05.06.2012 reworking of the 17.05.2011 e-mail from me (Warren Dicks) to Igor Mineyev which gave a one-page simplification of the latter's proof of *the SHNC*, Walter Neumann's strengthened form of the Hanna Neumann conjecture. Mineyev's proof (made available online 07.05.2011, published as *Submultiplicativity and the Hanna Neumann conjecture*, Ann. of Math. **175** (2012), 393–414) was independent of Joel Friedman's proof (made available online 01.05.2011 at http://arxiv.org/abs/1105.0129). Mineyev's proof deals with ℓ^2-numbers and yields a general result, while my simplification uses Bass-Serre theory and applies only to the proof of the SHNC. (Mineyev's *Groups, graphs, and the Hanna Neumann conjecture,* J. Topol. Anal. **4** (2012), 1–12, does little more than combine part of Mineyev's proof and most of my simplification to give an intermediate-length proof of the SHNC that uses neither Hilbert modules nor graphs of groups.) The evolution of the reworked simplification given here benefitted from observations made, in chronological order, by Igor Mineyev, Martin Lustig, Yago Antolín, Lluís Bacardit, Gilbert Levitt, Oleg Bogopolski, Ilya Kapovich, Armando Martino, George Bergman, Zoran Šunić and Joel Friedman.

An ordered group. • We first turn to Satz I in Wilhelm Magnus' *Beziehungen zwischen Gruppen und Idealen in einem speziellen Ring*, Math. Ann. **111** (1935), 259–280, and transcribe the two-variable case of the argument.

Let $\mathbb{Z}\langle \dot{x},\, \dot{y}\rangle$ denote the free associative ring on two variables $\dot{x},\, \dot{y}$. Let $\mathfrak{a}$ denote the two-sided ideal generated by $\dot{x},\, \dot{y}$. Set $\mathbb{Z}\langle\langle \dot{x},\, \dot{y}\rangle\rangle := \varprojlim(\,\mathbb{Z}\langle \dot{x},\, \dot{y}\rangle/\mathfrak{a}^n : n \geqslant 1)$. Let $\{\dot{x},\, \dot{y}\}^*$ denote the free multiplicative monoid on $\dot{x},\, \dot{y}$. An element a of $\mathbb{Z}\langle\langle \dot{x},\, \dot{y}\rangle\rangle$ is just a function $\{\dot{x},\, \dot{y}\}^* \to \mathbb{Z}$, $\omega \mapsto a[\omega]$, expressed as a formal infinite sum $\sum\limits_{\omega\in\{\dot{x},\,\dot{y}\}^*} \omega{\cdot}a[\omega]$ to facilitate ring operations. The power-series rings $\mathbb{Z}[[\dot{x}]]$ and $\mathbb{Z}[[\dot{y}]]$ are subrings of $\mathbb{Z}\langle\langle \dot{x},\, \dot{y}\rangle\rangle$, as is $\mathbb{Z}\langle \dot{x},\, \dot{y}\rangle$. Let $\mathrm{PU}(\mathbb{Z}\langle\langle \dot{x},\, \dot{y}\rangle\rangle)$ denote the group of units with constant term 1. Set

$$x := 1{+}\,\dot{x} \in \mathrm{PU}(\mathbb{Z}[[\dot{x}]]),\ y := 1{+}\,\dot{y} \in \mathrm{PU}(\mathbb{Z}[[\dot{y}]]), \text{ and } F := \langle x, y\rangle \leqslant \mathrm{PU}(\mathbb{Z}\langle\langle \dot{x},\, \dot{y}\rangle\rangle).$$

If $a = x^{m_1}y^{n_1}\cdots x^{m_j}y^{n_j}$, then $a[(\dot{x}\dot{y})^j] = m_1n_1\cdots m_jn_j$, as can be seen by expressing a as

$$\Big(1{+}m_1\dot{x}{+}\tbinom{m_1}{2}\dot{x}^2{+}\cdots\Big)\ \Big(1{+}n_1\dot{y}{+}\tbinom{n_1}{2}\dot{y}^2{+}\cdots\Big)\ \cdots\ \Big(1{+}m_j\dot{x}{+}\tbinom{m_j}{2}\dot{x}^2{+}\cdots\Big)\ \Big(1{+}n_j\dot{y}{+}\tbinom{n_j}{2}\dot{y}^2{+}\cdots\Big).$$

Thus, $x^{m_1}y^{n_1}\cdots x^{m_j}y^{n_j} \neq 1$ whenever $j \geqslant 1$ and $m_1n_1\cdots m_jn_j \neq 0$. It follows that $F = \langle x, y \mid \quad \rangle$.

