Now, we can prove that $G_1$ is not linear without quoting a big theorem.

**Proposition 3.68.** A finitely generated linear torsion group $G$ is periodic.

**Proof.** Using Mal'cev's observation (3.39), we assume that $G \leq \text{GL}_n(k)$ where $k$ is a purely transcendental extension of its prime field $k_0$.

Consider $g \in G \leq \text{GL}_n(k)$ and let $\mu_g(t)$ be its minimal polynomial. We aim to show that only finitely many polynomials arise this way. This will imply the proposition since the order of an endomorphism only depends on its minimal polynomial.

Since $g$ is torsion, we have, say, $g^n = 1$. Since every root of $\mu_g(t)$ is an eigenvalue for $g$, we find, that roots of $\mu_g(t)$ are roots of unity. Hence these roots are algebraic integers over $k_0$. The coefficients of $\mu_g$ are elementary symmetric functions of the roots. Since $k$ is a purely transcendental extension of $k_0$, these coefficients belong to $k_0$. For finite $k_0$ we are done now.

For $k_0 = \mathbb{Q}$, everything takes place inside $\mathbb{C}$. Thus, we can talk about absolute values. Roots of unity have absolute value 1. This implies a bound on the absolute values of coefficients of $\mu_g(t)$. On the other hand, these coefficients are integers, as we have seen. Therefore, we have only finitely many numbers from which to chose our coefficients. \hfill q.e.d.

**Corollary 3.69.** $G_1$ is not linear.

### 3.3.2 Congruence Subgroups

Put $G_1(s) := \ker(G_1 \to \text{Aut}(T^*_2(s)))$.

**Observation 3.70.** Since every non-trivial tree automorphism has to act non-trivially on some finite subtree, we have

$$\bigcap_{s \geq 1} G_1(s) = 1.$$
**Definition 3.71.** A **congruence subgroup** of \( G_1 \) is a group that contains the groups \( G_1(s) \) for \( s \) large enough.

**Remark 3.72.** This definition is reminiscent of arithmetic groups. The congruence subgroups of \( \text{SL}_n(Z) \) are those normal subgroups that contain a kernel of a congruence-homomorphism \( \text{SL}_n(Z) \to \text{SL}_n(Z_m) \). It turns out that for \( n \geq 3 \), every non-central normal subgroup in \( \text{SL}_n(Z) \) is a congruence subgroup.

We aim to prove

**Theorem 3.73 (Congruence Subgroup Property).** Every non-trivial normal subgroup of \( G_1 \) is a congruence subgroup.

**Remark 3.74.** This is to say that any quotient of \( G_1 \) is a quotient of one of the finite groups \( \text{Aut}(T_2^*(s)) \). In particular, all proper quotients of \( G_1 \) are finite, i.e., it is just infinite.

**Definition 3.75.** A group is **just infinite** if it is infinite but all its proper quotients are finite.

**Lemma 3.76.** The canonical homomorphism

\[
G_1(2) \to (G_1 \times G_1) \times (G_1 \times G_1) = G_1^4
\]

is onto in each coordinate.

**Proof.** Check

\[
((\sigma, \gamma), (\gamma, \sigma)) = \gamma \sigma \delta \sigma \beta \\
((\gamma, \sigma), (\sigma, \gamma)) = \sigma \gamma \sigma \delta \sigma \beta \sigma \\
((\beta, \beta), (\beta, \beta)) = \sigma \gamma \sigma \gamma \sigma \gamma \\
\]

and recall that \( G_1 = \langle \sigma, \beta, \gamma \rangle \). \( \text{q.e.d.} \)

**Lemma 3.77.** For any level \( s \), the action of \( G_1 \) on the set of vertices of level \( s \) is transitive.
Proof. This is an easy induction. Since $\sigma \in G_1$, we are done for $s = 1$.

