g][N, G, G].$$

Proof. By Lemma 1.12, we have

$$[nn', g] = [n, g]^{n'}[n', g] = [n, g][n, g]^{-1}[n, g]^{n'}[n', g] = [n, g][n, g, n'][n', g].$$

Using the fact that $N \trianglelefteq G$, we observe

$$[m, a, b]^c = [m^c, a^c, b^c] \in [N, G, G]$$

for all $m \in N$ and $a, b, c \in G$, and see $[N, G, G] \trianglelefteq G$. Hence we get

$$\begin{aligned}[n, g][n, g, n'][n', g] &= [n, g][n', g][n', g]^{-1}[n, g, n'][n', g] \\ &\in [n, g][n', g][N, G, G].\end{aligned}$$

We obtain

$$[n, gg'] = [n, g'][n, g]^{g'} = [n, g'][n, g][n, g, g'] \in [n, g'][n, g][N, G, G]$$

even faster. $\square$

We need another useful observation for the important proposition afterwards:

Lemma 7.2. *Let $Y \trianglelefteq_T X$ be non-trivial, shift-invariant, and of infinite T-index and let $s \in Y^*$ be an element of minimal width whose shifts generate Y as a T-group. Then*

$$[Y, X] = \left\langle \varsigma^k([s', x_N]) \,\middle|\, k, N \in \mathbb{Z}, s' \in \langle s \rangle_T, x_N \in X_N \right\rangle^X .$$

Proof. As X normalizes Y, we have $[Y, X] \trianglelefteq X$. Since $[Y, X]$ is a normal and shift-invariant subgroup of X, we observe

$$[Y, X] \supseteq \left\langle \varsigma^k([s', x_N]) \,\middle|\, k, N \in \mathbb{Z}, s' \in \langle s \rangle_T, x_N \in X_N \right\rangle^X .$$

To see that $[Y, X]$ is indeed the given normal closure, let $[y, x] \in [Y, X]$ and assume without loss of generality that $x \neq 1$. Then there exist $r \in \mathbb{N}$, $k_1, \ldots, k_r \in \mathbb{Z}$, and $s_1, \ldots, s_r \in \langle s \rangle_T$ such that $y = \varsigma^{k_1}(s_1) \cdot \ldots \cdot \varsigma^{k_r}(s_r)$ by Corollary 3.44 as well as $m \leqslant n \in \mathbb{Z}$ and $x_m, \ldots, x_n \in X$ such that $x = x_m \cdot \ldots \cdot x_n$ is the normal form of x. By applying both parts of Lemma 1.12 successively on $[\varsigma^{k_1}(s_1) \cdot \ldots \cdot \varsigma^{k_r}(s_r), x_m \cdot \ldots \cdot x_n]$, we see that $[y, x]$ is the product of elements of the form $\varsigma^k([s', x_N])^{x'}$, where $k, N \in \mathbb{Z}$, $s' \in \langle s \rangle_T$, $x_N \in X_N$, and $x' \in X$. Hence, the normal closure on the right-hand side above contains $[y, x]$. Since $[Y, X]$ is generated by commutators of that form, we have

$$[Y, X] \subseteq \left\langle \varsigma^k([s', x_N]) \,\middle|\, k, N \in \mathbb{Z}, s' \in \langle s \rangle_T, x_N \in X_N \right\rangle^X ,$$

and thus equality. □

Recall for the upcoming proof that the rooted groups of Ξ are abelian by Lemma 3.10 and that the commutator relations take the more pleasant form of Corollary 3.24.

Proposition 7.3. *([9, Lemma 9.2]) Let $Y \trianglelefteq_T X$ be shift-invariant and of infinite T-index. Then $[Y, X, X] = [Y, X]$.*

Proof. If $Y = \{1\}$, then the equality clearly holds. Therefore, let Y be non-trivial in the following.

We attend to the case where $Y_{odd} = \varnothing$, i.e. for all $y \in Y^*$ we have $\mu(y) \in 2\mathbb{N}$. Again, the case $Y_{even} = \varnothing$ follows by nearly the same reasoning while looking at Ξ^{pc} and the details are left out.

Since Y is normal in X, we infer $[Y, X] \leq Y$, and the inclusion $[Y, X, X] \subseteq [Y, X]$ surely holds.

Let $s \in Y^*$ be an element of minimal width whose shifts generate Y as a T-group (see Proposition 3.42 (i)). By shift-invariance we may assume that $s \in Y_{0,\omega_0-1}$. For the remaining inclusion $[Y,X] \subseteq [Y,X,X]$ it suffices to show, since $[Y,X,X]$ is a shift-invariant and normal subgroup of X, that $[s', x_N] \in [Y,X,X]$ for all $s' \in \langle s \rangle_T$ and for all $x_N \in X_N$ with $N \in \mathbb{N}$ by Lemma 7.2.

