# A GRAPH-THEORETIC PROOF FOR WHITEHEAD'S SECOND FREE-GROUP ALGORITHM

WARREN DICKS*

ABSTRACT. J. H. C. Whitehead's second free-group algorithm determines whether or not two given elements of a free group lie in the same orbit of the automorphism group of the free group. The algorithm involves certain connected graphs, and Whitehead used three-manifold models to prove their connectedness; later, Rapaport and Higgins & Lyndon gave group-theoretic proofs.

Combined work of Gersten, Stallings, and Hoare showed that the three-manifold models may be viewed as graphs. We give the direct translation of Whitehead's topological argument into the language of graph theory.

## 1. Minimal background

Whitehead(1936b) gave an algorithm which, with input two finite sequences $S_1$, $S_2$ of elements (or conjugacy classes of elements) of a finite-rank free group $F$, outputs either an $F$-automorphism $\varphi$ such that $\varphi(S_1) = S_2$ or an assurance that no such $\varphi$ exists. More importantly, he introduced certain connected graphs that have been of great interest to group theorists. His nine-page proof of connectedness used a three-manifold model for each $F$-automorphism. Rapaport(1958) gave a twenty-page group-theoretic proof of connectedness, and Higgins & Lyndon(1962, 1974) gave one of five pages; these proofs led the way to an even deeper understanding of $F$-automorphisms.

Gersten(1987) constructed a graph model for each $F$-automorphism, and Stallings(1983) pointed out a connection between Gersten's model and Whitehead's. Krstić(1989) used Cayley trees to simplify Gersten's construction. Hoare(1990) gave an explicit description of Whitehead's model in terms of Gersten's. Below, we give the resulting translation of Whitehead's topological argument into the language of graph theory.[1] This argument concerns changes of bases (free-generating sets) rather than automorphisms, and ours may be the first treatment of Gersten's graphs that does not mention group morphisms.

All of the following will apply throughout.

1.1. **Notation.** Set $\mathbb{N} := \{0, 1, 2, \ldots\}$.

Let $F$ be a finite-rank free group. By a *straight word in* $F$, we mean an element of $F$; by a *cyclic word in* $F$, we mean the $F$-conjugacy class of an element of $F$; and, by a *word in* $F$, we mean a straight-or-cyclic word in $F$. Let $R$ be a finite set of words in $F$. Let

2010 *Mathematics Subject Classification.* Primary: 20E05; Secondary:20F10, 20E08.

*Key words and phrases.* Free group, Whitehead algorithm, Gersten graph.

*Research supported by MINECO (Spain) through project number MTM2014-53644-P.

[1] For his earlier free-group algorithm, Whitehead(1936a) also used three-manifold models, to prove his celebrated cutvertex lemma. Hoare(1988) gave the second proof, using Gersten's graphs in place of manifolds. Dicks(2014, 2017), refining work of Stong(1997), proved a more general result by tricolouring a Cayley tree. The elegant folding theorem of Heusener & Weidmann(2014) leads to a yet more general result.

$X$ and $Y$ be $F$-bases. In Section 2, we shall recall the value $h(X) := \sum_{r\in R} X\text{-length}(r) \in \mathbb{N}$. We write $X^{\pm 1} := X \cup X^{-1}$. We say that $Y$ is a *Whitehead transform of* $X$ if there exists some $x \in X^{\pm 1}$ such that $Y \subseteq \{1, x\}\cdot X\cdot\{1, x^{-1}\}$. We say that $X$ *is a local-minimum point for* $h$ if $h(X) \leqslant h(X')$ for each Whitehead transform $X'$ of $X$. □

In Section 3, we shall use Gersten's graphs to define a value $\mathrm{d}(X, Y) \in \mathbb{N}$ that Whitehead used tacitly. What the topological portion of Whitehead's argument shows is precisely

(1.1) if $X$ and $Y$ are local-minimum points for $h$, then either $X^{\pm 1} = Y^{\pm 1}$ or some Whitehead transform $Y'$ of $Y$ satisfies $h(Y') = h(Y)$ and $\mathrm{d}(X, Y') < \mathrm{d}(X, Y)$.

This will be stated as Theorem 3.3 below, and our sole objective is to give a self-contained graph-theoretic proof that copies Whitehead's. All the other parts of his article are graph theoretic or group theoretic, and we shall not discuss them. However, Whitehead leaves the main consequence of (1.1) unsaid, and it is as follows.

Let us say that $Y$ is an *$F$-neighbour* of $X$ if either $Y^{\pm 1} = X^{\pm 1}$ (whence $h(Y) = h(X)$) or $Y$ is a Whitehead transform of $X$. Let $\Gamma(F)$ denote the graph with vertices the $F$-bases and with edges joining $F$-neighbours. Let $\Gamma(h)$ denote the subgraph of $\Gamma(F)$ with vertices the local-minimum points for $h$ and with edges joining $F$-neighbours. It is obvious, but important, that $h$ is constant on each connected subgraph of $\Gamma(h)$, and that a simple algorithm outputs a strictly $h$-decreasing $\Gamma(F)$-path starting at any given $\Gamma(F)$-vertex and stopping when $\Gamma(h)$ is reached. Now suppose that $X$ is a local-minimum point for $h$ and that $h(Y) \leqslant h(Z)$ for each $F$-basis $Z$. By induction on $\mathrm{d}(X, Y)$, it follows from (1.1) that there exists some ($h$-constant) $\Gamma(h)$-path from $Y$ to $X$. On varying $X$, we find that $\Gamma(h)$ is connected, which may be considered to be the main result of Whitehead(1936b); it greatly generalizes the result of Nielsen(1919) that $\Gamma(F)$ itself is connected.

The connectedness of $\Gamma(F)$ was used in the arguments of Whitehead, Rapaport, Higgins & Lyndon, and Gersten. However, Krstić did not use it, and this will permit us to prove (1.1) without using it.

## 2. Review of Cayley trees

2.1. **Definitions.** By a *graph*, we mean a quintuple $(\Gamma, \mathrm{V}\Gamma, \mathrm{E}\Gamma, \iota, \tau)$ such that $\Gamma$ is a set, $\mathrm{V}\Gamma$ and $\mathrm{E}\Gamma$ are disjoint subsets of $\Gamma$ whose union is $\Gamma$, and $\iota$ and $\tau$ are maps from $\mathrm{E}\Gamma$ to $\mathrm{V}\Gamma$. We use the same symbol $\Gamma$ to denote both the graph and the set. We call $\mathrm{V}\Gamma$ and $\mathrm{E}\Gamma$ the *vertex-set* and *edge-set* of $\Gamma$ respectively, and call their elements $\Gamma$-*vertices* and $\Gamma$-*edges* respectively. The maps $\iota$ and $\tau$ are called the *initial* and *terminal* incidence functions respectively.

Each $e \in \mathrm{E}\Gamma$ has an inverse in the free group $\langle\, \mathrm{E}\Gamma \mid \emptyset \rangle$, and we set $\iota(e^{-1}) := \tau(e)$ and $\tau(e^{-1}) := \iota(e)$. For each $v \in \mathrm{V}\Gamma$, by the $\Gamma$-*valence* of $v$, we mean $\big|\{e \in (\mathrm{E}\Gamma)^{\pm 1} : \iota e = v\}\big|$.

By a $\Gamma$-*path*, we mean a sequence of the form $p = (v_0, e_1, v_1, e_2, v_2, \ldots, v_{\ell-1}, e_\ell, v_\ell)$, where $\ell \in \mathbb{N}$ and, for each $i \in \{1, 2, \ldots, \ell\}$, $e_i \in (\mathrm{E}\Gamma)^{\pm 1}$, $v_{i-1} = \iota e_i$, and $v_i = \tau e_i$. We sometimes abbreviate $p$ to $(e_1, e_2, \ldots, e_\ell)$, even if $\ell = 0$ when $v_0$ is specified. The path $p$ is said to be *from* $v_0$ *to* $v_\ell$, and to have *length* $\ell$. For each $e \in \mathrm{E}\Gamma$, by the *number of times* $p$ *traverses* $e$, we mean $\big|\{i \in \{1, 2, \ldots, \ell\} : e_i \in \{e\}^{\pm 1}\}\big|$. We call the element $e_1 e_2 \cdots e_\ell$ of $\langle\, \mathrm{E}\Gamma \mid \emptyset \rangle$ the $\Gamma$-*label of* $p$. If $v_\ell = v_0$, then we say that $p$ is a *closed* path *based at* $v_0$. If $e_i \neq e_{i-1}^{-1}$ for each $i \in \{2, 3, \ldots, \ell\}$, then we say that $p$ is a *reduced* path.

For $v, w \in \mathrm{V}\Gamma$, let $\Gamma[v,w]$ denote the set of all $\Gamma$-paths from $v$ to $w$; we then have the *inversion* map $\Gamma[v,w] \to \Gamma[w,v]$, $p \mapsto p^{-1}$, where $(e_1, e_2, \ldots, e_\ell)^{-1} := (e_\ell^{-1}, \ldots, e_2^{-1}, e_1^{-1})$. For $u, v, w \in \mathrm{V}\Gamma$, we have the *concatenation* map $\Gamma[u,v] \times \Gamma[v,w] \to \Gamma[u,w]$, $(p_1, p_2) \mapsto p_1 \# p_2$, where $(e_1, e_2, \ldots, e_\ell)\#(e'_1, e'_2, \ldots, e'_m) := (e_1, e_2, \ldots, e_\ell, e'_1, e'_2, \ldots, e'_m)$. If a $\Gamma$-path $p$ is closed and $p\#p$ is reduced, we say that $p$ is *cyclically reduced*.

We say that $\Gamma$ is a *tree* if $\mathrm{V}\Gamma \neq \emptyset$ and, for all $v, w \in \mathrm{V}\Gamma$, there exists a unique reduced $\Gamma$-path from $v$ to $w$. We say that $\Gamma$ is *connected* if, for all $v, w \in \mathrm{V}\Gamma$, there exists a $\Gamma$-path from $v$ to $w$. By a *component* of $\Gamma$, we mean a maximal nonempty connected subgraph of $\Gamma$. Thus, $\Gamma$ equals the disjoint union of its components. We say that $\Gamma$ is a *forest* if each component of $\Gamma$ is a tree. Thus, $\Gamma$ is not a forest if and only if some closed $\Gamma$-path traverses some $\Gamma$-edge exactly once.

For any group $G$, we say that $\Gamma$ is a *left $G$-graph* if $\mathrm{V}\Gamma$ and $\mathrm{E}\Gamma$ are left $G$-sets, and $\iota$ and $\tau$ are left-$G$-set morphisms; *right $G$-graphs* are defined similarly. □

Recall that $F$ is a finite-rank free group, and that $X$ and $Y$ are $F$-bases. The finite-rank hypothesis will not be used in this section.