On level $s \geq 2$, we have a left half and a right half. The swap $\sigma$ interchanges these halves. Hence we only have to see that the action on the left half is transitive. This follows by induction from the fact that for any $\xi \in G_1$, there is a partner $\zeta \in G$ such that $(\xi, \zeta) \in G_1(1)$. So we can act transitively on the left half by going into the left subtree and choosing an appropriate $\xi$. \hfill q.e.d.

Proposition 3.78. For the commutator subgroup of $K$, we have

$$G_1(5) \leq [K,K] \leq G_1(3).$$

Proof. Note that $[K,K]$ is normal in $G_1$ since inner automorphism take commutators to commutators. As a normal subgroup, it is generated by the commutators of the three generators $t$, $v$, and $w$. So let us compute these commutators first:

$$[v,w] = [(t,1),(1,t)]$$
$$= 1$$
$$[v,t] = [(t,1),(\gamma \sigma, \sigma \gamma)]$$
$$= (w,1)$$
$$= ((1,t),(1,1))$$

$$\xi := [w,t^{-1}] = [(1,t),(\sigma \gamma, \gamma \sigma)]$$
$$= (1,w)$$
$$= ((1,1),(1,t))$$

From this computation we infer:

- $[K,K] \leq G_1(3)$ since $t \in G_1(1)$.
- $(1 \times 1) \times (1 \times K) \leq [K,K]$. This follows from (3.76) and the fact that $K$ is the normal closure of $t$. 

3.26
On the other hand, we have
\[
\sigma \xi \sigma = ((1, t), (1, 1)) \\
\beta \sigma \xi \sigma \beta = ((t, 1), (1, 1)) \\
\sigma \beta \sigma \xi \sigma \beta \sigma = ((1, 1), (t, 1))
\]
whence we have
\[
(1 \times K) \times (1 \times 1) \leq [K, K] \\
(K \times 1) \times (1 \times 1) \leq [K, K] \\
(1 \times 1) \times (K \times 1) \leq [K, K]
\]
It follows that \( K \times K \times K \times K \leq [K, K] \). Thus
\[
G_1(5) \leq G_1(4) \times G_1(4) \leq G_1(3)^4 \leq K^4 \leq [K, K]
\]
which proves the claim. \( \text{q.e.d.} \)

**Exercise 3.79.** Prove that
\[
K \times K \times K \times K = [K, K]
\]
and determine the index of \([K, K]\) in \(K\).

**Lemma 3.80.** Fix \( \xi = (\xi_1, \ldots, \xi_2^2) \in G_1(t) - G_1(t + 1) \) such that \( \xi_1 \notin G_1(1) \). Then, for any two elements \( \kappa_1, \kappa_2 \in K \), we have
\[
[[\xi, \kappa_1'], \kappa_2'] = \left(\left([\kappa_1^{-1}, \kappa_2], 1,1,\ldots,1\right)_2^2\right)
\]
where
\[
\kappa_i' := (\kappa_i, 1, 1, \ldots, (1, 1))_{2^2} \in K.
\]

**Proof.** We observe \( \xi_1 = (\zeta_0, \zeta_1)\sigma \), and compute
\[
[[\xi, \kappa_1']] = \left(\left(\kappa_1^{-1}, \zeta_1\kappa_1^{-1}\zeta_1^{-1}, 1, \ldots, 1\right)_2^2\right)
\]
and the claim follows. \( \text{q.e.d.} \)
Corollary 3.81. $G_1$ has trivial center.

Proof. As $G_1$ acts transitively on the vertices of any fixed level (3.77), any non-trivial element in $G_1$ is conjugate to an element satisfying the hypotheses of (3.80). For a central element, however, the double commutator had to be trivial which contradicts $G_1(5) \leq [K,K]$ (3.78). \hspace{1cm} \text{q.e.d.}

Now we are ready to prove that $G_1$ has the congruence subgroup property.

Proof of Theorem (3.73). Let $N$ be a normal subgroup and $\xi \in N$ be a non-trivial element.