• We next turn to the fourth sentence of George M. Bergman's *Ordering coproducts of groups and semigroups,* J. Algebra **133** (1990), 313–339, and transcribe a lexicographic case of the argument.

Endow $\{\dot{x},\, \dot{y}\}^*$ with the length-lexicographic ordering $\prec$ with $\dot{x} \prec \dot{y}$. Then $\{\dot{x},\, \dot{y}\}^*$ is an ordered monoid that is well-ordered. For each $a \in \mathbb{Z}\langle\langle \dot{x},\, \dot{y}\rangle\rangle{-}\{0\}$, let ω_a denote the $\prec$-smallest element of $\{\omega \in \{\dot{x},\, \dot{y}\}^* : a[\omega] \neq 0\}$. Set $P := \{a \in \mathbb{Z}\langle\langle \dot{x},\, \dot{y}\rangle\rangle{-}\{0\} : a[\omega_a] > 0\}$. Then $\mathbb{Z}\langle\langle \dot{x},\, \dot{y}\rangle\rangle = -P \vee \{0\} \vee P$, where $\vee$ denotes disjoint union. Also, P is closed under addition and multiplication. Thus, $\mathbb{Z}\langle\langle \dot{x},\, \dot{y}\rangle\rangle$ is an ordered ring with positive cone P, and we again let $\prec$ denote the ordering. Since $\mathrm{PU}(\mathbb{Z}\langle\langle \dot{x},\, \dot{y}\rangle\rangle)$ is then the group of positive units, it is an ordered group, as is the subgroup F; here, $(f_1 \prec f_2) \Leftrightarrow (f_2{-}f_1 \in P) \Leftrightarrow (1 \prec f_1^{-1}f_2)$.

• *Remarks.* A similar construction gives a Bergman-Magnus ordered free group of any desired rank. References cited on the third page of Bernhard H. Neumann's *On ordered division rings,* Trans. Amer. Math. Soc. **66** (1949), 202–252, show that orderings of free groups were found by G. Birkhoff (1946, unpubl.), A. Tarski (1946, unpubl.), H. Shimbireva (publ. 1947), K. Iwasawa (publ. 1948), and B. H. Neumann (publ. earlier in 1949).

Notation. For any free group G, $\bar{\mathrm{r}}(G) := \max\{\mathrm{rank}(G){-}1, 0\}$. We use the symbol ∞ to denote $\aleph_0$.

Throughout, $(F, \prec)$ will denote the ordered group defined above. We view the diagonal subgroup $\Delta \leqslant F\times F$ as an alter ego of F. We extend $\prec$ from Δ to $F\times F$ using the lexicographic ordering with $\prec$ in each coordinate. For $g \in \Delta$ and $d,\, e \in F\times F$, if $d \prec e$, then $gd \prec ge$. If a subset E of $F\times F$ has a $\prec$-largest element e, we write $\max(E) := e$; the usage of $\min(E)$ is analogous. We form the left F-graph with vertex F-set F, edge Δ-set $F\times F$, and incidence functions given by $\iota(f_1, f_2) := f_1$ and $\tau(f_1, f_2) := f_2$. The *Cayley tree of F with respect to* $\{x, y\}$, denoted $\mathbb{T}(F, \{x, y\})$, is the F-subgraph with vertex set F and edge set $\{(f, fz) : f \in F, z \in \{x, y\}\}$.

Remark. It can be shown that $\mathbb{T}(F, \{x, y\})$ is a tree; Magnus' argument shows that $\mathbb{T}(F, \{x, y\})$ is acyclic.

If v, w are vertices of $\mathbb{T}(F, \{x, y\})$, $[v, w]$ denotes the $\subset$-smallest subtree of $\mathbb{T}(F, \{x, y\})$ containing $\{v, w\}$.

Suppose that T is a subgraph of $\mathbb{T}(F, \{x, y\})$. A *line* is a tree in which each vertex has valence two, and a *subline of* T is a subgraph of T which is a line. An edge e of T is said to be a *bridge of* T if e is the $\prec$-largest edge of some subline of T. We let $\mathrm{V}T$, $\mathrm{E}T$, and $\mathrm{B}T$ denote the sets of vertices, edges, and bridges of T, respectively; if T is a tree, we let $\mathrm{I}T$ denote the set of components of the forest $T{-}\mathrm{B}T$, called *islands of* T.