2.2. **Definitions.** For any $g \in F$, we let $\cdot g$ and $g\cdot$ denote the permutations $F \to F$ given by $v \mapsto vg$ and $v \mapsto gv$ respectively. For any subset $S$ of $F$, we write $\cdot S := \{\cdot g : g \in S\}$ and $S\cdot := \{g\cdot : g \in S\}$.

We let $F\curvearrowleft Y$ denote the (Cayley) graph with vertex-set $F$ and edge-set $F \times \cdot Y$, for which each edge $(v, \cdot y)$ has initial vertex $v$ and terminal vertex $vy$. The $(F\curvearrowleft Y)$-paths $(v, (v, \cdot y), vy)$ and $(vy, (v, \cdot y)^{-1}, v)$ are depicted as $v \xrightarrow{\cdot y} vy$ and $vy \xrightarrow{\cdot y^{-1}} v$ respectively. An $(F\curvearrowleft Y)$-path $p$ will sometimes be depicted in the form

$$v \xrightarrow{\cdot y_1} vy_1 \xrightarrow{\cdot y_2} vy_1y_2 \to \cdots \to vy_1y_2\cdots y_{\ell-1} \xrightarrow{\cdot y_\ell} vy_1y_2\cdots y_{\ell-1}y_\ell,$$

for a unique $Y^{\pm1}$*-sequence* $\sigma = (y_1, y_2, \ldots, y_\ell)$, that is, an $\ell$-tuple of elements of $Y^{\pm1}$ for some $\ell \in \mathbb{N}$. We call $\sigma$ the *right $Y^{\pm1}$-label* of $p$. We say that $\sigma$ is *reduced* if $y_i \neq y_{i-1}^{-1}$ for each $i \in \{2, 3, \ldots, \ell\}$, and that $\sigma$ is *cyclically reduced* if $(y_1, y_2, \ldots, y_\ell, y_1, y_2, \ldots, y_\ell)$ is reduced. Thus, $p$ is a reduced $(F\curvearrowleft Y)$-path if and only its right $Y^{\pm1}$-label is a reduced $Y^{\pm1}$-sequence.

We let $X\curvearrowright F$ denote the graph with vertex-set $F$ and edge-set $X\cdot \times F$, for which each edge $(x\cdot, v)$ has initial vertex $v$ and terminal vertex $xv$. The $(X\curvearrowright F)$-paths $(v, (x\cdot, v), xv)$ and $(xv, (x\cdot, v)^{-1}, v)$ are depicted as $v \xrightarrow{x\cdot} xv$ and $xv \xrightarrow{x^{-1}\cdot} v$ respectively. An $(X\curvearrowright F)$-path $p$ will sometimes be depicted in the form

$$v \xrightarrow{x_1\cdot} x_1v \xrightarrow{x_2\cdot} x_2x_1v \to \cdots \to x_{\ell-1}\cdots x_2x_1v \xrightarrow{x_\ell\cdot} x_\ell x_{\ell-1}\cdots x_2x_1v,$$

for a unique $X^{\pm1}$-sequence $\sigma = (x_\ell, \ldots, x_2, x_1)$, called the *left $X^{\pm1}$-label* of $p$. Again, $p$ is a reduced $(X\curvearrowright F)$-path if and only its left $X^{\pm1}$-label is a reduced $X^{\pm1}$-sequence.

We let $X\curvearrowright F\curvearrowleft Y$ denote the graph with vertex-set $F$ and edge-set the (disjoint) union of $X\cdot \times F$ and $F \times \cdot Y$, with initial and terminal vertices as before. Thus, $X\curvearrowright F$ and $F\curvearrowleft Y$ are subgraphs of $X\curvearrowright F\curvearrowleft Y$ which are being amalgamated over their common vertex-set $F$. □

Dehn(1910) initiated the study of Cayley graphs of infinite groups, particularly surface groups, and he must have known the following at the start.

2.3. **Theorem.** *The left $F$-graph $F\curvearrowleft Y$ is a tree.*

*Proof* (Fox(1953), streamlined by Dicks(1980)). Set $T := F\curvearrowleft Y$. For each $(v, y) \in F \times Y$, set $v\otimes y := (v, \cdot y) \in F \times \cdot Y = \mathrm{E}\,T$; thus, $\iota(v\otimes y) = v$ and $\tau(v\otimes y) = vy$.

Clearly, $T$ is nonempty.

Let $\sim$ denote the inclusion-smallest equivalence relation on $\mathrm{V}T$ such that $\iota(v\otimes y) \sim \tau(v\otimes y)$ for each $T$-edge $v\otimes y$. There exists a left-$F$-set isomorphism between the set of components of $T$ and the set of equivalence classes of $\sim$. Also, $\sim$ is the inclusion-smallest equivalence relation on $F$ such that $v \sim vy$ for each $(v, y) \in F \times Y$. In particular, the equivalence class $[1]$ of 1 satisfies $[1] = [y] = y{\cdot}[1]$ for each $y \in Y$. Hence, the subgroup $\{f \in F : f{\cdot}[1] = [1]\}$ of $F$ includes $Y$. Thus, for all $f \in \langle Y \rangle = F$, $[1] = f{\cdot}[1] = [f]$. Hence, $[1] = F$. Thus, $T$ is connected.

For each set $S$, we let $\mathbb{Z}[S]$ denote the free $\mathbb{Z}$-module on $S$. The maps $\iota, \tau : \mathrm{E}T \to \mathrm{V}T$ induce $\mathbb{Z}$-module morphisms $\hat{\iota}, \hat{\tau} : \mathbb{Z}[\mathrm{E}T] \to \mathbb{Z}[\mathrm{V}T]$. For each closed $T$-path $p$ which traverses some $T$-edge exactly once, the abelianization map $\langle \mathrm{E}T \mid \emptyset \rangle \to \mathbb{Z}[\mathrm{E}T]$ carries the $T$-label of $p$ to a nonzero element of the kernel of $\hat{\tau}{-}\hat{\iota}$. Thus, to show that $T$ is a tree, it suffices to show that $\hat{\tau}{-}\hat{\iota}$ is injective. Using the natural left $F$-action on $\mathbb{Z}[\mathrm{E}T]$, we may form the semi-direct-product group $\left(\begin{smallmatrix} F & \mathbb{Z}[\mathrm{E}T] \\ \{0\} & \{1\} \end{smallmatrix}\right)$ with matrix-style multiplication, wherein each element $\left(\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix}\right)$ is denoted $\lceil a, b \rceil$. Since $Y$ is an $F$-basis, there exists a unique group morphism $F \to \left(\begin{smallmatrix} F & \mathbb{Z}[\mathrm{E}T] \\ \{0\} & \{1\} \end{smallmatrix}\right)$, $f \mapsto \lceil \varphi f, \alpha f \rceil$, such that $\lceil \varphi y, \alpha y \rceil = \lceil y, 1\otimes y \rceil$ for each $y \in Y$. For all $f$, $g \in F$,

$$\lceil \varphi(fg), \alpha(fg) \rceil = \lceil \varphi f, \alpha f \rceil \lceil \varphi g, \alpha g \rceil = \lceil (\varphi f)(\varphi g), (\varphi f)(\alpha g) + \alpha f \rceil.$$

Then $\varphi : F \to F$ is the identity map, since $\varphi y = y$ and $\varphi(fg) = (\varphi f)(\varphi g)$. The map $\alpha : F \to \mathbb{Z}[\mathrm{E}T]$ satisfies $\alpha y = 1\otimes y$ and $\alpha(fg) = (\varphi f)(\alpha g) + \alpha f$. Thus, we have a map $\alpha : \mathrm{V}T \to \mathbb{Z}[\mathrm{E}T]$ such that, for each $v\otimes y \in \mathrm{E}T$,

$$\alpha\big(\tau(v\otimes y)\big) - \alpha\big(\iota(v\otimes y)\big) = \alpha(vy) - \alpha(v) = (\varphi v)(\alpha y) = (v)(1\otimes y) = v\otimes y.$$

Now $\alpha$ induces a $\mathbb{Z}$-module morphism $\hat{\alpha} : \mathbb{Z}[\mathrm{V}T] \to \mathbb{Z}[\mathrm{E}T]$, and the composite

$$\mathbb{Z}[\mathrm{E}T] \xrightarrow{\hat{\tau}-\hat{\iota}} \mathbb{Z}[\mathrm{V}T] \xrightarrow{\hat{\alpha}} \mathbb{Z}[\mathrm{E}T]$$

is the identity map on $\mathbb{Z}[\mathrm{E}T]$, since it carries each $v\otimes y \in \mathrm{E}T$ to itself. Hence, $\hat{\tau}{-}\hat{\iota}$ is injective, as desired. $\square$

2.4. **Definitions.** For each straight word $r$ in $F$, there exists some reduced $Y^{\pm 1}$-sequence $(y_1, y_2, \ldots, y_\ell)$ such that $y_1 y_2 \cdots y_\ell = r$. Here,

$$1 \xrightarrow{\cdot y_1} y_1 \xrightarrow{\cdot y_2} y_1 y_2 \rightarrow \cdots \rightarrow y_1 y_2 \cdots y_{\ell-1} \xrightarrow{\cdot y_\ell} y_1 y_2 \cdots y_\ell = r$$

is a reduced $(F{\curvearrowleft}Y)$-path from 1 to $r$, which is unique by Theorem 2.3. Thus, $(y_1, y_2, \ldots, y_\ell)$ is unique, and we call it the reduced $Y^{\pm 1}$-sequence *for* $r$. We set $Y$-length$(r) := \ell$ and $Y_{|y}$-length$(r) := \big|\{i \in \{1, 2, \ldots, \ell\} : y_i \in \{y\}^{\pm 1}\}\big|$, for each $y \in Y^{\pm 1}$.

For each cyclic word $r$ in $F$, there exists some cyclically reduced $Y^{\pm 1}$-sequence $(y_1, \ldots, y_\ell)$ such that $y_1 y_2 \cdots y_\ell \in r$. Here, $(y_1, \ldots, y_\ell)$ is unique up to cyclic permutation, as may be seen by considering another such sequence, a conjugation equality, and any possible cancellation therein. We set $Y$-length$(r) := \ell$ and $Y_{|y}$-length$(r) := \big|\{i \in \{1, 2, \ldots, \ell\} : y_i \in \{y\}^{\pm 1}\}\big|$ for $y \in Y^{\pm 1}$.