As $\cap G_1(s) = 1$, we find $t$ such that $\xi \in G_1(t) - G_1(t+1)$. Since the action of $G_1$ on the vertices of any fixed level is transitive (3.77), we can conjugate $\xi$ such that it satisfies the hypotheses of (3.80). For arbitrary elements $\kappa_1$ and $\kappa_2$, we have

$$[[\xi, \kappa_1^t], \kappa_2^t] = ([\kappa_1^{-1}, \kappa_2], 1, 1, \ldots, 1)_{G_1}.$$ 

This element belongs to the normal closure of $\xi$. Since $G_1(5) \leq [K,K]$, it follows that

$$G_1(5) \times 1 \times \cdots \times 1 \leq N$$

where we have $2^{t+1}$ factors. Using (3.77) again, it follows that

$$G_1(5) \times \cdots \times G_1(5) \leq N.$$ 

This implies $G_1(t + 1 + 5) \leq N$. \hspace{1cm} \text{q.e.d.}

3.4 Intermediate Growth

Definition 3.82. We call a word in the letters $\{\sigma, \beta, \gamma, \delta\}$ reduced if the letter $\sigma$ alternates with the other letters. For each word $w$, there is a unique reduced word $\overline{w}$ obtained by applying the following to processes:
1. Delete an occurrence of $aa$ for any letter $a \in \{\sigma, \beta, \gamma, \delta\}$. This process shortens the word by two letters.

2. Replace a pair of distinct letters in $\{\beta, \gamma, \delta\}$ by the third one.
   This process shortens the word by one letter.

For any word $w$ that represents an element in $G_1(s)$, we can alternatingly reduce and split as follows to obtain the following cascade

\[
\overline{w} = (w_0, w_1) \\
\overline{w_0} = (w_0, w_{0,1}) \\
\overline{w_1} = (w_1, w_{1,1}) \\
\vdots \\
\overline{w_{1,1}} = (w_{1,0}, w_{1,1}) \\
\vdots \\
\overline{w_{1,...,1}} = (w_{1,...,0}, w_{1,...,1}).
\]

**Lemma 3.83 ([Gr85]).** For every $w \in G_1(3)$,

\[
\sum_{i,j,k \in \{0,1\}} |\overline{w}_{i,j,k}| \leq \frac{3}{4} |w| + 8
\]

**Proof.** First, let us see why the statement should be true. In each step, we rewrite a word as a pair. We know what happens to the non-$\sigma$-letters. They cycle through $\beta \mapsto \gamma \mapsto \delta \mapsto 1$. Hence, after three steps, all non-$\sigma$-letters will disappear. Note that these letters account for half of the length of $\overline{w}$. So how does the $\frac{3}{4}$ arise? On our way down to level 3, we do not only rewrite words as pairs, we also reduce words. Because of this, we might lose track of two letters that merge into one new letter. In this way, a letter either disappears or merges with another letter contributing at least a shortening of $\overline{w}$ by half a letter. Since half of the letters do this, the word is shortened by $\frac{1}{4}$.

The real proof is all about bookkeeping. First observe

\[
|w_0| + |w_1| \leq |\overline{w}| + 1 - C_\delta(\overline{w})
\]

3.29
the “+1” being there for $\overline{w}$ is reduced but not necessarily cyclicly reduced. Now we reduce the $w_i$ and have

$$|\overline{w_0}| + |\overline{w_1}| \leq |\overline{w}| + 1 - C_\delta(\overline{w}) - r_1$$  \hspace{1cm} (3)

where

$$r_1 := (|w_0| - |\overline{w_0}|) + (|w_1| - |\overline{w_1}|).$$

This transformation may seem silly, but the key observation in this proof is that $r_1$, and its relatives defined farther down the road, have a double meaning: one the one hand side, these reduction counts measure our progress in shortening the words, on the other hand, they measure the loss of control that we have over the letters in the words. The reason is that the reduction processes are cancellations which cause letters to disappear or change. This is good (for we make progress) and bad (for we lose control) at the same time. The key issue is to keep track of it so that we can see how it works out in the end. To illustrate how $r_1$ helps to keep track of things, let us observe that

$$C_\gamma(\overline{w_0}) + C_\delta(\overline{w_0}) + C_\gamma(\overline{w_1}) + C_\delta(\overline{w_1}) \geq C_\beta(\overline{w}) + C_\gamma(\overline{w}) - 2r_1.$$  \hspace{1cm} (4)