Remark. Clearly, if T' is a subgraph of T, then $\mathrm{B}T' \subseteq \mathrm{B}T$.

Suppose that $G \leqslant F$. We write $G_T := \{g \in G : gT = T\} \leqslant G$; then T is a G_T-subgraph of $\mathbb{T}(F, \{x, y\})$. We let $\mathbb{T}(G)$ denote the intersection of all the G-subtrees of $\mathbb{T}(F, \{x, y\})$.

Remarks. Clearly, $\mathrm{B}\mathbb{T}(G)$ is a G-subset of $\mathrm{E}\mathbb{T}(G)$. Clearly, if $H \leqslant G$, then $\mathbb{T}(H) \subseteq \mathbb{T}(G)$ and, hence, $\mathrm{B}\mathbb{T}(H) \subseteq \mathrm{B}\mathbb{T}(G)$. Clearly, if $G = \{1\}$, then $\mathbb{T}(G) = \emptyset$. It can be shown that if $\mathrm{rank}(G) = 1$, then $\mathbb{T}(G)$ is a line. Hence, if $G \neq \{1\}$, then $\mathbb{T}(G)$ is the $\subset$-smallest G-subtree of $\mathbb{T}(F, \{x, y\})$. It can be shown that if $\mathrm{rank}(G) < \infty$, then $|G\backslash \mathrm{E}\mathbb{T}(G)| < \infty$.

The island theorem. *If* $\{1\} \neq G \leqslant F$ *and* $C \in \mathrm{I}\mathbb{T}(G)$*, then the chain of implications* (a)$\Rightarrow$(b)$\Rightarrow$(c) *holds for the following conditions:* (a) $\mathrm{rank}(G) < \infty$*;* (b) $G_C \neq \{1\}$*;* (c) $\mathrm{rank}(G_C) = 1$.

Proof. • (a)$\Rightarrow$(b). By (a), $|G\backslash \mathrm{E}\mathbb{T}(G)| < \infty$. Let δC denote the set of all bridges of $\mathbb{T}(G)$ incident to $\mathrm{V}C$.

Case 1: $|\delta C| = 0$. In this event, C is a component of $\mathbb{T}(G)$, $C = \mathbb{T}(G)$, $G_C = G \neq \{1\}$, and (b) holds.

Case 2: $0 < |\delta C| < \infty$. In this event, $\min(\delta C)$ exists. As $\min(\delta C)$ is a bridge of $\mathbb{T}(G)$, there exists some subline L of $\mathbb{T}(G)$ with $\max(\mathrm{E}L) = \min(\delta C)$. Clearly, $|\mathrm{E}L \cap \delta C| = 1$. Hence, one of the two (infinite) components of $L - \delta C$ is $L \cap C$. Then $|\mathrm{E}C| = \infty > |G\backslash \mathrm{E}\mathbb{T}(G)|$. Thus, there exist $d, e \in \mathrm{E}C$ such that $d \neq e$ and $Gd = Ge$. Here, there exists $g \in G$ such that $gd = e$, and then $e \in g\mathrm{E}C \cap \mathrm{E}C$. Hence, the islands gC and C of $\mathbb{T}(G)$ must be equal. Thus, $g \in G_C$. Moreover, $g \neq 1$, since $gd = e \neq d$. Hence, $g \in G_C - \{1\}$ and (b) holds.

Case 3: $|\delta C| = \infty$. For each $e \in \delta C$, there is a unique incidence function $\nu_e \in \{\iota, \tau\}$ such that $\nu_e e \in \mathrm{V}C$. Since $|\delta C| = \infty > |\{\iota, \tau\} \times (G\backslash \mathrm{E}\mathbb{T}(G))|$, there exist $d, e \in \delta C$ such that $d \neq e$, $\nu_d = \nu_e$, and $Gd = Ge$. Here, there exists $g \in G$ such that $gd = e$. Thus, $g\mathrm{V}C \ni g(\nu_d d) = \nu_d(gd) = \nu_e e \in \mathrm{V}C$. As in Case 2, (b) holds.

