Handbook of Set Theory

Foreman, Kanamori, and Magidor (eds.)

September 7, 2005

Contents

I. Elementary embeddings and algebra

Patrick Dehornoy

It has been observed for many years that computations with elementary embeddings entail some purely algebraic features—as opposed to the logical nature of the embeddings themselves. The key point is that the operation of applying an embedding to another one satisfies, when defined, the self-distributivity law $x(yz) = (xy)(xz)$. Using the specific properties of the elementary embeddings and their critical ordinals, hence under some large cardinal hypotheses, R. Laver established two purely algebraic results about sets equipped with a self-distributive operation (LD-systems), namely the decidability of the associated word problem in 1989, and the unboundedness of the periods in some finite LD-systems in 1993. The large cardinal assumption was eliminated from the first result by P. Dehornoy in 1992, using an argument that led to unexpected results about Artin braid groups; as for the second of Laver's results, no proof in ZF has been discovered so far, and the only result known to date is that it cannot be proved in Primitive Recursive Arithmetic.

1. Iterations of an elementary embedding

Our aim is to study the algebraic operation obtained by *applying* an elementary embedding to another one. For $j, k : V \prec M$, we can apply j to any set-restriction of k, and, in good cases, the images of these restrictions cohere so as to form a new elementary embedding that we shall denote by $j[k]$. It is then easy to see that the application operation so defined satisfies various algebraic identities.

Convention: All elementary embeddings we consider here are supposed to be distinct from the identity. An easy rank argument shows that every such embedding moves some ordinal; in particular, the least ordinal moved by j is called the critical ordinal of j, and denoted $\mathrm{crit}(j)$.

1.1. Kunen's bound and Axiom (I3)

If j is an elementary embedding of V into a proper subclass M, then $j[j]$, whenever it is defined, is an elementary embedding of M into a proper subclass M' of M, and it is not clear that $j[j]$ can be in turn applied to j, whose set-restrictions need not belong to M in general. So, if we wish the application operation on elementary embeddings to be everywhere defined, we should consider embeddings where the source and the target models coincide. Here comes an obstruction.

1.1 Proposition (Kunen [15]). *(AC) There is no $j : V \prec V$.*

Proof. Assume $j : V \prec V$. Let $\kappa_0 = \mathrm{crit}(j)$, and, recursively, $\kappa_{n+1} = j(\kappa_n)$. Write $\lambda = \sup_n \kappa_n$. By standard arguments, each κ_n is an inaccessible cardinal, so λ is a strong limit cardinal. Fix an injection i_n of $\mathcal{P}(\kappa_n)$ into λ. Then the mapping $X \mapsto (i_n(X \cap \kappa_n))_{n\in\omega}$ defines an injection of $\mathcal{P}(\lambda)$ into λ^ω. Using AC, we fix an enumeration $(\gamma_\xi, X_\xi)_{\xi<\nu}$ of $\lambda \times [\lambda]^\lambda$, and then inductively construct an injective sequence $(s_\xi)_{\xi<\nu}$ in λ^ω such that s_ξ belongs to $[X_\xi]^\omega$: this is possible because the cardinality of $\lambda \times [\lambda]^\lambda$ equals that of λ^ω. Let $f : \lambda^\omega \to \lambda$ be defined by $f(s) = \gamma_\xi$ for $s = s_\xi$, and $f(s) = 0$ for s not of the form s_ξ. Let $X \in [\lambda]^\lambda$. Then, for each $\gamma < \lambda$, there exists $\xi < \nu$ satisfying $(\gamma, X) = (\gamma_\xi, X_\xi)$. For this ξ, we have $s_\xi \in [X]^\omega$ by hypothesis, and $f(s_\xi) = \gamma_\xi$. Hence the function f, which lies in $V_{\lambda+2}$, has the property that the range of $f{\restriction}X^\omega$ is λ for every X in $[\lambda]^\lambda$.

Let us consider $j(f)$. We have $j(\lambda) = \sup_n \kappa_{n+1} = \lambda$, hence $j(f)$ is a function of λ into itself, and, as j is elementary, $j(f)$ has the property that, for every X in $[\lambda]^\lambda$, the range of $j(f){\restriction}X^\omega$ is λ. Now, let X be the set $\{\theta < \lambda\ ;\ \theta \in \mathrm{Im}(j)\}$. For every s in X^ω, we have $s_n \in \mathrm{Im}(j)$ for every n, hence $s = j(s')$ for some s', and $j(f)(s) = j(f)(j(s')) = j(f(s')) \in X$. As X is a proper subset of λ, the range of $j(f){\restriction}X^\omega$ is not λ, and we have got a contradiction. $\dashv$

We are thus led to considering weaker assumptions, involving embeddings that are defined on ranks rather than on the whole universe.

1.2 Definition (Gaifman, Solovay-Reinhardt-Kanamori [21]). Axiom (I3): For some δ, there exists $j : V_\delta \prec V_\delta$.