The reason for this inequality is, of course, that the letters $\gamma$ and $\delta$ in the components $w_i$ derive from letters $\beta$ and $\gamma$ in $\overline{w}$ and we loose them only due to cancellations. Passing from $w_i$ to $\overline{w_i}$, we loose track of two letters in every cancellation but, in the worst case, we might shorten the words only by one letter — this explains the factor of 2.

Let us descend one level. So we are dealing with $w_{i,j}$ defined by $\overline{w_i} = (w_{i,0}, w_{i,1})$ and the reductions $\overline{w_{i,j}}$. Doing the same as in establishing (3), we find

$$\sum_{i,j \in \{0,1\}} |\overline{w_{i,j}}| \leq (|\overline{w_0}| + 1 - C_\delta(\overline{w_0})) + (|\overline{w_1}| + 1 - C_\delta(\overline{w_1})) - r_2$$  \hspace{1cm} (5)

where

$$r_2 := \sum_{i,j \in \{0,1\}} (|w_{i,j}| - |\overline{w_{i,j}}|).$$
Let us combine (3) and (5). We obtain
\[
\sum_{i,j \in \{0,1\}} |w_{i,j}| \leq \|w\| + 3 - C_\delta (\|w\|) - r_1 - C_\delta (\|w_0\|) - C_\delta (\|w_1\|) - r_2.
\] (6)

Descending to level 3 finally yields:
\[
\sum_{i,j,k \in \{0,1\}} |w_{i,j}| \leq \|w\| + 7 - C_\delta (\|w\|) - r_1 - C_\delta (\|w_0\|) - C_\delta (\|w_1\|) - r_2 - \sum_{i,j \in \{0,1\}} C_\delta (\|w_{i,j}\|)
\] (7)

where we do not account for the reductions in the third step. It turns out that we get away with a more sloppy estimate this time.

Now we derive the level 2 companion to (4). We find
\[
\sum_{i,j \in \{0,1\}} C_\delta (\|w_{i,j}\|) \geq C_\gamma (\|w_0\|) + C_\gamma (\|w_1\|) - 2r_2
\]
\[
= C_\gamma (\|w_0\|) + C_\delta (\|w_0\|) + C_\gamma (\|w_1\|) + C_\delta (\|w_1\|) - C_\delta (\|w_0\|) - C_\delta (\|w_1\|) - 2r_2
\]
\[
\geq C_\beta (\|w\|) + C_\gamma (\|w\|) - 2r_1 - C_\delta (\|w_0\|) - C_\delta (\|w_1\|) - 2r_2.
\]

Here we used (4) in the last estimate. We can use this inequality to embark on (7). Substitution in the last term yields:
\[
\sum_{i,j,k \in \{0,1\}} |w_{i,j}| \leq \|w\| + 7 - C_\delta (\|w\|) - C_\beta (\|w\|) - C_\gamma (\|w\|) + r_1 + r_2.
\]

Now we use the fact that the non-\(\sigma\)-letters in \(\bar{w}\) account for half of its length — maybe we are off by one if the word starts and ends with \(\sigma\). Thus, we infer
\[
\sum_{i,j,k \in \{0,1\}} |w_{i,j}| \leq \frac{1}{2} \|w\| + 8 + r_1 + r_2.
\]

On the other hand, we obtain from (7) just by dropping terms:
\[
\sum_{i,j,k \in \{0,1\}} |w_{i,j}| \leq \|w\| + 7 - r_1 - r_2
\]

The proof is finished by averaging the last two estimates: \(-r_1\) and \(r_i\) average out to 0, whereas \(\|w\|\) and \(\frac{1}{2} \|w\|\) have average \(\frac{3}{2} \|w\|\). Finally, the average of 7 and 8 is 8.

q.e.d.