Recall that $R$ is a finite set of words in $F$. We set $h(Y) := \sum_{r\in R} Y$-length$(r)$ and $h(Y_{|y}) := \sum_{r\in R} Y_{|y}$-length$(r)$. It is clear that $h(Y_{|y^{-1}}) = h(Y_{|y})$ and $h(Y) = \sum_{y\in Y} h(Y_{|y})$. $\square$

## 3. Gersten's graphs and Whitehead's proof

3.1. **Definitions.** Consider any subset $V$ of $F$. We let $X\curvearrowright V$, $V\curvearrowleft Y$, and $X\curvearrowright V\curvearrowleft Y$ denote the full subgraphs of $X\curvearrowright F$, $F\curvearrowleft Y$, and $X\curvearrowright F\curvearrowleft Y$ with vertex-set $V$ respectively, where a subgraph $\Gamma_0$ of a graph $\Gamma$ is said to be *full* if $\Gamma_0$ contains every $\Gamma$-edge whose initial and terminal vertices lie in $\Gamma_0$. By Theorem 2.3, $X\curvearrowright F$ and $F\curvearrowleft Y$ are trees; thus, $X\curvearrowright V$ and $V\curvearrowleft Y$ are forests. A subset of $X\curvearrowright F\curvearrowleft Y$ is said to be 1-*containing* it it contains 1. We say that $V$ is an $(X,Y)$-*translator* if $V$ is a 1-containing $F$-generating set such that $X\curvearrowright V$ and $V\curvearrowleft Y$ are trees. In this event, we let $(X\curvearrowright V\curvearrowleft Y)_{\geqslant 3}$ denote the set of elements of $V-\{1\}$ which have $(X\curvearrowright V\curvearrowleft Y)$-valence at least 3. Notice that $|V-\{1\}| \geqslant \operatorname{rank}(F)$, since $V-\{1\}$ generates $F$.

Clearly, $F$ itself is an $(X,Y)$-translator. Let $\kappa$ denote the minimum value for $|V-\{1\}|$ as $V$ ranges over the set of all $(X,Y)$-translators. If $\kappa > \operatorname{rank}(F)$, we define $\mathrm{d}(X,Y) := \kappa$. Otherwise, $\kappa = \operatorname{rank}(F)$, and we then define $\mathrm{d}(X,Y)$ to be the minimum value for $\big|(X\curvearrowright V\curvearrowleft Y)_{\geqslant 3}\big|$ as $V$ ranges over the set of all $(X,Y)$-translators of cardinal $1+\operatorname{rank}(F)$. □

3.2. **Lemma** (Gersten(1987))**.** $\mathrm{d}(X,Y)\in\mathbb{N}$.

*Proof* (Krstić(1989), here streamlined). For each finite 1-containing subset $W$ of $F$, we let $\breve{\mathbb{X}}W$ and $\breve{\mathbb{Y}}W$ denote the vertex-sets of the 1-containing components of the forests $X\curvearrowright W$ and $W\curvearrowleft Y$ respectively; also, we let $\overline{\mathbb{X}}W$ and $\overline{\mathbb{Y}}W$ denote the vertex-sets of the tree-closures of $W$ in the trees $X\curvearrowright F$ and $F\curvearrowleft Y$ respectively, where the *tree-closure of* $W$ in a tree is the inclusion-smallest subtree which includes $W$. We have now defined four self-maps of the set of finite 1-containing subsets of $F$.

Set $\tilde{Y} := \{1\}\cup Y^{\pm 1}$ and $V := \breve{\mathbb{Y}}\overline{\mathbb{X}}\,\overline{\mathbb{Y}}\,\overline{\mathbb{X}}\,\tilde{Y}$.

Then $V$ is a finite 1-containing subset of $F$, $V\curvearrowleft Y$ is a tree, and

$$V = \breve{\mathbb{Y}}\big(\overline{\mathbb{X}}(\overline{\mathbb{Y}}\,\overline{\mathbb{X}}\,\tilde{Y})\big) \supseteq \breve{\mathbb{Y}}\big((\overline{\mathbb{Y}}\,\overline{\mathbb{X}}\,\tilde{Y})\big) = \overline{\mathbb{Y}}\,\overline{\mathbb{X}}\,\tilde{Y} \supseteq \overline{\mathbb{X}}\,\tilde{Y} \supseteq \tilde{Y}. \tag{3.1}$$

In particular, $V$ is an $F$-generating set.

We now prove that $(\breve{\mathbb{X}}V)\cdot\tilde{Y} \subseteq \breve{\mathbb{X}}(V\cdot\tilde{Y})$. Let $y\in\tilde{Y}$ and $v\in\breve{\mathbb{X}}V$; thus, $V\supseteq\overline{\mathbb{X}}\{v,1\}$. Then $V\supseteq\overline{\mathbb{X}}\{v,1,y^{-1}\}$, since $V\supseteq\overline{\mathbb{X}}\,\tilde{Y}$, by (3.1). Now $V\cdot\tilde{Y}\supseteq(\overline{\mathbb{X}}\{v,1,y^{-1}\})\cdot y = \overline{\mathbb{X}}\{v\cdot y, y, 1\}$, since $X\curvearrowright F$ is a right $F$-tree. Thus, $v\cdot y\in\breve{\mathbb{X}}(V\cdot\tilde{Y})$, as desired.

It follows from the definition of $\breve{\mathbb{Y}}$ that $V$ is the inclusion-smallest 1-containing subset of $F$ such that $\overline{\mathbb{X}}\,\overline{\mathbb{Y}}\,\overline{\mathbb{X}}\,\tilde{Y}\cap V\cdot\tilde{Y} \subseteq V$. Now $\breve{\mathbb{X}}V$ is a 1-containing subset of $V$, and

$$\overline{\mathbb{X}}\,\overline{\mathbb{Y}}\,\overline{\mathbb{X}}\,\tilde{Y}\cap(\breve{\mathbb{X}}V)\cdot\tilde{Y} \;\subseteq\; \breve{\mathbb{X}}(\overline{\mathbb{X}}\,\overline{\mathbb{Y}}\,\overline{\mathbb{X}}\,\tilde{Y})\cap\breve{\mathbb{X}}(V\cdot\tilde{Y}) \;\subseteq\; \breve{\mathbb{X}}(\overline{\mathbb{X}}\,\overline{\mathbb{Y}}\,\overline{\mathbb{X}}\,\tilde{Y}\cap V\cdot\tilde{Y}) \;\subseteq\; \breve{\mathbb{X}}(V).$$

It follows from the minimality property of $V$ that $\breve{\mathbb{X}}V = V$. Thus, $X\curvearrowright V$ is a tree.

Hence, $V$ is a finite $(X,Y)$-translator. □

3.3. **Theorem** (Whitehead(1936b))**.** *With* Notation 1.1*, if $X$ and $Y$ are local-minimum points for $h$, then either $X^{\pm 1} = Y^{\pm 1}$ or some Whitehead transform $Y'$ of $Y$ satisfies $h(Y') = h(Y)$ and* $\mathrm{d}(X,Y') < \mathrm{d}(X,Y)$.

*Proof* (Whitehead(1936b), here translated). For all $v,g\in F$, we let $v \overset{X:g\cdot}{-\cdots\rightarrowtail} g\cdot v$ denote the unique reduced $(X\curvearrowright F)$-path from $v$ to $g\cdot v$, and $v \overset{\cdot g:Y}{-\cdots\rightarrowtail} v\cdot g$ denote the unique reduced $(F\curvearrowleft Y)$-path from $v$ to $v\cdot g$. If $g=1$, then these paths have length zero.

We shall obtain information about Whitehead transforms of $Y$ that are constructed using a procedure that depends on $\mathrm{d}(X,Y)$. We begin by describing features that apply whenever we have an $(X,Y)$-translator $V$.

For each $x \in X^{\pm 1}$, we set $\hat{\iota}_X x := x^{-1}{\cdot}V \cap V$ and $\hat{\tau}_X x := x{\cdot}V \cap V = x{\cdot}\hat{\iota}_X$. For each $y \in Y^{\pm 1}$, we set $\hat{\iota}_Y y := V{\cdot}y^{-1} \cap V$ and $\hat{\tau}_Y y := V{\cdot}y \cap V = \hat{\iota}_Y{\cdot}y$.

Consider any $y \in Y^{\pm 1}$. We shall now show that $X{\frown}(\hat{\iota}_Y y)$ and $X{\frown}(\hat{\tau}_Y y)$ are subtrees of the tree $X{\frown}V$, and that $\big(X{\frown}(\hat{\iota}_Y y)\big){\cdot}y = X{\frown}(\hat{\tau}_Y y)$. We first show that $\hat{\iota}_Y y \neq \emptyset$. Since $V$ generates $F$, there exists some $u \in V - \langle Y - \{y\}^{\pm 1}\rangle$. Let $(y_1, y_2, \ldots, y_\ell)$ be the reduced $Y^{\pm 1}$-sequence for $u$; thus, there exists some $k \in \{1, 2, \ldots, \ell\}$ such that $\{y_k\}^{\pm 1} = \{y\}^{\pm 1}$. The reduced $(F{\frown}Y)$-path from 1 to $u$ is then

$$1 = u_0 \xrightarrow{\cdot y_1} u_1 \xrightarrow{\cdot y_2} u_2 \cdots \xrightarrow{\cdot y_\ell} u_\ell = u;$$

this is a $(V{\frown}Y)$-path, since the endpoints lie in $V$, and the subpath $u_{k-1} \xrightarrow{\cdot y_k} u_k$ meets $\hat{\iota}_Y y$, as desired. Now consider any $v, w \in \hat{\iota}_Y y$. Then $v{\cdot}y, w{\cdot}y \in \hat{\tau}_Y y$. Let $(x_\ell, x_{\ell-1}, \ldots, x_1)$ be the reduced $X^{\pm 1}$-sequence for $w{\cdot}v^{-1} = (w{\cdot}y){\cdot}(v{\cdot}y)^{-1}$. The reduced $(X{\frown}F)$-paths

$$v = v_0 \xrightarrow{x_1 \cdot} v_1 \cdots \xrightarrow{x_\ell \cdot} v_\ell = w \quad \text{and} \quad v{\cdot}y = v_0{\cdot}y \xrightarrow{x_1 \cdot} v_1{\cdot}y \cdots \xrightarrow{x_\ell \cdot} v_\ell{\cdot}y = w{\cdot}y$$

are $(X{\frown}V)$-paths, since their endpoints lie in $V$. Thus, $\{v_0, v_1, \ldots, v_\ell\}{\cdot}\{1, y\} \subseteq V$. This proves that $X{\frown}(\hat{\iota}_Y y)$ is a subtree of the tree $X{\frown}V$. Also, $\big(X{\frown}(\hat{\iota}_Y y)\big){\cdot}y = X{\frown}(\hat{\tau}_Y y)$, and $X{\frown}(\hat{\tau}_Y y)$ is a subtree of the tree $X{\frown}V$.