Let s' be an element in the group generated by the T-orbit of s and x_N an element of X_N. First, we perform an induction on $N \geqslant 0$.
If $0 \leqslant N \leqslant \omega_0 - 1$, then

$$[s', x_N] \in Y_{\min\{0,N\}+1,\max\{\omega_0-1,N\}-1} = Y_{1,\omega_0-2} = \{1\} \subseteq [Y,X,X]$$

by the corollary we refereed to beforehand.

For the induction step, let $N \geqslant \omega_0$ and assume that $[s'', x_M] \in [Y,X,X]$ for all $s'' \in \langle s \rangle_T$ and for all $x_M \in X_M$ with $M < N$. Let $[s', x_N]$ be non-trivial. Using the fact that $[Y,X] \leq Y$ together with Corollary 3.44 and Proposition 3.50, we can write the commutator in sorted form, i.e.

$$\varsigma^{j_1}(s_1) \cdot \ldots \cdot \varsigma^{j_r}(s_r) = [s', x_N] \in Y_{2,N-1}$$

with some integers $0 < j_1 < \ldots < j_r \leqslant \frac{N-\omega_0}{2}$ and elements $s_i \in \langle s \rangle_T$ for all $1 \leqslant i \leqslant r$, where $1 \leqslant r \in \mathbb{N}$ and $s_1 \neq 1 \neq s_r$. The strict lower bound is due to the assumption $Y_{odd} = \varnothing$ and the upper bound of these integers is chosen this way because

$$\nu(\varsigma^{j_i}(s_i)) = \omega(\varsigma^{j_i}(s_i)) + \mu(\varsigma^{j_i}(s_i)) - 1 = \omega_0 + 2j_i - 1 \leqslant N - 1$$

is equivalent to $j_i \leqslant \frac{N-\omega_0}{2}$ for all $1 \leqslant i \leqslant r$. For now, let $i \neq 1$. We have

$$N + 2(j_1 - j_i) = N - 2(j_i - j_1) < N.$$

Let $x'_N \in X_N$ be arbitrary. By our induction hypothesis used on $s_i \in \langle s \rangle_T$ and $\varsigma^{j_1-j_i}(x'_N) \in X_{N-2(j_i-j_1)}$ together with the shift-invariance of $[Y,X,X]$ we infer

$$[\varsigma^{j_i}(s_i), \varsigma^{j_1}(x'_N)] = \varsigma^{j_i}([s_i, \varsigma^{j_1-j_i}(x'_N)]) \in [Y,X,X]$$

for all $2 \leqslant i \leqslant r$. Repeatedly using the first statement in Lemma 7.1 and then the just now stated observation, we compute

$$\begin{aligned}
[[s', x_N], \varsigma^{j_1}(x'_N)] &= [\varsigma^{j_1}(s_1) \cdot \ldots \cdot \varsigma^{j_r}(s_r), \varsigma^{j_1}(x'_N)] \\
&\in [\varsigma^{j_1}(s_1), \varsigma^{j_1}(x'_N)][\varsigma^{j_2}(s_2) \cdot \ldots \cdot \varsigma^{j_r}(s_r), \varsigma^{j_1}(x'_N)][Y, X, X] \\
&= [\varsigma^{j_1}(s_1), \varsigma^{j_1}(x'_N)][\varsigma^{j_2}(s_2), \varsigma^{j_1}(x'_N)] \\
&\quad \cdot [\varsigma^{j_3}(s_3) \cdot \ldots \cdot \varsigma^{j_r}(s_r), \varsigma^{j_1}(x'_N)][Y, X, X] \\
&= \ldots \\
&= [\varsigma^{j_1}(s_1), \varsigma^{j_1}(x'_N)] \cdot \ldots \cdot [\varsigma^{j_r}(s_r), \varsigma^{j_1}(x'_N)][Y, X, X] \\
&= [\varsigma^{j_1}(s_1), \varsigma^{j_1}(x'_N)] \cdot \ldots \cdot [\varsigma^{j_{r-1}}(s_{r-1}), \varsigma^{j_1}(x'_N)][Y, X, X] \\
&= \ldots \\
&= [\varsigma^{j_1}(s_1), \varsigma^{j_1}(x'_N)][Y, X, X].
\end{aligned}$$

As $[[s', x_N], \varsigma^{j_1}(x'_N)] \in [Y, X, X]$, we conclude

$$\varsigma^{j_1}([s_1, x'_N]) = [\varsigma^{j_1}(s_1), \varsigma^{j_1}(x'_N)] \in [Y, X, X].$$

We infer $[s_1, x'_N] \in [Y, X, X]$ by shift-invariance.