3.31
Lemma 3.84.

\[ P_m(n) := \left| \{ (k_1, \ldots, k_m) \in \mathbb{N}^m \mid \sum k_i < n \} \right| \leq n^m \]

Proof. Observe

\[ P_m(n) = \sum_{k=1}^n P_{m-1}(k) \leq \sum_{k=1}^n k^{m-1} \leq n^m \]

and induct. \[ \text{q.e.d.} \]

Theorem 3.85 (Grigorchuk). \( G_1 \) has intermediate growth.

Proof. We already observed that \( G_1 \) does not have polynomial growth (3.63). Thus, we have to prove that its growth is subexponential.

We consider the generating set \( \Sigma := \{ \sigma, \beta, \gamma, \delta \} \) for \( G_1 \).

Assume the \( \beta_{\Sigma} \) is exponential, i.e., we have

\[ O(L^n) \leq \beta_{\Sigma}(n) \]

for some \( L > 1 \). Since \( G_1(3) \) is of finite index, there is a constant \( K \) such that

\[ O(L^n) \leq \beta_{\Sigma}(n) \]
\[ \leq |\{ \xi \in G_1(3) \mid |\xi| \leq n + K \}| \]
\[ \leq \sum_{i_1 + \cdots + i_6 \leq \frac{3}{4}(n+K)+9} \beta_{\Sigma}(i_1) \cdots \beta_{\Sigma}(i_6) \]
\[ \leq P_8 \left( \frac{3}{4}(n+K)+9 \right) O\left( L^{\frac{3}{4}(n+K)+9} \right) \]
\[ = O\left( L^{\frac{3}{4}n} \right) \]

which contradicts \( L > 1 \). \[ \text{q.e.d.} \]

3.5 Presentations

Here, we will outline a strategy for obtaining a presentation of \( G_1 \).

The starting point is the group

\[ \Gamma := \langle \sigma, \beta, \gamma, \delta \mid \sigma^2 = \beta^2 = \gamma^2 = \delta^2 = \beta \gamma \delta = 1 \rangle \cong C_2 * \mathbb{V} \]
which has an obvious homomorphism onto $G_1$. Let $\Gamma_1$ and $\Gamma_K$ denote the preimages of $G_1(1)$ and $K$, respectively, and meditate on the following diagram:

$$
\begin{array}{ccl}
N & \hookrightarrow & \Gamma \\
\uparrow & & \uparrow \\
\Gamma_1 \times \Gamma_1 & \xleftarrow{\psi=(\psi_L, \psi_R)} & \Gamma_1 \\
\uparrow & & \uparrow \\
\Gamma_K \times \Gamma_K & \to & K \\
\end{array}
$$

Our goal is do determine the kernel $N$ of the epimorphism $\pi$. Let $N_0 := 1 \leq \Gamma_1 \leq \Gamma$, and put

$$
N_{i+1} := \psi^{-1}(N_i \times N_i).
$$

Note that $N_1$ is the kernel of $\psi$. Put

$$
N_\infty := \bigcup_i N_i.
$$

**Lemma 3.86.** $N = N_\infty$.

**Proof.** The inclusion $N_\infty \leq N$ is clear. To see the other inclusion, let $w$ be a reduced word in $N$. Note that for $|w| \geq 3$, we have

$$
|\psi_L(w)| \leq \frac{|w|}{2} \quad \text{and} \quad |\psi_R(w)| \leq \frac{|w|}{2}.
$$

Since a non-trivial reduced word of length $\leq 2$ in $\Gamma$ is not in $N$, we infer that

$$
\psi_{[\log_2(|w|)]}(w) = (1, \ldots, 1_{2^{[\log_2(|w|)]}}).
$$

This implies $w \in N_{[\log_2(|w|)]} \leq N_\infty$. q.e.d.

**Observation 3.87.** Since $\psi_L(w) = \psi_R(\sigma w \sigma)$ and $\beta, \gamma, \delta \in \Gamma_1$, all the $N_i$ are normal in $\Gamma$ and we have $N_{i+1} = \psi_R^{-1}(N_i)$. q.e.d.