Analogous assertions hold for $(\hat{\iota}_X x){\frown}Y$ and $(\hat{\tau}_X x){\frown}Y$.

Consider any $v, w \in V$ and any $(X{\frown}V{\frown}Y)$-path $p$ from $v$ to $w$. Let $(x_1{\cdot}, x_2{\cdot}, \ldots, x_\ell{\cdot})$ be the sequence of $X^{\pm 1}{\cdot}$-labels encountered along $p$. We call the $X^{\pm 1}$-sequence $(x_\ell, \ldots, x_2, x_1)$ the *left $X^{\pm 1}$-label of $p$*, and call $g := x_\ell \cdots x_2 x_1$ the *left $F$-label of $p$*. Let $({\cdot}y_1, {\cdot}y_2, \ldots, {\cdot}y_{\ell'})$ be the sequence of ${\cdot}Y^{\pm 1}$-labels encountered along $p$. We call the $Y^{\pm 1}$-sequence $(y_1, y_2, \ldots, y_{\ell'})$ the *right $Y^{\pm 1}$-label of $p$*, and call $g' := y_1 y_2 \cdots y_{\ell'}$ the *right $F$-label of $p$*. It is not difficult to see that $gvg' = w$ in $F$. We may use ordinary path reductions and assume that $p$ is a reduced $(X{\frown}V{\frown}Y)$-path without changing the left and right $F$-labels. If the right $Y^{\pm 1}$-label of $p$ is still not a reduced $Y^{\pm 1}$-sequence, then $p$ has some subpath $p'$ of the form

$$u \xrightarrow{\cdot y} u{\cdot}y \overset{X:h\cdot}{-\cdots\rightarrow} h{\cdot}u{\cdot}y \xrightarrow{\cdot y^{-1}} h{\cdot}u,$$

for some $h \in F - \{1\}$. Since $X{\frown}V$ is a tree, we have the $(X{\frown}V)$-path $p''$ which is $u \overset{X:h\cdot}{-\cdots\rightarrow} h{\cdot}u$. The $(X{\frown}V{\frown}Y)$-path obtained from $p$ by replacing $p'$ with $p''$ is said to be a *right $Y$-reduction of $p$*. This gives a shorter $(X{\frown}V{\frown}Y)$-path from $v$ to $w$ with the same left and right $F$-labels, the same left $X^{\pm 1}$-label, and a shorter right $Y^{\pm 1}$-label. Similar considerations give *left $X$-reductions of $p$*. Any $(X{\frown}V{\frown}Y)$-path yields an $(X{\frown}V{\frown}Y)$-path with reduced left $X^{\pm 1}$- and right $Y^{\pm 1}$-labels after applying ordinary, left $X$-, and right $Y$-reductions sufficiently often.

Similar considerations apply for *cyclic* ordinary, left $X$-, and right $Y$-reductions of closed $(X{\frown}V{\frown}Y)$-paths; these operations may change where the path is based.

We write $\mathrm{Paths}(X{\frown}V{\frown}Y)$ to denote the set of all $(X{\frown}V{\frown}Y)$-paths. We construct a map $F \to \mathrm{Paths}(X{\frown}V{\frown}Y)$ which assigns to each $g \in F$ a closed $(X{\frown}V{\frown}Y)$-path based at 1 whose left $X^{\pm 1}$-label is the reduced $X^{\pm 1}$-sequence for $g^{-1}$, and whose right $Y^{\pm 1}$-label is the reduced $Y^{\pm 1}$-sequence for $g$. One way to do this is first to choose, for each $x \in X$, some $v_x \in \hat{\iota}_X x$, and then the $(X{\frown}V{\frown}Y)$-path

$$1 \overset{\cdot x\cdot v_x:Y}{-\cdots\rightarrow} x{\cdot}v_x \xrightarrow{x^{-1}\cdot} v_x \overset{\cdot v_x^{-1}:Y}{-\cdots\rightarrow} 1$$

has left $X^{\pm 1}$-label $(x^{-1})$, which is the reduced $X^{\pm 1}$-sequence for $x^{-1}$. Using inversion and concatenation of paths, we may now assign to each $g \in F$ a closed $(X{\frown}V{\frown}Y)$-path based at 1 whose left $X^{\pm 1}$-label is the reduced $X^{\pm 1}$-sequence for $g^{-1}$. The left $F$-label is then $g^{-1}$, and the right $F$-label must then be $g$. By applying right $Y$-reductions, we obtain a closed $(X\curvearrowright V \curvearrowleft Y)$-path based at 1 whose left $X^{\pm 1}$-label is still the reduced $X^{\pm 1}$-sequence for $g^{-1}$, whose right $Y^{\pm 1}$-label is a reduced $Y^{\pm 1}$-sequence, and whose right $F$-label is still $g$. We call this the *chosen* $(X\curvearrowright V \curvearrowleft Y)$*-path representing* $g$. The reduced $Y^{\pm 1}$-sequence for $g$ and the reverse of the reduced $X^{\pm 1}$-sequence for $g^{-1}$ have been interlaced to form a closed $(X\curvearrowright V \curvearrowleft Y)$-path based at 1. For our counting purposes, the reverse of the reduced $X^{\pm 1}$-sequence for $g^{-1}$ contains the same information as the reduced $X^{\pm 1}$-sequence for $g$; previous authors amalgamated $F\curvearrowleft(X^{-1})$ and $F\curvearrowleft Y$ over their vertex-sets via the inversion map on $F$.

We now construct a map $R \to \mathrm{Paths}(X\curvearrowright V \curvearrowleft Y)$. We map each straight word $r$ contained in $R$ to the chosen $(X\curvearrowright V \curvearrowleft Y)$-path representing $r$. For each cyclic word $r$ contained in $R$, we choose an element $g$ of $r$, and consider the chosen $(X\curvearrowright V \curvearrowleft Y)$-path representing $g$, and apply cyclic ordinary, left $X$-, and right $Y$-reductions, until we get a closed $(X\curvearrowright V \curvearrowleft Y)$-path whose right $Y^{\pm 1}$-label is a cyclically reduced $Y^{\pm 1}$-sequence and whose left $X^{\pm 1}$-label is a cyclically reduced $X^{\pm 1}$-sequence; then the right $F$-label is a conjugate of $g$, and the left $F$-label is a conjugate of $g^{-1}$. We call this the *chosen* $(X\curvearrowright V \curvearrowleft Y)$*-path representing* $r$.

Our map $R \to \mathrm{Paths}(X\curvearrowright V \curvearrowleft Y)$ gives $R$ the structure of a set of closed $(X\curvearrowright V \curvearrowleft Y)$-paths, with the straight words being based at 1. We may now speak of the number of times an element of $R$ traverses a given $(X\curvearrowright V \curvearrowleft Y)$-edge $e$, and by summing over all elements of $R$, we may speak of the number of times $R$ traverses $e$, and denote the number by $\widetilde{h}(e)$.

For each length-one $(X\curvearrowright V)$-path $v \overset{x\cdot}{\rightarrowtail} w$, we let $v \overset{x\cdot}{\rightleftharpoons} w$ denote the $(X\curvearrowright V)$-edge it traverses, and set $\widetilde{h}(v \overset{x\cdot}{\rightarrowtail} w) := \widetilde{h}(v \overset{x\cdot}{\rightleftharpoons} w)$; thus, $w \overset{x^{-1}\cdot}{\rightleftharpoons} v$ equals $v \overset{x\cdot}{\rightleftharpoons} w$, and $\widetilde{h}(w \overset{x^{-1}\cdot}{\rightarrowtail} v)$ equals $\widetilde{h}(v \overset{x\cdot}{\rightarrowtail} w)$. For any element $x$ of $X^{\pm 1}$, and subsets $V_0$ and $V_1$ of $V$, we set

$$\widetilde{h}(V_0 \overset{x\cdot}{\rightarrowtail} V_1) := \sum_{v \in V_0 \cap (x^{-1}\cdot V_1)} \widetilde{h}(v \overset{x\cdot}{\rightarrowtail} x\cdot v).$$

Notice that $\widetilde{h}(V \overset{x\cdot}{\rightarrowtail} V) = \widetilde{h}(\hat{\iota}_X x \overset{x\cdot}{\rightarrowtail} \hat{\tau}_X x) = h(X_{|x})$, since the left $X^{\pm 1}$-labels of the chosen $(X\curvearrowright V \curvearrowleft Y)$-paths are reduced, and cyclically reduced for cyclic words.

Analogous notation applies with $Y$ in place of $X$.

For any $x_* \in X^{\pm 1}$, $y_* \in Y^{\pm 1}$, and $v_* \in \hat{\iota}_X x_*$, we say that $(v_*, x_*, y_*)$ is a *first-stage triple*, and associate to it all of the following data.

The $(X\curvearrowright V)$-edge $v_* \overset{x_*\cdot}{\rightleftharpoons} x_*\cdot v_*$ is called the *disconnecting* edge. Let $V_0$ denote the vertex-set of the 1-containing component of the forest $(X\curvearrowright V) - \{v_* \overset{x_*\cdot}{\rightleftharpoons} x_*\cdot v_*\}$, and set $V_1 := V - V_0$, the vertex-set of the other component. We let $\chi : V \to \{0,1\}$ denote the characteristic function of $V_1$; thus, $v \in V_{\chi(v)}$ for each $v \in V$. We define a map $\hat{\chi} : \hat{\iota}_Y(Y^{\pm 1}) \to \{0,1\}$ as follows. For $j \in \{1,2\}$, let $Y^{\pm 1}_{j\text{-part}}$ denote the set of those $y \in Y^{\pm 1} - \{y_*\}$ such that $\chi$ restricted to $\hat{\iota}_Y y$ takes exactly $j$ values. For each $y \in Y^{\pm 1}_{1\text{-part}}$, $\chi$ restricted to $\hat{\iota}_Y y$ takes exactly one value, and we define $\hat{\chi}(\hat{\iota}_Y y)$ to be that value. Let $\chi_F : F \to \{0,1\}$ denote the characteristic function of the vertex-set of that component of $(X\curvearrowright F) - \{v_* \overset{x_*\cdot}{\rightleftharpoons} x_*\cdot v_*\}$ which does *not* contain 1; the restriction of $\chi_F$ to $V$ is then $\chi$. For each $y \in Y^{\pm 1}_{2\text{-part}}$, we define $\hat{\chi}(\hat{\iota}_Y y) := \chi_F(v_*\cdot y^{-1})$. To complete the definition of the map $\hat{\chi} : \hat{\iota}_Y(Y^{\pm 1}) \to \{0,1\}$, we set $\hat{\chi}(\hat{\iota}_Y y_*) := 1 - \hat{\chi}(\hat{\iota}_Y y_*^{-1})$.