Since x'_N was arbitrary, the element s_1 is contained in the set

$$U := \{x \in X \mid \forall z \in X_N : [x, z] \in [Y, X, X]\}.$$

We show that U is, in fact, a T-subgroup of X. Let $z \in X_N$ be arbitrary. Then $1 \in U$ as $[1, z] = 1 \in [Y, X, X]$. Furthermore, if $x, y \in U$, then

$$[x^{-1}, z] = x[z, x]x^{-1} = [z, x]^{x^{-1}} \in [Y, X, X]^{x^{-1}} = [Y, X, X]$$

and, by Lemma 1.12,

$$[xy, z] = [x, z]^y [y, z] \in [Y, X, X],$$

so that $x^{-1}, xy \in [Y, X, X]$. Hence, U is a subgroup of X. It is T-invariant since $t^{-1}.X_N = X_N$ and

$$[t.x, z] = t.[x, t^{-1}.z] \in t.[Y, X, X] = [Y, X, X]$$

for all $t \in T$, $x \in U$, and $z \in X_N$. As a T-subgroup containing s_1, U fully contains the T-subgroup $\langle s_1 \rangle_T$ and therefore the element s and its T-span $\langle s \rangle_T$ by Lemma 3.46. Particularly, this implies $[s', x_N] \in [Y, X, X]$ for all $s' \in \langle s \rangle_T$ and $x_N \in X_N$. Thus, we have proven that $[s', x_N] \in [Y, X, X]$ for all $s' \in \langle s \rangle_T$ and for all $x_N \in X_N$ with $N \geqslant 0$.

A similar induction shows $[s', x_N] \in [Y, X, X]$ for all $s' \in \langle s \rangle_T$ and for all $x_N \in X_N$ with $N < 0$. We only outline some parts.

In the base case $N = -1$ we observe $[s', x_N] \in Y_{0,\omega_0-2} = \{1\} \subseteq [Y, X, X]$. For the inductive step we use the equivalence

$$[s', x_N] \in [Y, X, X] \Leftrightarrow [x_N, s'] = [s', x_N]^{-1} \in [Y, X, X]$$

as well as an analogous sorted expression

$$\varsigma^{j_1}(s_1) \cdot \ldots \cdot \varsigma^{j_r}(s_r) = [x_N, s'] \in Y_{N+1,\omega_0-2}$$

with $\frac{N+1}{2} \leqslant j_1 < \ldots < j_r < 0$ and elements $s_i \in \langle s \rangle_T$ for all $1 \leqslant i \leqslant r$, where $1 \leqslant r \in \mathbb{N}$ and $s_1 \neq 1 \neq s_r$. By using $[Y, X, X] = [X, [Y, X]]$, the second statement of Lemma 7.1, and the induction hypothesis on the elements $[\varsigma^{j_r}(x'_N), \varsigma^{j_i}(s_i)] = \varsigma^{j_i}([\varsigma^{j_r - j_i}(x'_N), s_i])$ for all $1 \leqslant i \leqslant r-1$, we infer for the largest integer j_r that

$$[\varsigma^{j_r}(x'_N), [x_N, s']] \in [\varsigma^{j_r}(x'_N), \varsigma^{j_r}(s_r)][Y, X, X]$$

for arbitrary $x'_N \in X_N$. Following the arguments in the induction for $N \geqslant 0$, we infer $[x'_N, s_r] \in [Y, X, X]$, and therefore $[x_N, s'] = [s', x_N]^{-1} \in [Y, X, X]$ for all $s' \in \langle s \rangle_T$ and for all $x_N \in X_N$ with $N < 0$.

This proves $[Y, X] \subseteq [Y, X, X]$, and finally equality. □

Definition 7.4. A T-subgroup Y of X will be called *lower shift-invariant* if $\varsigma^{-1}(Y) \subseteq Y$. Equally, we will call Y *higher shift-invariant* if $\varsigma(Y) \subseteq Y$.

Observe that a T-subgroup is shift-invariant if and only if it is lower and higher shift-invariant. An example for a higher shift-invariant T subgroup is $Z = \langle Y_{k,\infty} \rangle^X$, $k \in \mathbb{Z}$, as seen in the proof of Proposition 6.1.

With this definition we state an adapted version of Lemma 9.5 of [9].

Lemma 7.5. *Let $Y \trianglelefteq_T X$ be shift-invariant with $[Y, X] \neq \{1\}$. Assume further that Y is of infinite T-index in X and T-generated by the shifts of $s \in Y^*$ with $k := \nu(s)$. Define $N := Y_{-\infty,k}$ and $N_0 := [N, X_{-\infty,k}]$. Then N_0 is a non-trivial, proper, and normal T-subgroup of N. Furthermore, both groups are lower shift-invariant and their quotient N/N_0 is finitely T-generated.*

Proof. First we show $N_0 \trianglelefteq_T N$. As Y and $X_{-\infty,k}$ are T-invariant, so are N and, being a commutator group of two T-invariant groups, N_0. Since Y is normal in X, we have $X_{-\infty,k} \subseteq N_X(Y)$ and therefore $N \trianglelefteq X_{-\infty,k}$. This implies $N_0 = [N, X_{-\infty,k}] \leq_T N$. Furthermore, N and $X_{-\infty,k}$ are invariant under conjugation by elements of N; hence, N_0 is a normal T-subgroup of N.
We have $N_0 \leq \delta(X_{-\infty,k}) \leq X_{-\infty,k-1}$ by the corollary of Proposition 3.22, so that $s \in N$ but $s \notin N_0$. Thus, $N_0 \triangleleft_T N$ and N/N_0 is non-trivial.