Our next goal is to find a subset $R$ that generates $N$ as a normal subgroup of $\Gamma$, so that we can turn the identity $G_1 = \Gamma/N$ into

3.33
a presentation for $G_i$. Our starting point will be a finite set $R_1 \subset \Gamma$ that generates $N_1$ as a normal subgroup in $\Gamma$: To find $R_1$, note that the image of $\psi$ is a finitely presented group since it has finite index in the finitely presented group $\Gamma \times \Gamma$. Hence we can find a finite presentation by the Reidemeister-Schreier method (A.13). Thus $N_1$ is finitely generated as a normal subgroup of $\Gamma_i$ and we can find explicitly a finite set $R_0$ that generates $N_1$ as a normal subgroup of $\Gamma_1$. Improving upon this, you can construct a finite, and possibly smaller, set that generates $N_1$ as a normal subgroup of $\Gamma$. We will not carry this out, but a good choice for $R_1$ is known.

**Fact 3.88 ([Gri98]).** $R_1 = \{(\sigma \delta)^4, (\sigma \gamma)^8, (\sigma \delta \sigma \gamma)^4\}$ generates $N_1$ as a normal subgroup of $\Gamma$.

The next step is to determine, in terms of $R_1$, our normal generating set $R$ for $N$. The problem is: $N_{i+1}$ is defined as a preimage, thus we have everything backwards. Taking iterated preimages does not lend itself to a nice description of a set $R$. The following fact will allow us to overcome this obstacle.

**Miracle 3.89.** The letter substitution

$$
\begin{align*}
\sigma & \mapsto \sigma \gamma \sigma \\
\beta & \mapsto \delta \\
\gamma & \mapsto \beta \\
\delta & \mapsto \gamma
\end{align*}
$$

defines an endomorphism $\Phi : \Gamma \to \Gamma$ which satisfies

$$
\Phi(w) = (1, w) \quad \text{for all } w \in \Gamma_K.
$$

To see this, just evaluate $\Phi$ on $t$, $v$, and $w$ and check that you obtain $(1, t)$, $(1, v)$, and $(1, w)$.

It follows that $N_{i+1}$ is the normal closure of $\Phi(N_i)$. Hence, $N$ is the normal closure of $R := \bigcup_i \Phi(R_i)$. Thus, the observation that $\Phi((\sigma \delta)^4) = (\sigma \gamma)^8$ completes our sketchy proof of

3.34
Theorem 3.90 ([Lys85],[Gri98]). The First Grigorchuk Group $G_1$ has a presentation

$$G_1 = \langle \sigma, \beta, \gamma, \delta \mid \sigma^2 = \beta^2 = \gamma^2 = \delta^2 = \beta \gamma \delta = \Phi^i((\sigma \delta)^4) = \Phi^i((\sigma \delta \gamma \sigma \gamma)^4) = 1, (i \geq 0) \rangle.$$ 

**Definition 3.91.** A group is co-Hopfian if every injective endomorphism is onto.

**Exercise 3.92.** The endomorphism $\Phi: \Gamma \to \Gamma$ descends to an injective endomorphism of $G_1$. The index of the image is infinite. Thus $G_1$ is not co-Hopfian.

**Remark 3.93.** As $G_1$ is finitely generated and residually finite, it is automatically Hopfian (2.22).

**Exercise 3.94.** Show that, for all $i \geq 1$,

$$\Phi^i((\sigma \delta)^4) \in N_{i+1} - N_i.$$ 

**Corollary 3.95.** For all $i$, the inclusion $N_i \leq N_{i+1}$ is strict. In particular, $G_1$ is not finitely presented.

**Proof.** If $G_1$ was finitely presented, all but finitely many relators in the presentation (3.90) would be redundant. Hence the inclusions $N_i \leq N_{i+1}$ would stabilize. \(\text{q.e.d.}\)

### 3.6 The Conjugacy Problem

**Theorem 3.96.** The conjugacy problem for $G_1$ is solvable.