Let $y_\dagger$ denote the element of $\{y_*\}^{\pm 1}$ such that $\hat{\chi}(\hat{\iota}_Y y_\dagger) = 0$; hence, $\hat{\chi}(\hat{\tau}_Y y_\dagger) = \hat{\chi}(\hat{\iota}_Y y_\dagger^{-1}) = 1$. For each $y \in Y^{\pm 1} - \{y_\dagger\}^{\pm 1}$, we set $y' := y_\dagger^{\hat{\chi}(\hat{\iota}_Y y)}\cdot y\cdot y_\dagger^{-\hat{\chi}(\hat{\tau}_Y y)}$, while, for each $y \in \{y_\dagger\}^{\pm 1}$, we set $y' := y$. We then set $Y' := \{y' \mid y \in Y\}$. Thus, $Y'$ is a Whitehead transform of $Y$. Since $Y$ is a local-minimum point for $h$, $h(Y) \leqslant h(Y')$. It is not difficult to see from the definition of $Y'$ that, for each $y \in Y^{\pm 1} - \{y_\dagger\}^{\pm 1}$, $h(Y'_{|y'}) \geqslant h(Y_{|y})$. Similarly, $h(Y_{|y}) \geqslant h(Y'_{|y'})$, and, hence, equality holds. Now

$$0 \leqslant h(Y') - h(Y) = h(Y'_{|y'_\dagger}) - h(Y_{|y_\dagger}). \tag{3.2}$$

We next define a map $\xi : \operatorname{Paths}(X \curvearrowright V \curvearrowleft Y) \to \operatorname{Paths}(X \curvearrowright F \curvearrowleft Y')$. It suffices to define $\xi$ on $V$ and on the set of length-one $(X \curvearrowright V \curvearrowleft Y)$-paths, and then concatenate paths.

We define $\xi$ on $V$ by

$$V = V_0 \cup V_1 \to V_0 \cup V_1 \cdot y_\dagger^{-1} \subseteq F, \qquad v \mapsto \xi(v) := v \cdot y_\dagger^{-\chi(v)}.$$

Consider a length-one $(X \curvearrowright V \curvearrowleft Y)$-path of the form $v \xrightarrow{\cdot y} w$, $y \in Y^{\pm 1}$. We define $\xi(v \xrightarrow{\cdot y} w)$ to be $\xi(v) \xrightarrow{\cdot \xi(v)^{-1} \cdot \xi(w) : Y'} \cdots \xi(w)$. Notice that

$$y_\dagger^{\hat\chi(\hat\iota_Y y)} \cdot y \cdot y_\dagger^{-\hat\chi(\hat\tau_Y y)} = y'^{\delta(y)} \text{ where } \delta(y) := \begin{cases} 1 & \text{if } y \in Y^{\pm 1} - \{y_\dagger\}^{\pm 1}, \\ 0 & \text{if } y \in \{y_\dagger\}^{\pm 1}. \end{cases}$$

As $\xi(v) = v \cdot y_\dagger^{-\chi(v)}$ and

$$\xi(w) = w \cdot y_\dagger^{-\chi(w)} = v \cdot y \cdot y_\dagger^{-\chi(w)} = \xi(v) \cdot y_\dagger^{\chi(v)} \cdot y \cdot y_\dagger^{-\chi(w)} = \xi(v) \cdot y_\dagger^{\chi(v) - \hat\chi(\hat\iota_Y y)} \cdot y'^{\delta(y)} \cdot y_\dagger^{\hat\chi(\hat\tau_Y y) - \chi(w)},$$

$$\xi(v \xrightarrow{\cdot y} w) \text{ equals } \xi(v) \xrightarrow{\cdot y_\dagger'^{\chi(v) - \hat\chi(\hat\iota_Y y)} \cdot y'^{\delta(y)} \cdot y_\dagger'^{\hat\chi(\hat\tau_Y y) - \chi(w)} : Y'} \cdots \xi(w). \tag{3.3}$$

Consider now a length-one $(X \curvearrowright V \curvearrowleft Y)$-path of the form $v \xrightarrow{x \cdot} w$, $x \in X^{\pm 1}$. Here,

$$\xi(v) = v \cdot y_\dagger^{-\chi(v)} \text{ and } \xi(w) = w \cdot y_\dagger^{-\chi(w)} = x \cdot v \cdot y_\dagger^{-\chi(w)} = x \cdot \xi(v) \cdot y_\dagger^{\chi(v) - \chi(w)}.$$

(3.4) We shall define $\xi(v \xrightarrow{x \cdot} w)$ to be $\xi(v) \xrightarrow{x \cdot} x \cdot \xi(v) \xrightarrow{\cdot y_\dagger'^{\chi(v) - \chi(w)} : Y'} \cdots \xi(w)$

$$\text{or } \xi(v) \xrightarrow{\cdot y_\dagger'^{\chi(v) - \chi(w)} : Y'} \cdots \xi(v) \cdot y_\dagger'^{\chi(v) - \chi(w)} \xrightarrow{x \cdot} \xi(w).$$

Recall that $\chi(v) = \chi(w)$ unless $v \overset{x \cdot}{\rightleftharpoons} w$ is the disconnecting edge $v_* \overset{x_* \cdot}{\rightleftharpoons} x_* \cdot v_*$. If $\chi(v) = \chi(w)$, then $\xi(v \xrightarrow{x \cdot} w)$ equals $\xi(v) \xrightarrow{x \cdot} \xi(w)$. Later, we shall have enough information to choose between the two options and define $\xi(v_* \xrightarrow{x_* \cdot} x_* \cdot v_*)$ precisely.

Now $\xi$ will convert $(X \curvearrowright V \curvearrowleft Y)$-paths into $(X \curvearrowright F \curvearrowleft Y')$-paths without changing the left $X^{\pm 1}$-labels, and, hence, without changing the left $F$-labels. Since $\xi(1) = 1 \cdot y_\dagger^{-\chi(1)} = 1$, we see that $R$ is now represented by closed $(X \curvearrowright F \curvearrowleft Y')$-paths. We do not claim that the right $Y'^{\pm 1}$-labels are reduced, but the image of $R$ in $\operatorname{Paths}(X \curvearrowright F \curvearrowleft Y')$ does give an upper bound for $h(Y'_{|y'_\dagger})$. On carefully considering (3.4) and (3.3), and noting that the $y'^{\delta(y)}$-terms contribute no $y'_\dagger$-terms, we see that

$$h(Y'_{|y'_\dagger}) \leqslant \widetilde{h}(v_* \xrightarrow{x_* \cdot} x_* \cdot v_*) + \sum_{y \in Y^{\pm 1}} \widetilde{h}(V_{1 - \hat\chi(\hat\iota_Y y)} \xrightarrow{\cdot y} V).$$

Since $h(Y_{|y_\dagger}) = h(Y_{|y_*})$, we see from (3.2) that

$$0 \leqslant h(Y') - h(Y) \leqslant \widetilde{h}(v_* \xrightarrow{x_* \cdot} x_* \cdot v_*) - h(Y_{|y_*}) + \sum_{y \in Y^{\pm 1}} \widetilde{h}(V_{1 - \hat\chi(\hat\iota_Y y)} \xrightarrow{\cdot y} V). \tag{3.5}$$

We now consider two cases.

**Case 1:** $\mathrm{d}(X, Y) \leqslant \operatorname{rank} F$.

Here, we assume that $|V - \{1\}| = \operatorname{rank} F$ and $\big|(X \curvearrowright V \curvearrowleft Y)_{\geqslant 3}\big| = \mathrm{d}(X, Y)$.

Since $V \curvearrowleft Y$ is a tree, we have $\sum_{y \in Y} |\hat\iota_Y y| = |\operatorname{E}(V \curvearrowleft Y)| = |V| - 1 = |Y|$. For each $y \in Y$, $|\hat\iota_Y y| \geqslant 1$; hence, $|\hat\iota_Y y| = 1$. Here in Case 1, for each $y \in Y^{\pm 1}$, we write $\iota_Y y$ to denote the unique element of $\hat\iota_Y y$, and similarly for $\tau_Y y$, and analogously with $X$ in place of $Y$.

As an abelian group, $F/[F,F]$ is freely generated by the image of any $F$-basis. Hence, there exists a unique map $n_{X,Y} : X \times Y \to \mathbb{Z}$, $(x,y) \mapsto n_{x,y}$, such that, for each $y \in Y$,

$$y{\cdot}[F,F] = \textstyle\prod_{x\in X}\big((x{\cdot}[F,F])^{n_{x,y}}\big) \text{ in } F/[F,F];$$

we set $X$-absupp$(y) := \{x \in X \mid n_{x,y} \neq 0\}$. By choosing bijections from $\{1,2,\ldots,\operatorname{rank} F\}$ to $X$ and to $Y$, we may view the map $n_{X,Y}$ as an invertible matrix over $\mathbb{Z}$, and view every bijection $\varphi : X \xrightarrow{\sim} Y$, $x \mapsto \varphi x$, as a permutation of $\{1,2,\ldots,\operatorname{rank} F\}$. Then

$$\sum_{\varphi:X\xrightarrow{\sim}Y} (\operatorname{sign}(\varphi){\cdot}\textstyle\prod_{x\in X} n_{x,\varphi x}) = \operatorname{Det}(n_{X,Y}) \in \{1,-1\}.$$

There thus exists some bijection $\psi : X \xrightarrow{\sim} Y$ such that $\prod_{x\in X} n_{x,\psi x} \neq 0$; we fix such a $\psi$ throughout Case 1. Hence, $x \in X$-absupp$(\psi x)$ for each $x \in X$.

Consider any $x_* \in X$, and set $y_* := \psi(x_*) \in Y$ and $v_* := \iota_X x_* \in V$. We say that $(v_*, x_*, y_*)$ is a *second-stage Case* 1 *triple*. We have all the data associated to a first-stage triple.