To see that $N_0 \neq \{1\}$, we show that $\{1\} \neq [Y, X] = \bigcup_{n\in\mathbb{N}} \varsigma^n(N_0)$. Indeed, by shift-invariance of Y, we observe

$$\begin{aligned}\bigcup_{n\in\mathbb{N}} \varsigma^n(N_0) &= \bigcup_{n\in\mathbb{N}} [\varsigma^n(Y_{-\infty,k}), \varsigma^n(X_{-\infty,k})] \\ &= \bigcup_{n\in\mathbb{N}} [Y_{-\infty,k+2n}, X_{-\infty,k+2n}] \subseteq [Y, X].\end{aligned}$$

For the other inclusion it suffices to show that $[y, x] \in \bigcup_{n\in\mathbb{N}} \varsigma^n(N_0)$ for all $y \in Y^*$ and $x \in X^*$, since the union of the shifts of N_0 is a group. In fact, the union is T-invariant by (M5) and the T-invariance of N_0, but this is not needed here. The subgroup properties are readily verified by using the fact that a shift $\varsigma^n(N_0)$ contains all lower shifts of N_0, i.e. the shifts $\varsigma^m(N_0)$ for $m \leqslant n$. Set $l := \max\{\nu(y), \nu(x)\}$. Then either $l \leqslant k$ and

$$[y, x] \in [Y_{-\infty,l}, X_{-\infty,l}] \subseteq N_0$$

or $k < l$ and there exists $n' \in \mathbb{N}$ such that $l \leqslant k + 2n'$ and

$$[y, x] \in [Y_{-\infty,k+2n'}, X_{-\infty,k+2n'}] \subseteq \varsigma^{n'}(N_0).$$

We infer $[Y, X] = \bigcup_{n\in\mathbb{N}} \varsigma^n(N_0)$ and $\{1\} \neq N_0$ by our premise.

Since $\varsigma^{-1}(N) = Y_{-\infty,k-2} < N$, the T-group N is lower shift-invariant. Let $[y, x]$ be a commutator in the generating set of N_0. We compute

$$\varsigma^{-1}([y, x]) = [\varsigma^{-1}(y), \varsigma^{-1}(x)] \in [Y_{-\infty,k-2}, X_{-\infty,k-2}] \leq N_0,$$

hence $\varsigma^{-1}(N_0) \subseteq N_0$, and N_0 is also lower shift-invariant.

It remains to show that N/N_0 is finitely T-generated. By Lemma 3.47 and the choice of s, the lower shift-invariant subgroup N is T-generated by the shifts $\varsigma^j(s)$ with $j \leqslant 0$. Now, let $m < k$ be maximal with the property that there exists $z \in N_0^* \leq X_{-\infty,k-1}$ such that $\mu(z) = m$. We claim that $N = N_0 Y_{m+2,k}$.
Assume that this is not the case. Then there exists $a \in N \setminus N_0 Y_{m+2,k}$ with $\mu(a)$ maximal. In particular, we have $\mu(a) < m + 2$ (in fact even $\leqslant m$ by the infinite T-index of Y), as otherwise $a \in Y_{m+2,k} \subseteq N_0 Y_{m+2,k}$. Thus, since N_0 is lower shift-invariant and $\mu(a) \leqslant \mu(z)$, there is $b \in N_0$ such that $\mu(b) = \mu(a)$. Hence, both a and b are contained in $Y_{\mu(a),k}$.
Moreover, as $N \cap Y_{\mu(a),k} = Y_{\mu(a),k}$ and $N_0 \cap Y_{\mu(a),k} = (N_0 \cap Y) \cap X_{\mu(a),k}$ are non-trivial T-subgroups of X, Corollary 3.32 (ii) implies that

$$\pi_{\mu(a)}(Y_{\mu(a),k}) = X_{\mu(a)} = \pi_{\mu(a)}(N_0 \cap Y_{\mu(a),k}).$$

Hence, we may assume that $b_{\mu(a)} = a^{-1}_{\mu(a)}$, so that we get $\mu(a) < \mu(ab)$ similar to the proof of assertion (v) of Lemma 3.33. By the maximal choice of $\mu(a)$ we infer $ab \in N_0 Y_{m+2,k}$, and by $b \in N_0$ further $a \in N_0 Y_{m+2,k}$, a contradiction.