• (b)$\Rightarrow$(c). Choose a pair $(v, h) \in \mathrm{V}C \times (G_C - \{1\})$ for which $[v, hv]$ is $\subset$-minimal. Then $v \in \mathrm{V}\mathbb{T}(\langle h \rangle)$. By replacing h with h^{-1} if necessary, we further arrange that $\max(\mathrm{E}[v, h^{-1}v]) \succ \max(\mathrm{E}[v, hv])$. We shall show that $G_C = \langle h \rangle$ by showing that each $g \in G$ lies in $\langle h \rangle \cup (G - G_C)$. Choose a pair $(u, w) \in \mathrm{V}\mathbb{T}(\langle h \rangle) \times g\mathrm{V}\mathbb{T}(\langle h \rangle)$ for which $[u, w]$ is $\subset$-minimal. There exist unique $i, j \in \mathbb{Z}$ such that $h^{-i}u,\ h^{-j}(g^{-1}w) \in \mathrm{V}[v, hv] - \{hv\}$.

Case 1: $u = w$. Here, $[h^{-i}w,\ (h^{-j}g^{-1}h^{i})(h^{-i}w)] = [h^{-i}u,\ h^{-j}(g^{-1}w)] \subset [v, hv]$. Also, $g \in G_C$, since $C \cap gC \supseteq \mathbb{T}(\langle h \rangle) \cap g\mathbb{T}(\langle h \rangle) \supseteq \{w\}$. By the $\subset$-minimality of $[v, hv]$, $h^{-j}g^{-1}h^{i} = 1$. Now, $g = h^{i-j} \in \langle h \rangle$.

Case 2: $u \neq w$. Here, $\mathbb{T}(\langle h \rangle) \cap g\mathbb{T}(\langle h \rangle) = \emptyset$ and $[h^{i+1}v, gh^{j+1}v] = [h^{i+1}v, u] \cup [u, w] \cup [w, gh^{j+1}v]$. Set

$$M := [v, hv] \cup [hv, h^2v] \cup [h^2v, h^3v] \cup \cdots \qquad \text{and} \qquad L := h^{i+1}M \cup gh^{j+1}M \cup [h^{i+1}v, gh^{j+1}v].$$

Then L is a line. For each $n \in \mathbb{Z}$, set $e_n := \max(\mathrm{E}[h^n v, h^{n+1}v])$; as it was arranged that $e_{-1} \succ e_0$, we see that $e_n = h^{n+1}e_{-1} \succ h^{n+1}e_0 = e_{n+1}$. Hence, $\max(\mathrm{E}M) = e_0$ and $\max(\mathrm{E}L) = \max(\{e_{i+1}, ge_{j+1}\} \cup \mathrm{E}[h^{i+1}v, gh^{j+1}v])$. Then $\max(\mathrm{E}L)$ is a bridge of $\mathbb{T}(G)$ that must lie in $L - (C \cup gC)$, which is finite. Hence, $C \cup gC$ is not connected. ∎

The bridge theorem. *If* $G \leqslant F$*, then* $|G\backslash \mathrm{B}\mathbb{T}(G)| = \bar{\mathrm{r}}(G)$.

Proof. *Case* 1: $G = \{1\}$. In this event, we can say $|G\backslash \mathrm{B}\mathbb{T}(G)| = |G\backslash \mathrm{E}\mathbb{T}(G)| = \mathrm{rank}(G) = \bar{\mathrm{r}}(G) = 0$.

Case 2: $|G\backslash \mathrm{B}\mathbb{T}(G)| = \infty$. In this event, we can say $|G\backslash \mathrm{B}\mathbb{T}(G)| = |G\backslash \mathrm{E}\mathbb{T}(G)| = \mathrm{rank}(G) = \bar{\mathrm{r}}(G) = \infty$.

Case 3: $G \neq \{1\}$ and $|G\backslash \mathrm{B}\mathbb{T}(G)| < \infty$. In particular, $\mathbb{T}(G)$ is a G-subtree of $\mathbb{T}(F, \{x, y\})$. Let $\overline{\mathbb{T}}(G)$ denote the G-tree with edge set $\mathrm{B}\mathbb{T}(G)$ and vertex set $\mathrm{I}\mathbb{T}(G)$ viewed as the quotient G-tree of $\mathbb{T}(G)$ obtained by collapsing each island of $\mathbb{T}(G)$ to a vertex. By the Bass-Serre structure theorem, the G-tree $\overline{\mathbb{T}}(G)$ gives rise to a graph of groups $(\mathbf{G}(-), Y)$ whose fundamental group, $\pi(\mathbf{G}(-), Y)$, can be identified with G. Here, we have the following information: $|\mathrm{E}Y| = |G\backslash \mathrm{B}\mathbb{T}(G)| < \infty$; $|\mathrm{V}Y| \leqslant |\mathrm{E}Y| + 1 < \infty$; if $v \in \mathrm{V}Y$, then $\mathbf{G}(v) = G_C$ for some $C \in \mathrm{I}\mathbb{T}(G)$; and if $e \in \mathrm{E}Y$, then $\mathbf{G}(e) = \{1\}$. By collapsing a maximal subtree of Y to a single vertex, or by using the presentation of $\pi(\mathbf{G}(-), Y)$, we find that $G = G_0 * G_1$ where $G_0 = \mathop{\ast}\limits_{v \in \mathrm{V}Y} \mathbf{G}(v)$ and G_1 is a free group of rank $|\mathrm{E}Y| - |\mathrm{V}Y| + 1$. By the island theorem (b)$\Rightarrow$(c), the vertex groups are cyclic, and, hence, $\mathrm{rank}(G) < \infty$. By the island theorem (a)$\Rightarrow$(c), the vertex groups have rank one, and $\bar{\mathrm{r}}(G) = \mathrm{rank}(G) - 1 = |\mathrm{E}Y| = |G\backslash \mathrm{B}\mathbb{T}(G)|$. ∎