Let us first show that, for each $y \in Y^{\pm1}$, $\hat\chi(\hat\iota_Y y) = \chi(\iota_Y y)$. Clearly $Y^{\pm1}_{\text{2-part}} = \emptyset$; hence, if $y \in Y^{\pm1}{-}\{y_*\} = Y^{\pm1}_{\text{1-part}}$, then $\hat\chi(\hat\iota_Y y) = \chi(\iota_y y)$, as desired. It remains to consider $y_*$. Now

$$\hat\chi(\hat\iota_Y y_*) = 1 - \hat\chi(\hat\iota_Y y_*^{-1}) = 1 - \chi(\iota_Y y_*^{-1}) = 1 - \chi(\tau_Y y_*),$$

and it suffices to show that $\chi(\tau_Y y_*) \neq \chi(\iota_Y y_*)$. Let $(x_\ell, \ldots, x_2, x_1)$, $\ell \in \mathbb{N}$, be the reduced $X^{\pm1}$-sequence for $(\tau_Y y_*){\cdot}(\iota_Y y_*)^{-1}$. Then

$$\iota_Y y_*{\cdot}y_*{\cdot}(\iota_Y y_*)^{-1} = (\tau_Y y_*){\cdot}(\iota_Y y_*)^{-1} = x_\ell \cdots x_2 x_1.$$

Hence, $y_*{\cdot}[F,F] = \prod_{k=1}^{\ell}(x_k{\cdot}[F,F])$. Since $x_* \in X$-absupp$(\psi(x_*))$ and $\psi(x_*) = y_*$, there exists some $k \in \{1,2,\ldots,\ell\}$ such that $\{x_k\}^{\pm1} = \{x_*\}^{\pm1}$. The reduced $(X{\curvearrowright}F)$-path

$$\iota_Y y_* = v_0 \xrightarrow{x_1\cdot} v_1 \xrightarrow{x_2\cdot} \cdots \xrightarrow{x_{\ell-1}\cdot} v_{\ell-1} \xrightarrow{x_\ell\cdot} v_\ell = x_\ell \cdots x_1{\cdot}\iota_Y y_* = \tau_Y y_*$$

is the unique reduced $(X{\curvearrowright}V)$-path from $\iota_Y y_*$ to $\tau_Y y_*$, and it traverses $v_{k-1} \overset{x_k\cdot}{\rightleftharpoons} v_k$, which is $v_* \overset{x_*\cdot}{\rightleftharpoons} x_*{\cdot}v_*$, which is the disconnecting edge. Hence, $\chi(\iota_Y y_*) \neq \chi(\tau_Y y_*)$, as desired.

Since $y_* = \psi(x_*)$, here in Case 1, (3.5) takes the form

$$0 \leqslant h(Y'){-}h(Y) \leqslant h(X_{|x_*}){-}h(Y_{|\psi(x_*)}). \tag{3.6}$$

Since $x_*$ is arbitrary, $0 \leqslant h(X_{|x}){-}h(Y_{|\psi x})$ for each $x \in X$. Thus,

$$0 \leqslant \sum_{x\in X}\big(h(X_{|x}){-}h(Y_{|\psi x})\big) = h(X) - h(Y).$$

By the interchangeability of $X$ and $Y$, we then have $h(X) - h(Y) = 0$. It follows in turn that $h(X_{|x}){-}h(Y_{|\psi x}) = 0$ for each $x \in X$. By (3.6), $h(Y') = h(Y)$, as desired.

Consider the subcase where, for each $y \in Y^{\pm1}$ such that $\iota_Y y = 1$, the element $\tau_Y y$ of $V$ has $(X{\curvearrowright}V{\curvearrowleft}Y)$-valence exactly two, and therefore $(X{\curvearrowright}V)$-valence exactly one and $(V{\curvearrowleft}Y)$-valence exactly one. The latter means that $V^{\pm1}{-}\{1\} = Y^{\pm1}$, and the former then means that $V^{\pm1}{-}\{1\} = X^{\pm1}$. We then have $X^{\pm1} = Y^{\pm1}$, which is one of the desired conclusions; here, $\mathrm{d}(X,Y) = 0$. It remains to consider the subcase where there exists some $y_\dagger \in Y^{\pm1}$ such that $\iota_Y y_\dagger = 1$ and $\tau_Y y_\dagger \in (X{\curvearrowright}V{\curvearrowleft}Y)_{\geqslant3}$. We fix such a $y_\dagger$, and take $y_* \in Y \cap \{y_\dagger\}^{\pm1}$, $x_* := \psi^{-1}(y_*)$, and $v_* := \iota_X x_*$. We say that $(v_*, x_*, y_*)$ is a *third-stage Case* 1 *triple*.

By (3.3), for each $y \in Y^{\pm1} - \{y_\dagger\}^{\pm1}$, we have $\xi(\iota_Y y \xrightarrow{\cdot y} \tau_Y y)$ equals $\xi(\iota_Y y) \xrightarrow{\cdot y'} \xi(\tau_Y y)$, while $\xi(\iota_Y y_\dagger \xrightarrow{\cdot y_\dagger} \tau_Y y_\dagger)$ equals $\xi(\iota_Y y_\dagger) \overset{\cdot 1:Y'}{-\cdots\to} \xi(\tau_Y y_\dagger)$.

Set $x_\dagger := x_*^{(-1)^{\chi(v_*)}}$; thus, $\iota_X x_\dagger \overset{x_\dagger\cdot}{\rightleftharpoons} \tau_X x_\dagger$ equals $v_* \overset{x_*\cdot}{\rightleftharpoons} x_*{\cdot}v_*$, $\chi(\iota_X x_\dagger) = 0$, and $\chi(\tau_X x_\dagger) = 1$. In (3.4), for each $x \in X^{\pm1} - \{x_\dagger\}^{\pm1}$, $\xi(\iota_X x \xrightarrow{x\cdot} \tau_X x)$ equals $\xi(\iota_X x) \xrightarrow{x\cdot} \xi(\tau_X x)$, while we now choose $\xi(\iota_X x_\dagger \xrightarrow{x_\dagger\cdot} \tau_X x_\dagger)$ to be equal to $\xi(\iota_X x_\dagger) \xrightarrow{x_\dagger\cdot} \tau_X x_\dagger \xrightarrow{\cdot y_\dagger'^{-1}} \xi(\tau_X x_\dagger)$.

Set $V' := \xi(V) \cup \hat{\tau}_X x_\dagger = V_0 \cup V_1{\cdot}y_\dagger^{-1} \cup \hat{\tau}_X x_\dagger \subseteq F$. We shall see that $V'$ is an $(X, Y')$-translator. Since $\hat{\tau}_X x_\dagger \subseteq V_1$, we see that $V'$ is a finite, 1-containing, $F$-generating set. Thus, $|V'| \geqslant |V|$. Since $\hat{\iota}_Y y_\dagger \subseteq V_0 \cap V_1{\cdot}y_\dagger^{-1}$, we see that $|V'| = |V|$ and $\tau_X x_\dagger \notin V_0 \cup V_1{\cdot}y_\dagger^{-1}$.

It is clear that $\xi\big(\mathrm{Paths}(X{\curvearrowright}V{\curvearrowleft}Y)\big) \subseteq \mathrm{Paths}(X{\curvearrowright}V'{\curvearrowleft}Y')$. Let us examine the graphs $X{\curvearrowright}V'{\curvearrowleft}Y'$, $X{\curvearrowright}V'$, and $V'{\curvearrowleft}Y'$. From the form that $\xi$ takes here, we see that $X{\curvearrowright}V'{\curvearrowleft}Y'$ is obtained from $X{\curvearrowright}V{\curvearrowleft}Y$ by first subdividing the edge $\iota_X x_\dagger \overset{x_\dagger\cdot}{\rightleftharpoons} \tau_X x_\dagger$, and secondly collapsing the edge $\iota_Y y_\dagger \overset{\cdot y_\dagger}{\rightleftharpoons} \tau_Y y_\dagger$. The graph $X{\curvearrowright}V'$ is thus obtained from the tree $X{\curvearrowright}V$ by first removing an edge, leaving two components with vertex-sets $V_0$ and $V_1$, secondly identifying one vertex of $V_0$ with one vertex of $V_1$, and thirdly attaching one new vertex and one new edge incident to the new vertex and an old vertex. Hence, $X{\curvearrowright}V'$ is a tree. The graph $V'{\curvearrowleft}Y'$ is obtained from the tree $V{\curvearrowleft}Y$ by first collapsing one edge identifying its vertices, and secondly attaching one new vertex and one new edge incident to the new vertex and an old vertex. Thus, $V'{\curvearrowleft}Y'$ is a tree. Hence, $V'$ is an $(X, Y')$-translator.

Finally, $\big|(X{\curvearrowright}V{\curvearrowleft}Y)_{\geqslant 3}\big| > \big|(X{\curvearrowright}V'{\curvearrowleft}Y')_{\geqslant 3}\big|$, since the newly created vertex has $(X{\curvearrowright}V'{\curvearrowleft}Y')$-valence two, while the two old vertices which become identified are $\tau_Y y_\dagger \in (X{\curvearrowright}V{\curvearrowleft}Y)_{\geqslant 3}$ and $\iota_Y y_\dagger = 1$. Hence, $\mathrm{d}(X, Y) > \mathrm{d}(X, Y')$.

**Case 2:** $\mathrm{d}(X, Y) > \operatorname{rank} F$.

Here, we assume that $|V - \{1\}| = \mathrm{d}(X, Y)$. Hence, $|V - \{1\}| > \operatorname{rank} F = |Y|$.

Since $V{\curvearrowleft}Y$ is a tree, $\sum\limits_{y\in Y} |\hat{\iota}_Y y| = |\,\mathrm{E}(V{\curvearrowleft}Y)| = |V| - 1 > |Y|$. There then exists some $y_* \in Y^{\pm1}$ such that $|\hat{\iota}_Y y_*| \geqslant 2$. The tree $X{\curvearrowright}(\hat{\iota}_Y y_*)$ must then contain some edge $v_* \overset{x_*\cdot}{\rightleftharpoons} x_*{\cdot}v_*$, and then the tree $X{\curvearrowright}(\hat{\tau}_Y y_*) = \big(X{\curvearrowright}(\hat{\iota}_Y y_*)\big){\cdot}y_*$ contains the edge $v_*{\cdot}y_* \overset{x_*\cdot}{\rightleftharpoons} x_*{\cdot}v_*{\cdot}y_*$, giving a diagram

$$\begin{array}{ccc} v_*{\cdot}y_* & \xrightarrow{x_*\cdot} & x_*{\cdot}v_*{\cdot}y_* \\ {\scriptstyle \cdot y_*}\big\uparrow & & \big\uparrow{\scriptstyle \cdot y_*} \\ v_* & \xrightarrow{x_*\cdot} & x_*{\cdot}v_* \end{array}$$

of length-one $(X{\curvearrowright}V{\curvearrowleft}Y)$-paths. We say that $(v_*, x_*, y_*)$ is a *second-stage Case* 2 *triple.* We now have all the data associated with a first-stage triple.

If $y_*^{-1} \in Y_{\text{1-part}}$, then $\hat{\chi}(\hat{\iota} y_*^{-1}) = \chi(v_*{\cdot}y_*)$, because $v_*{\cdot}y_* \in \hat{\iota} y_*^{-1}$. If $y_*^{-1} \in Y_{\text{2-part}}^{\pm1}$, then, by definition, $\hat{\chi}(\hat{\iota} y_*^{-1}) = \chi_F(v_*{\cdot}y_*) = \chi(v_*{\cdot}y_*)$. This proves that $\hat{\chi}(\hat{\iota} y_*^{-1}) = \chi(v_*{\cdot}y_*)$; hence, $\hat{\chi}(\hat{\iota} y_*) = 1 - \chi(v_*{\cdot}y_*)$.