Now, observe that $N_0 \cap Y_{m+2,k}$ is normalized by $Y_{m+2,k} \subseteq N$ as $N_0 \trianglelefteq N$. An application of Proposition 3.26 (ii) yields

$$N/N_0 = N_0 Y_{m+2,k}/N_0 \cong_T Y_{m+2,k}/(N_0 \cap Y_{m+2,k}).$$

Since $Y_{m+2,k}$ is finitely T-generated by Lemma 5.4, the quotient on the right-hand side has the same property by Lemma 3.27. At last, N/N_0 is finitely T-generated by Corollary 3.28. □

The following lemma serves to outsource one specific part of the proof of the main result.

Lemma 7.6. *We have* $\delta(X) \leq C_X([X, X, X])$.

Proof. Let Y denote the subgroup $[X, X, X]$. If $Y = \{1\}$, then $C_X(Y) = X$ and there is nothing to prove. Otherwise, Y is a non-trivial, normal, shift- and T-invariant subgroup of X. In particular, this implies $\delta(X) \neq 1$ as well, since $Y = \gamma_2(X) \trianglelefteq \delta(X)$. Recall that $\delta(X)$ is of infinite T-index by the last chapter's main result Theorem 6.8. Thus, we can apply Proposition 7.3 to $\delta(X)$ and infer

$$Y = [X, X, X] = [\delta(X), X] = [\delta(X), X, X] = [X, X, X, X] = [Y, X].$$

Moreover, as $Y \trianglelefteq \delta(X)$ and $\delta(X)$ is of infinite T-index, also Y is of infinite T-index by Corollary 3.29. Hence, all requirements of Proposition 5.10 are fulfilled and there exists $m \in \mathbb{Z}$ such that $Y_{-\infty,m} \trianglelefteq X$ or $Y_{m,\infty} \trianglelefteq X$. We may assume that the second case holds. Furthermore, the group Y is abelian by Corollary 5.12.

If $\delta(X)$ satisfies one of the If-conditions in Lemma 5.9, then it is also abelian and naturally centralizes Y. If not, then using Corollary 5.7, there exists $i \in \mathbb{Z}$ such that $\delta(X) = \langle \delta(X)_{0,i} \rangle^X$. Since $d^x = d[d, x]$ holds for all $d \in \delta(X)_{0,i}$ and all $x \in X$, we get

$$\delta(X) = \langle \delta(X)_{0,i} \rangle^X \subseteq \delta(X)_{0,i}[\delta(X)_{0,i}, X] \subseteq \delta(X)_{0,i}[\delta(X), X] = \delta(X)_{0,i} Y$$

and by $\delta(X)_{0,i}, Y \subseteq \delta(X)$ the equality $\delta(X) = \delta(X)_{0,i} Y$. Applying the shift ς^j for $j \in \mathbb{Z}$ we observe $\delta(X) = \delta(X)_{2j,2j+i} Y$ for all $j \in \mathbb{Z}$.

Now, let $s \in Y^*$ be a shortest word with $\mu(s) = m$ whose non-negative shifts T-generate $Y_{m,\infty}$, i.e. $Y_{m,\infty} = \left\langle \varsigma^k(s') \,\middle|\, k \in \mathbb{N}_0, s' \in \langle s \rangle_T \right\rangle$ by Lemma 3.47. Furthermore, let $j \in \mathbb{Z}$ such that $2j + i < m$. As $Y_{m,\infty} \trianglelefteq X$ by our assumption, we have $[d, s] \in Y_{m,\infty}$ for all $d \in \delta(X)_{2j,2j+i}$. But

$$\nu([d, s]) \leqslant \max\{2j + i, \nu(s)\} - 1 \leqslant \max\{m, \nu(s)\} - 1 = \nu(s) - 1$$

by Proposition 3.22, so that $\omega([d,s]) < \omega(s)$ and $[d,s] = 1$, i.e. $\delta(X)_{2j,2j+i}$ commutes with s. It follows that $\delta(X) = \delta(X)_{2j,2j+i}Y \leq C_X(\{s\})$ as $Y \leq \delta(X)$ is abelian. By T-invariance of $\delta(X)$ we see that $\delta(X)$ centralizes the T-orbit of s. Furthermore, applying Lemma 1.12 to a commutator between an element of $\delta(X)$ and a product of elements in $T(s) \cup T(s)^{-1}$ we see that $\langle s \rangle_T$ is centralized by $\delta(X)$. Since the derived subgroup is shift-invariant and Y is generated by the shifts of elements in $\langle s \rangle_T$, we analogously infer $\delta(X) \leq C_X(Y)$. □

Recall Definition 1.19 and let M be a G-module. Furthermore, let M be T-invariant. If the action of G on M is, for example, by T-equivariant automorphisms, then

$$\begin{aligned} t.[m,g] &= t.(m^{-1}(m.g)) = (t.m^{-1})(t.(m.g)) \\ &= (t.m)^{-1}((t.m).g) = [t.m,g] \in [M,G] \end{aligned}$$

for all $t \in T$, $m \in M$, and $g \in G$. Thus, $[M,G]$ is also T-invariant.