The Friedman-Mineyev SHNC theorem. *Let* G *be a free group, let* H *and* K *be subgroups of* G*, and let* S *be a subset of* G *such that the map* $S \to H\backslash G/K$, $s \mapsto HsK$*, is bijective. Then* $\sum\limits_{s \in S} \bar{\mathrm{r}}(H^s \cap K) \leqslant \bar{\mathrm{r}}(H) \cdot \bar{\mathrm{r}}(K)$.

Proof. Any free factor of G that contains $H \cup K$ also contains each $s \in S$ such that $H^s \cap K \neq \{1\}$, by malnormality. If H or K is uncountable, then the desired inequality holds. It follows that we may assume that G is countable. We then identify G with $\langle x^{y^i} : 0 \leqslant i < \mathrm{rank}(G) \rangle \leqslant F$.

If $s \in S$, then $\mathrm{B}\mathbb{T}(H^s \cap K) \subseteq \mathrm{B}\mathbb{T}(H^s) \cap \mathrm{B}\mathbb{T}(K) = (s^{-1}\mathrm{B}\mathbb{T}(H)) \cap \mathrm{B}\mathbb{T}(K)$. Suppose that $s_1, s_2 \in S$, $e_1 \in \mathrm{B}\mathbb{T}(H^{s_1} \cap K)$, and $e_2 \in \mathrm{B}\mathbb{T}(H^{s_2} \cap K)$; then we have the following chain of equivalences.

$$\begin{aligned}
&s_1 = s_2 \text{ and } (H^{s_1} \cap K)e_1 = (H^{s_2} \cap K)e_2 \\
&\quad \Leftrightarrow s_1 = s_2 \text{ and } \exists (h, k) \in H \times K \text{ such that } h^{s_1} = k \text{ and } ke_1 = e_2 \\
&\quad \Leftrightarrow \exists (h, k) \in H \times K \text{ such that } hs_1k^{-1} = s_2 \text{ and } ke_1 = e_2 \\
&\quad \Leftrightarrow \exists (h, k) \in H \times K \text{ such that } hs_1e_1 = s_2e_2 \text{ and } ke_1 = e_2 \\
&\quad \Leftrightarrow Hs_1e_1 = Hs_2e_2 \text{ and } Ke_1 = Ke_2.
\end{aligned}$$

Thus, we have a well-defined, injective map of sets

$$\bigvee_{s \in S} ((H^s \cap K)\backslash \mathrm{B}\mathbb{T}(H^s \cap K)) \hookrightarrow (H\backslash \mathrm{B}\mathbb{T}(H)) \times (K\backslash \mathrm{B}\mathbb{T}(K)), \ \ (H^s \cap K)e \mapsto (Hse, Ke) \text{ for } s \in S, e \in \mathrm{B}\mathbb{T}(H^s \cap K).$$

Now, $\sum\limits_{s \in S} \bar{\mathrm{r}}(H^s \cap K) \overset{\text{bridge thm.}}{=} \left| \bigvee\limits_{s \in S} ((H^s \cap K)\backslash \mathrm{B}\mathbb{T}(H^s \cap K)) \right| \leqslant \left| (H\backslash \mathrm{B}\mathbb{T}(H)) \times (K\backslash \mathrm{B}\mathbb{T}(K)) \right| \overset{\text{bridge thm.}}{=} \bar{\mathrm{r}}(H) \cdot \bar{\mathrm{r}}(K)$. ∎