Let $\mathrm{west}_{(v_*,x_*)}(\hat{\iota}_Y y_*)$ and $\mathrm{east}_{(v_*,x_*)}(\hat{\iota}_Y y_*)$ denote the vertex-sets of the components of $\big(X{\curvearrowright}(\hat{\iota}_Y y_*)\big) - \{v_* \overset{x_*\cdot}{\rightleftharpoons} x_*{\cdot}v_*\}$ which contain $v_*$ and $x_*{\cdot}v_*$ respectively. Let $\mathrm{proper}(v_*, x_*, y_*)$ denote the intersection of $\hat{\iota}_Y y_*$ with the component of $\big(X{\curvearrowright}F\big) - \{v_* \overset{x_*\cdot}{\rightleftharpoons} x_*{\cdot}v_*\}$ which does *not* contain $v_*{\cdot}y_*$ and, hence, intersects $V$ in $V_{1-\chi(v_*\cdot y_*)}$. Since $\hat{\chi}(\hat{\iota} y_*) = 1 - \chi(v_*{\cdot}y_*)$, $\mathrm{proper}(v_*, x_*, y_*) = \hat{\iota}_Y y_* \cap V_{\hat{\chi}(\hat{\iota}_Y y_*)} \in \{\mathrm{west}_{(v_*,x_*)}(\hat{\iota}_Y y_*), \mathrm{east}_{(v_*,x_*)}(\hat{\iota}_Y y_*)\}$.

Let $\mathrm{south}_{(v_*,y_*)}(\hat{\iota}_X x_*)$ and $\mathrm{north}_{(v_*,y_*)}(\hat{\iota}_X x_*)$ denote the vertex-sets of the components of $\big((\hat{\iota}_X x_*){\curvearrowleft}Y\big) - \{v_* \overset{\cdot y_*}{\rightleftharpoons} v_*{\cdot}y_*\}$ which contain $v_*$ and $v_*{\cdot}y_*$ respectively. It is not difficult to show that $\text{south}_{(v_*,y_*)}(\hat{\iota}_X x_*) = \{v_*\}$ if and only if $v_*$ has $(X\frown V)$-valence one, if and only if $Y^{\pm 1}_{\text{2-part}} = \emptyset$.

We now consider an arbitrary $y \in Y^{\pm 1}_{\text{2-part}}$. Thus, $\hat{\chi}(\hat{\iota}_Y y) = \chi_F(v_* \cdot y^{-1})$. We have a diagram

$$\begin{array}{ccc}
v_*\cdot y_* & \xrightarrow{x_*\cdot} & x_*\cdot v_*\cdot y_* \\
{\scriptstyle \cdot y_*}\uparrow & & \uparrow{\scriptstyle \cdot y_*} \\
v_* & \xrightarrow{x_*\cdot} & x_*\cdot v_* \\
{\scriptstyle \cdot y}\downarrow & & \downarrow{\scriptstyle \cdot y} \\
v_*\cdot y & \xrightarrow{x_*\cdot} & x_*\cdot v_*\cdot y
\end{array}$$

of length-one $(X\frown V \frown Y)$-paths. Notice that $(v_*\cdot y, x_*, y^{-1})$ is a second-stage Case 2 triple, and that $\text{proper}(v_*\cdot y, x_*, y^{-1})$ is then the intersection of $\hat{\iota}_Y y^{-1}$ with that component of $\big(X\frown F\big) - \{v_*\cdot y \overset{x_*\cdot}{\rightleftharpoons} x_*\cdot v_*\cdot y\}$ which does *not* contain $v_*$. On right multiplying by $y^{-1}$, we see that $\big(\text{proper}(v_*\cdot y, x_*, y^{-1})\big)\cdot y^{-1}$ is the intersection of $\hat{\iota}_Y y$ with that component of $\big(X\frown F\big) - \{v_* \overset{x_*\cdot}{\rightleftharpoons} x_*\cdot v_*\}$ which does *not* contain $v_*\cdot y^{-1}$ and, hence, intersects $V$ in $V_{1-\chi_F(v_*\cdot y^{-1})}$. Since $\hat{\chi}(\hat{\iota}_Y y) = \chi_F(v_*\cdot y^{-1})$, $\big(\text{proper}(v_*\cdot y, x_*, y^{-1})\big)\cdot y^{-1} = \hat{\iota}_Y y \cap V_{1-\hat{\chi}(\hat{\iota}_Y y)}$.

Now

$$\begin{aligned}
&\sum_{y\in Y^{\pm 1}} \widetilde{h}(V_{1-\hat{\chi}(\hat{\iota}_Y y)} \xrightarrow{\cdot y} V) \\
&= \widetilde{h}(V_{1-\hat{\chi}(\hat{\iota}_Y y_*)} \xrightarrow{\cdot y_*} V) + \sum_{y\in Y^{\pm 1}_{\text{2-part}}} \widetilde{h}(V_{1-\hat{\chi}(\hat{\iota}_Y y)} \xrightarrow{\cdot y} V) \\
&= h(Y_{|y_*}) - \widetilde{h}(V_{\hat{\chi}(\hat{\iota}_Y y_*)} \xrightarrow{\cdot y_*} V) + \sum_{y\in Y^{\pm 1}_{\text{2-part}}} \widetilde{h}\Big(\big(\text{proper}(v_*\cdot y, x_*, y^{-1})\big)\cdot y^{-1} \xrightarrow{\cdot y} V\Big) \\
&= h(Y_{|y_*}) - \widetilde{h}\big(\text{proper}(v_*, x_*, y_*) \xrightarrow{\cdot y_*} V\big) + \sum_{y\in Y^{\pm 1}_{\text{2-part}}} \widetilde{h}\big(\text{proper}(v_*\cdot y, x_*, y^{-1}) \xrightarrow{\cdot y^{-1}} V\big).
\end{aligned}$$

Thus, here in Case 2, (3.5) takes the form

$$\begin{aligned}
(3.7)\qquad 0 \leqslant h(Y') - h(Y) \leqslant \widetilde{h}(v_* \xrightarrow{x_*\cdot} x_*\cdot v_*) - \widetilde{h}\big(\text{proper}(v_*, x_*, y_*) \xrightarrow{\cdot y_*} V\big) \\
+ \sum_{y\in Y^{\pm 1}_{\text{2-part}}} \widetilde{h}\big(\text{proper}(v_*\cdot y, x_*, y^{-1}) \xrightarrow{\cdot y^{-1}} V\big).
\end{aligned}$$

In particular,

$$\widetilde{h}\big(\text{proper}(v_*, x_*, y_*) \xrightarrow{\cdot y_*} V\big) \leqslant \widetilde{h}(v_* \xrightarrow{x_*\cdot} x_*\cdot v_*) + \sum_{y\in Y^{\pm 1}_{\text{2-part}}} \widetilde{h}\big(\text{proper}(v_*\cdot y, x_*, y^{-1}) \xrightarrow{\cdot y^{-1}} V\big).$$

Since

$$\widetilde{h}\big(\text{south}_{(v_*,y_*)}(\hat{\iota}_X x_*) \xrightarrow{x_*\cdot} V\big) = \widetilde{h}(v_* \xrightarrow{x_*\cdot} x_*\cdot v_*) + \sum_{y\in Y^{\pm 1}_{\text{2-part}}} \widetilde{h}\big(\text{south}_{(v_*\cdot y, y^{-1})}(\hat{\iota}_X x_*) \xrightarrow{x_*\cdot} V\big),$$

it may be seen by induction on $|\,\text{south}_{(v_*,y_*)}(\hat{\iota}_X x_*)|$ that

$$\widetilde{h}\big(\text{proper}(v_*, x_*, y_*) \xrightarrow{\cdot y_*} V\big) \leqslant \widetilde{h}\big(\text{south}_{(v_*,y_*)}(\hat{\iota}_X x_*) \xrightarrow{x_*\cdot} V\big).$$

Let us write

$$\widetilde{h}\text{-west} := \widetilde{h}\big(\text{west}_{(v_*,x_*)}(\hat{\iota}_Y y_*) \xrightarrow{\cdot y_*} V\big) \text{ and } \widetilde{h}\text{-south} := \widetilde{h}\big(\text{south}_{(v_*,y_*)}(\hat{\iota}_X x_*) \xrightarrow{x_*\cdot} V\big),$$

and similarly for $\widetilde{h}$-east and $\widetilde{h}$-north. We have shown that

$$\min\{\widetilde{h}\text{-west}, \widetilde{h}\text{-east}\} \leqslant \widetilde{h}\big(\mathrm{proper}(v_*, x_*, y_*) \xrightarrow{\cdot y_*} V\big) \leqslant \widetilde{h}\text{-south}\,.$$

Replacing $(v_*, x_*, y_*)$ with the second-stage Case 2 triple $(v_*{\cdot}y_*, x_*, y_*^{-1})$ interchanges south and north, and we find that $\min\{\widetilde{h}\text{-west}, \widetilde{h}\text{-east}\} \leqslant \widetilde{h}\text{-north}$. Hence,

$$\min\{\widetilde{h}\text{-west}, \widetilde{h}\text{-east}\} \leqslant \min\{\widetilde{h}\text{-south}, \widetilde{h}\text{-north}\}.$$

Interchanging $X$ and $Y$ interchanges south and west, as well as north and east, and we find

$$\min\{\widetilde{h}\text{-south}, \widetilde{h}\text{-north}\} \leqslant \min\{\widetilde{h}\text{-west}, \widetilde{h}\text{-east}\} \leqslant \widetilde{h}\big(\mathrm{proper}(v_*, x_*, y_*) \xrightarrow{\cdot y_*} V\big). \tag{3.8}$$