The action on any G-module M will be conjugation $m.g = m^g = g^{-1}mg$ in the following. Note that $(mn.g) = (m.g)(n.g)$ is indeed satisfied for all $m, n \in M$ and $g \in G$.
Under this action and if G and M are T-invariant, then $[M,G]$ is T-invariant again, since $t.[m,g] = [t.m, t.g] \in [M,G]$.

We will combine the important results that we prepared up until now to obtain the main result of this work which extends the main result of Grüninger, Horn, and Mühlherr (see Theorem 2.32 or the original [9, Theorem 3.4]).

Theorem 7.7. *Let $\Xi = (X, (X_k)_{k\in\mathbb{Z}}, \varsigma, T)$ be an irreducible and nilpotent $\mathbb{Z}$-system. Then X is nilpotent of class at most 2.*

Proof. As before, let Y denote the normal, shift- and T-invariant subgroup $[X,X,X]$ of X. We want to prove $Y = \{1\}$ by assuming and contradicting that Y is non-trivial in the following.

By the same reasoning given in the proof of the preceding lemma, we see that $Y \leq \delta(X)$ is of infinite T-index in X, $Y = [Y,X]$ is abelian, and there exists $m \in \mathbb{Z}$ such that $Y_{-\infty,m} \trianglelefteq X$ or $Y_{m,\infty} \trianglelefteq X$. We again may assume the second case. At last, Lemma 7.6 yields $\delta(X) \leq C_X(Y)$.

Let $s \in Y^*$ be a word of shortest width with $\mu(s) = m$ as in the proof above. We set

$$M := \varsigma(Y_{m,\infty}) = Y_{m+2,\infty} \trianglelefteq_T Y, \quad n := \nu(s), \quad \text{and} \quad Z := Y/M$$

for legibility. Note that M is a normal subgroup of X, since it is the homomorphic image of the normal subgroup $Y_{m,\infty}$, and that Z is a well-defined, abelian T-group. Since $M, Y \trianglelefteq X$, the conjugation action of X on the quotient Z is given by

$$(yM).x = x^{-1}yMx = y^x M$$

for all $y \in Y$ and $x \in X$. Hence, we can regard Z as an X- or $X_{-\infty,n}$-module (or even as an $X/\delta(X)$-module by $\delta(X) \leq C_X(Y)$).

By Corollary 3.53 there is $l \in \mathbb{Z}$ such that $Y = Y_{-\infty,l}Y_{m+2,\infty}$. By the observation following that corollary, we can choose $l = \nu(\varsigma(s)) - 2 = n$, so that $Y = Y_{-\infty,n}Y_{m+2,\infty}$. Let N denote the left-hand side group $Y_{-\infty,n} \trianglelefteq_T Y$, i.e. $Y = NM$. As $n - m + 1$ is the minimal width, we have

$$N \cap M = \begin{cases} \{1\}, & \text{if } n < m+2, \\ Y_{m+2,n} = \{1\}, & \text{if } n \geqslant m+2, \end{cases}$$

so that Y is the inner direct product of N and M. Thus, the quotient Z is T-isomorphic to N by Proposition 3.26 (ii), say via the T-isomorphism $F : N \to Z : y \mapsto yM$. Since $X_{-\infty,n}$ normalizes itself and Y, the abelian group N is normal in $X_{-\infty,n}$ and can also be viewed as an $X_{-\infty,n}$-module. As

$$F(y.x) = F(y^x) = y^x M = (yM)^x = F(y).x$$

for all $y \in N$ and $x \in X_{-\infty,n}$, we can regard both groups N and Z as isomorphic $X_{-\infty,n}$-modules. We further define

$$Z_0 := [Z, X_{-\infty,n}] \quad \text{and} \quad A := Z/Z_0.$$

The first group is an X-submodule of Z. Indeed, since $[Y, X] = Y \trianglelefteq X$ is abelian, the commutators $[y, x]$ for $y \in Y$ and $x \in X$ commute, so that the same is true for the generators

$$[yM, x] = (yM)^{-1}(yM)^x = y^{-1}My^xM = [y, x]M$$

of $[Z, X_{-\infty,n}]$, and Z_0 is abelian on the one hand. As Z is an X-module, we have $Z_0 \leq Z$. In particular, Z_0 is normal in Z. Since $\delta(X) \leq C_X(Y)$ acts trivially (via conjugation) on Y and thus on Z, we infer

$$\begin{aligned} [z, x]^w &= (z^{-1}z^x)^w = (z^{-1})^w z^{xw} = (z^w)^{-1} z^{wx[x,w]} \\ &= (z^w)^{-1}(z^{wx})^{[x,w]} = (z^w)^{-1} z^{wx} = [z^w, x] \in [Z, X_{-\infty,n}] = Z_0 \end{aligned}$$

for all $z \in Z$, $x \in X_{-\infty,n}$, and $w \in X$ on the other hand. Hence, conjugation of Z_0 by elements of X is a well-defined action. Thus, Z_0 is an X-submodule of Z and A is a well-defined X-module, again.