We now choose a third-stage Case 2 triple as follows. Consider the preceding $x_*$. Thus, $(\hat\iota_X x_*)\curvearrowleft Y$ is a finite tree that has at least one edge and, hence, at least two valence-one vertices. There then exists a valence-one $\big((\hat\iota_X x_*)\curvearrowleft Y\big)$-vertex $v_*$ such that $\widetilde{h}(v_* \xrightarrow{x_*\cdot} x_*{\cdot}v_*) \leqslant h(X_{|x_*})/2$. Taking $v_* \overset{x_*\cdot}{\rightleftharpoons} x_*{\cdot}v_*$ as the disconnecting edge determines a map $\chi : V \to \{0,1\}$. If $\chi(v_*) = 0$, we fix this $x_*$ and this $v_*$. If $\chi(v_*) = 1$, we replace $(x_*, v_*)$ with $(x_*^{-1}, x_*{\cdot}v_*)$, and then fix this new $x_*$ and $v_*$; then $\chi(v_*) = 0$. Now $\chi(x_*{\cdot}v_*) = 1$. Let $y_*$ denote the element of $Y^{\pm 1}$ such that $v_* \overset{\cdot y_*}{\rightleftharpoons} v_*{\cdot}y_*$ is the unique edge of $(\hat\iota_X x_*)\curvearrowleft Y$ that is incident to $v_*$. Now $(v_*, x_*, y_*)$ is a second-stage Case 2 triple, $Y^{\pm 1}_{\text{2-part}} = \emptyset$, $\widetilde{h}(v_* \xrightarrow{x_*\cdot} x_*{\cdot}v_*) \leqslant h(X_{|x_*})/2$, $\chi(v_*) = 0$, and $\chi(x_*{\cdot}v_*) = 1$; we say that $(v_*, x_*, y_*)$ is a *third-stage Case 2 triple.* Since $Y^{\pm 1}_{\text{2-part}} = \emptyset$,

$$\widetilde{h}\text{-south} = \widetilde{h}(v_* \xrightarrow{x_*\cdot} x_*{\cdot}v_*) \leqslant h(X_{|x_*}) - \widetilde{h}(v_* \xrightarrow{x_*\cdot} x_*{\cdot}v_*) = \widetilde{h}\text{-north};$$

thus, $\widetilde{h}(v_* \xrightarrow{x_*\cdot} x_*{\cdot}v_*) = \min\{\widetilde{h}\text{-south}, \widetilde{h}\text{-north}\}$. Also, since $Y^{\pm 1}_{\text{2-part}} = \emptyset$, it follows from (3.7) and (3.8) that $h(Y') = h(Y)$, as desired.

Set $V' := \xi(V) = V_0 \cup V_1{\cdot}y_\dagger^{-1}$. It suffices to show that $V'$ is an $(X, Y')$-translator with $|V'| < |V|$.

Since $Y^{\pm 1}_{\text{2-part}} = \emptyset$, (3.3) says that $\xi(v \xrightarrow{\cdot y} v{\cdot}y)$ equals $\xi(v) \xrightarrow{\cdot y'} \xi(v{\cdot}y)$ if $y \in Y^{\pm 1} - \{y_*\}^{\pm 1}$ and $v \in \hat\iota_Y y$, and that

$$\xi(v \xrightarrow{\cdot y_*} v{\cdot}y_*) \;=\; \begin{cases} \xi(v) \xrightarrow{\cdot 1:Y'} \xi(v{\cdot}y_*) & \text{if } v \in V_{\hat\chi(\hat\iota_Y y_*)} \cap \hat\iota_Y y_*, \\ \xi(v) \xrightarrow{\cdot y'_*} \xi(v{\cdot}y_*) & \text{if } v \in V_{1-\hat\chi(\hat\iota_Y y_*)} \cap \hat\iota_Y y_*. \end{cases}$$

By (3.4), $\xi(v \xrightarrow{x\cdot} x{\cdot}v)$ equals $\xi(v) \xrightarrow{x\cdot} \xi(x{\cdot}v)$ if $x \in X^{\pm 1}, v \in \hat\iota_X x$, and $v \overset{x\cdot}{\rightleftharpoons} x{\cdot}v$ is not equal to $v_* \overset{x_*\cdot}{\rightleftharpoons} x_*{\cdot}v_*$. It remains to specify $\xi(v_* \xrightarrow{x_*\cdot} x_*{\cdot}v_*)$. Clearly, $v_*{\cdot}y_* \overset{x_*\cdot}{\rightleftharpoons} x_*{\cdot}v_*{\cdot}y_*$ is not equal to $v_* \overset{x_*\cdot}{\rightleftharpoons} x_*{\cdot}v_*$; hence, $\chi(x_*{\cdot}v_*{\cdot}y_*) = \chi(v_*{\cdot}y_*) = 1 - \hat\chi(\hat\iota_Y y_*) \in \{0,1\}$. We then have two subcases.

If $\chi(v_*{\cdot}y_*) = \chi(x_*{\cdot}v_*{\cdot}y_*) = 1 - \hat\chi(\hat\iota_Y y_*) = 1$, then $y_\dagger = y_*$ and

$$\xi(v_*{\cdot}y_*) = v_*{\cdot}y_*{\cdot}y_\dagger^{-1} = v_*, \qquad \xi(x_*{\cdot}v_*{\cdot}y_*) = x_*{\cdot}v_*{\cdot}y_*{\cdot}y_\dagger^{-1} = x_*{\cdot}v_*,$$
$$\xi(v_*) = v_*, \qquad \xi(x_*{\cdot}v_*) = x_*{\cdot}v_*{\cdot}y_\dagger^{-1} = x_*{\cdot}v_*{\cdot}y_*^{-1}.$$

Here, $\xi : V \to V'$ is not injective, and we define $\xi(v_* \xrightarrow{x_*\cdot} x_*{\cdot}v_*)$ to be

$$\xi(v_*) \xrightarrow{x_*\cdot} \xi(x_*{\cdot}v_*{\cdot}y_*) \xrightarrow{\cdot y_*'^{-1}} \xi(x_*{\cdot}v_*)$$

in $\mathrm{Paths}(X\curvearrowright V'\curvearrowleft Y')$; notice that $\xi(x_*{\cdot}v_*{\cdot}y_*) \xrightarrow{\cdot y_*'^{-1}} \xi(x_*{\cdot}v_*)$ equals $\xi\big(x_*{\cdot}v_*{\cdot}y_* \xrightarrow{\cdot y_*^{-1}} x_*{\cdot}v_*\big)$.

If $\chi(v_*\cdot y_*) = \chi(x_*\cdot v_*\cdot y_*) = 1 - \hat{\chi}(\hat{\imath}_Y y_*) = 0$, then $y_\dagger = y_*^{-1}$ and

$$\xi(v_*\cdot y_*) = v_*\cdot y_*, \qquad \xi(x_*\cdot v_*\cdot y_*) = x_*\cdot v_*\cdot y_*,$$
$$\xi(v_*) = v_*, \qquad \xi(x_*\cdot v_*) = x_*\cdot v_*\cdot y_\dagger^{-1} = x_*\cdot v_*\cdot y_*.$$

Here, $\xi : V \to V'$ is not injective, and we define $\xi(v_* \xrightarrow{x_*\cdot} x_*\cdot v_*)$ to be

$$\xi(v_*) \xrightarrow{\cdot y'_*} \xi(v_*\cdot y_*) \xrightarrow{x_*\cdot} \xi(x_*\cdot v_*)$$

in $\mathrm{Paths}(X\curvearrowright V'\curvearrowleft Y')$; notice that $\xi(v_*) \xrightarrow{\cdot y'_*} \xi(v_*\cdot y_*)$ equals $\xi\big(v_* \xrightarrow{\cdot y_*} v_*\cdot y_*\big)$.

Since $\xi : V \to V'$ is surjective, but not injective, $|V'| < |V|$. Notice also that $y_* \in \langle V' \rangle$; hence, $V'$ generates $F$.

Let us examine the graphs $X\curvearrowright V'\curvearrowleft Y'$, $X\curvearrowright V'$, and $V'\curvearrowleft Y'$. From the form that $\xi$ takes here, we see that $X\curvearrowright V'\curvearrowleft Y'$ is obtained from $X\curvearrowright V\curvearrowleft Y$ by first removing one edge, secondly reattaching it elsewhere, and thirdly collapsing various edges. The graph $X\curvearrowright V'$ is thus obtained from the tree $X\curvearrowright V$ by first removing an edge, leaving components with vertex-sets $V_0$ and $V_1$, secondly reattaching the edge elsewhere, and thirdly identifying one or more vertices of $V_0$ with vertices of $V_1$. Hence, $X\curvearrowright V'$ is connected, and therefore a tree. The graph $V'\curvearrowleft Y'$ is obtained from the tree $V\curvearrowleft Y$ by collapsing edges; hence, $V'\curvearrowleft Y'$ is a tree. Thus, $V'$ is an $(X, Y')$-translator, and $\mathrm{d}(X, Y) > \mathrm{d}(X, Y')$. □

## References


Warren Dicks: Groups, trees and projective modules. Lecture Notes Math. **790**. Springer, Berlin (1980).

Warren Dicks: On free-group algorithms that sandwich a subgroup between free-product factors. J. Group Theory **17**, 13–28 (2014).

Warren Dicks: On Whitehead's first free-group algorithm, cutvertices, and free-product factorizations. `https://arxiv.org/abs/1704.05338`, 9 pages (2017).

M. Dehn: Über die Topologie des dreidimensionalen Raumes. Math. Ann. **69**, 137–168 (1910).

Ralph H. Fox: Free differential calculus I. Derivation in the free group ring. Ann. of Math. **57** (2), 547–560 (1953).

S. M. Gersten: Fixed points of automorphisms of free groups. Adv. in Math. **64**, 51–85 (1987).

Michael Heusener and Richard Weidmann: A remark on Whitehead's lemma. Preprint, 4 pages (2014).

P. J. Higgins and R. C. Lyndon: Equivalence of elements under automorphisms of a free group. Queen Mary College Mimeographed Notes, London, 5 pages (1962).

P. J. Higgins and R. C. Lyndon: Equivalence of elements under automorphisms of a free group. J. London Math. Soc. **8**, 254–258 (1974).

A. H. M. Hoare: On automorphisms of free groups I. J. London Math. Soc. (2) **38**, 277–285 (1988).

A. H. M. Hoare: On automorphisms of free groups II. J. London Math. Soc. (2) **42**, 226–236 (1990).

Sava Krstić: On graphs representing automorphisms of free groups. Proc. Amer. Math. Soc. **107**, 573–575 (1989).

J. Nielsen: Über die Isomorphismen unendlicher Gruppen ohne Relation. Math. Ann. **79**, 269–272 (1919).

Elvira Strasser Rapaport: On free groups and their automorphisms. Acta Math. **99**, 139–163 (1958).

John R. Stallings: Automorphisms of free groups in graph theory and topology. Talk given at the AMS meeting, SUNY, Albany, August 8–11 (1983).

Richard Stong: Diskbusting elements of the free group. Math. Res. Lett. **4**, 201–210 (1997).

J. H. C. Whitehead: On certain sets of elements in a free group. Proc. London Math. Soc. (2) **41**, 48–56 (1936a).

J. H. C. Whitehead: On equivalent sets of elements in a free group. Ann. of Math. (2) **37**, 782–800 (1936b).

Departament de Matemàtiques,
Universitat Autònoma de Barcelona,
E-08193 Bellaterra (Barcelona),
SPAIN

*Email address*: dicks@mat.uab.cat
*Home page*: http://mat.uab.cat/~dicks/