Note that A is T-invariant by the observation stated before the theorem and that $X_{-\infty,n}$ acts trivially on A by definition of Z_0, i.e. we have

$$(zZ_0)^x = z^x Z_0 = zz^{-1}z^x Z_0 = z[z,x]Z_0 = zZ_0$$

for all $z \in Z$ and $x \in X_{-\infty,n}$. Hence, we observe $[A,X] = [A, X_{n+1,\infty}]$.

The subgroup $N_0 := [N, X_{-\infty,n}]$ is also an $X_{-\infty,n}$-module and T-isomorphic to Z_0 via the restriction of F to N_0 with image Z_0. Therefore, the T-isomorphism F induces a T-epimorphism from N/N_0 onto $A = Z/Z_0$ by mapping yN_0 to $F(y)Z_0$, for $y \in N$, with kernel $\{N_0\}$, i.e. a T-isomorphism.

Now, the T-group A is non-trivial and finitely T-generated by combining Lemma 7.5 and Corollary 3.28. Hence, there exists $k \in \mathbb{Z}$ maximal such that $Y_{k,m+1}$ contains a (finite) set of representatives of the cosets that T-generate A. Particularly, we may assume $\mu(y_1) = k$ for at least one representative y_1 of a generator a_1. Using Corollary 3.24, we have

$$[Y_{k,m+1}, X_{n+1,\infty}] \leq Y \cap X_{k+1,\infty} = Y_{k+1,\infty}.$$

Let $a \in A$ be arbitrary with $a = (yM)Z_0$ for some $y \in Y_{k,m+1}$, and let $x \in X_{n+1,\infty}$. Then

$$[a,x] = ([y,x]M)Z_0 \subseteq (Y_{k+1,\infty}M)Z_0.$$

Hence, $[A,X] = [A, X_{n+1,\infty}] \subseteq (Y_{k+1,\infty}M)Z_0$. But $a_1 = (y_1 M)Z_0 \in (Y_{k,\infty}M)Z_0$ is not contained in $(Y_{k+1,\infty}M)Z_0$, so that $a_1 \notin [A,X]$ and $[A,X] < A$.

However, the application of Proposition 7.3 on $\delta(X)$ has yielded $[Y,X] = Y$. This in turn implies

$$\begin{aligned}[Z,X] &= \langle [yM,x] \mid y \in Y, x \in X\rangle = \langle [y,x]M \mid y \in Y, x \in X\rangle \\ &= ([Y,X]M)/M = (YM)/M = Y/M = Z.\end{aligned}$$

Moreover, using $[zZ_0, x] = [z,x]Z_0$ for $z \in Z$ and $x \in X$, we analogously compute

$$[A,X] = ([Z,X]Z_0)/Z_0 = Z/Z_0 = A.$$

But this is a contradiction to our previous strict inclusion $[A,X] < A$. Our initial assumption $Y \neq \{1\}$ must therefore be wrong. Hence, $[X,X,X]$ is trivial and X is nilpotent of class at most 2. □

In view of [9, Remark 3.5] our main theorem implies

Theorem 7.8. *([9, Theorem A]) The unipotent horocyclic group of a Moufang twin tree that yields an irreducible and nilpotent $\mathbb{Z}$-system is nilpotent of class at most 2.*

References

[1] Peter Abramenko and Kenneth S. Brown. *Buildings: Theory and Applications*. Graduate Texts in Mathematics, vol. 248. Springer Science+Business Media, LLC, 2010.

[2] Anthony E. Clement, Stephen Majewicz, and Marcos Zyman. *The Theory of Nilpotent Groups*. Birkhäuser/Springer International Publishing AG, Cham, 2017.

[3] Tom De Medts and Yoav Segev. Identities in Moufang sets. *Transactions of the American Mathematical Society*, 360(11): pp. 5831–5852, 2008.

[4] Tom De Medts and Yoav Segev. A course on Moufang sets. *Innovations in Incidence Geometry*, 9: pp. 79–122, 2009.

[5] Tom De Medts and Richard M. Weiss. Moufang sets and Jordan division algebras. *Mathematische Annalen*, 335: pp. 415–433, 2006.

[6] Reinhard Diestel. *Graphentheorie*. 4. Auflage, Springer Heidelberg Dordrecht London New York, 2010.

[7] John D. Dixon and Brian Mortimer. *Permutation Groups*. Graduate Texts in Mathematics, vol. 163. Springer-Verlag New York, 1996.

[8] Daniel Gorenstein. *Finite Groups*. Second Edition, AMS Chelsea Publishing Series. American Mathematical Society, 2007.

[9] Matthias Grüninger, Max Horn, and Bernhard Mühlherr. Moufang twin trees of prime order. *Advances in Mathematics*, 302: pp. 1–24, 2016.

[10] Christoph Hering, William M. Kantor, and Gary M. Seitz. Finite groups with a split BN-pair of rank 1. I. *Journal of Algebra*, 20(3): pp. 435–475, 1972.

[11] Kevin McCrimmon. Jordan algebras and their applications. *Bulletin of the American Mathematical Society*, 84(4): pp. 612–627, 1978.

[12] Bernhard Mühlherr. Locally split and locally finite twin buildings of 2-spherical type. *Journal für die reine und angewandte Mathematik*, 1999(511): pp. 119–143, 1999.

[13] Bernhard Mühlherr and Mark A. Ronan. Local to global structure in twin buildings. *Inventiones Mathematicae*, 122: pp. 71–81, 1995.

[14] Derek J.S. Robinson. *A Course in the Theory of Groups*. Graduate Texts in Mathematics, vol. 80. Second Edition, Springer Science+Business Media New York, 1996.

[15] Mark A. Ronan. *A Construction of Buildings with no Rank 3 Residues of Spherical Type*. In: Luigi Antonio Rosati (ed.). *Buildings and the Geometry of Diagrams*. Lecture Notes in Mathematics, vol. 1181, pp. 242–248. Springer-Verlag Berlin Heidelberg, 1986.

[16] Mark A. Ronan. *Lectures on Buildings: Updated and Revised*. The University of Chicago Press, Chicago and London, 2009.

[17] Mark A. Ronan and Jacques Tits. Building buildings. *Mathematische Annalen*, 278: pp. 291–306, 1987.

[18] Mark A. Ronan and Jacques Tits. Twin trees I. *Inventiones Mathematicea*, 116: pp. 463–479, 1994.

[19] Moritz Scholz. *Moufang-systems*. Unpublished manuscript, 13 pages, September 2014.

[20] Yoav Segev and Richard M. Weiss. On the action of the Hua subgroups in special Moufang sets. *Mathematical Proceedings of the Cambridge Philosophical Society*, 144(1): pp. 77–88, 2008.

[21] Jean-Pierre Serre. *Trees*. Translated from the French by John Stillwell. Springer Monographs in Mathematics. Corrected Second Printing, Springer-Verlag Berlin Heidelberg New York, 2003.

[22] Ernest Shult. On a class of doubly transitive groups. *Illinois Journal of Mathematics*, 16(3): pp. 434–445, 1972.

[23] Jacques Tits. *A Local Approach to Buildings*. In: Chandler Davis, Branko Grünbaum, and F. A. Sherk (Eds.). *The Geometric Vein - The Coxeter Festschrift*, pp. 519–547. Springer-Verlag New York Inc., 1981.

[24] Jacques Tits. *Immeubles de type affine*. In: Luigi Antonio Rosati (ed.). *Buildings and the Geometry of Diagrams*. Lecture Notes in Mathematics, vol. 1181, pp. 159–190. Springer-Verlag Berlin Heidelberg, 1986.

[25] Jacques Tits. Endliche Spiegelungsgruppen, die als Weylgruppen auftreten. *Inventiones Mathematicae*, 43: pp. 283–295, 1977.

[26] Jacques Tits. *Buildings of Sperical Type and Finite BN-Pairs*. Lecture Notes in Mathematics, vol. 386. 2nd Corrected Printing, Springer-Verlag Berlin Heidelberg, 1986.

[27] Jacques Tits. Uniqueness and presentation of Kac-Moody groups over fields. *Journal of Algebra*, 105(2): pp. 542–573, 1987.

[28] Jacques Tits. *Twin buildings and groups of Kac-Moody type*. In: Martin W. Liebeck and Jan Saxl (Eds.). *Groups, Combinatorics and Geometry*. London Mathematical Society Lecture Note Series, vol. 165, pp. 249–286. Cambridge University Press, Cambridge, 1992.

[29] Jacques Tits. *Arbres jumelés (cours 1995-1996)*. In: Résumés des cours au Collège de France 1973-2000. Documents Mathématique, vol. 12, pp. 275–298. Société Mathématique de France, 2013.

[30] Jacques Tits. *Immeubles jumelés (cours 1988-1989)*. In: Résumés des cours au Collège de France 1973-2000. Documents Mathématique, vol. 12, pp. 157–172. Société Mathématique de France, 2013.

[31] Jacques Tits and Richard M. Weiss. *Moufang Polygons*. Springer Monographs in Mathematics. Springer-Verlag Berlin Heidelberg, 2002.

Index

Statement of authorship

I declare that I have completed this dissertation single-handedly without the unauthorized help of a second party and only with the assistance acknowledged therein. I have appropriately acknowledged and cited all text passages that are derived verbatim from or are based on the content of published work of others, and all information relating to verbal communications. I consent to the use of an anti-plagiarism software to check my thesis. I have abided by the principles of good scientific conduct laid down in the charter of the Justus Liebig University Giessen „Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis“ in carrying out the investigations described in the dissertation.

Gießen, October 2021