Assume $j : V_\delta \prec V_\delta$. Let $\kappa_0 = \mathrm{crit}(j)$, and $\kappa_n = j^n(\kappa_0)$. The proof of 1.1 shows that, letting $\lambda = \sup_n \kappa_n$, it is impossible (at least if AC is true) that the function called f there belongs to the target model of j. The function f belongs to $V_{\lambda+2}$, so $\delta \geqslant \lambda + 2$ is impossible, and the only remaining possibilities for (I3) are $\delta = \lambda$, and $\delta = \lambda + 1$. The second possibility subsumes the first:

1.3 Lemma. *Assume $j : V_{\delta+1} \prec V_{\delta+1}$. Then we have $j{\restriction}V_\delta : V_\delta \prec V_\delta$.*

Proof. First, $j(\delta) < \delta$ is impossible, so we necessarily have $j(\delta) = \delta$, and, therefore, $j{\restriction}V_\delta$ maps V_δ to itself. As for elementarity, an easy induction shows that, for $\vec{a}$ in V_δ and Φ a first-order formula, $V_\delta \models \Phi(\vec{a})$ is equivalent to $V_{\delta+1} \models \Phi^{V_\delta}(\vec{a})$, and, therefore, $V_\delta \models \Phi(\vec{a})$ is equivalent to $V_{\delta+1} \models \Phi^{V_\delta}(\vec{a})$, hence to $V_{\delta+1} \models \Phi^{V_{j(\delta)}}(j(\vec{a}))$, and finally to $V_\delta \models \Phi(j(\vec{a}))$. $\dashv$

Thus, without loss of generality, we can restrict to the case $j : V_\lambda \prec V_\lambda$ in the sequel, *i.e.*, when using (I3), we can add the assumption that δ is the supremum of the cardinals $j^n(\mathrm{crit}(j))$.

Before turning to the core of our study, let us observe that Axiom (I3) lies very high in the hierarchy of large cardinals.

1.4 Proposition. *Assume $j : V_\delta \prec V_\delta$, with $\kappa = \mathrm{crit}(j)$. Then there exists a normal ultrafilter on κ concentrating on cardinals that are n-huge for every n.*

Proof. As above, let $\kappa_n = j^n(\kappa)$. Let $U_n = \{X \subseteq \mathcal{P}(\kappa_n); j``\kappa_n \in j(X)\}$. Then U_n is a κ-complete ultrafilter U_n on $\mathcal{P}(\kappa_n)$, and, for every $i < n$, the set $\{x \in \mathcal{P}(\kappa_n); ot(x \cap \kappa_{i+1}) = \kappa_i\}$ belongs to U_n, since its image under j is $\{x \in \mathcal{P}(\kappa_{n+1}); ot(x \cap \kappa_{i+2}) = \kappa_{i+1}\}$, which contains $j``\kappa_n$ as we have $j``\kappa_n \cap \kappa_{i+2} = j``\kappa_{i+1}$. By [14], Theorem 24.8, this means that κ is n-huge.

Then we use a classical reflection argument, especially easy here. Let $U = \{X \subseteq \kappa; \kappa \in j(X)\}$. Then U is a normal ultrafilter over κ. Let X_0 be the set of all cardinals below κ that are n-huge for every n. Then $j(X_0)$ is the set of all cardinals below $j(\kappa)$ that are n-huge for every n, which contains κ as was seen above. So X_0 belongs to U. $\dashv$

1.2. Operations on elementary embeddings

For λ a limit ordinal, we denote by $\mathcal{E}_\lambda$ the family of all $j : V_\lambda \prec V_\lambda$. In most cases, $\mathcal{E}_\lambda$ is empty, while Axiom (I3) precisely states that at least one set $\mathcal{E}_\lambda$ is nonempty.

For λ a limit ordinal, it is not true that a function $f : V_\lambda \to V_\lambda$ is an element of V_λ. However, we can approximate f by its restrictions $f{\restriction}V_\gamma$ with $\gamma < \lambda$, each of which belongs to V_λ. If g is (another) function defined on V_λ, then g can be applied to each restriction $f{\restriction}V_\gamma$. If g happens to be an elementary embedding, the images $g(f{\restriction}V_\gamma)$ form a coherent system, and, in this way, we can *apply* g to f.

1.5 Definition. For $j, k : V_\lambda \to V_\lambda$, the *application of j to k* is defined by

$$j[k] = \bigcup_{\gamma<\lambda} j(k{\restriction}V_\gamma).$$

This definition makes sense, as, by construction, $k{\restriction}V_\gamma$ belongs to $V_{k(\gamma)+3}$, and therefore to V_λ.

1.6 Lemma. *Assume $j, k \in \mathcal{E}_\lambda$. Then $j[k]$ belongs to $\mathcal{E}_\lambda$, and we have* $\mathrm{crit}(j[k]) = j(\mathrm{crit}(k))$.

Proof. When γ ranges over λ, the various mappings $k{\restriction}V_\gamma$ are compatible. As j is elementary, $j(k{\restriction}V_\gamma)$ is a partial mapping defined on $V_{j(\gamma)}$, and the partial mappings $j(k{\restriction}V_\gamma)$ and $j(k{\restriction}V_{\gamma'})$ associated with different ordinals γ, γ' agree on $V_{j(\gamma)} \cap V_{j(\gamma')}$. Hence $j[k]$ is a mapping of V_λ into itself.

Let $\Phi(\vec{x})$ be a first-order formula. For each γ in λ, we have

$$(\forall \vec{x} \in V_\gamma)(\Phi(\vec{x}) \Leftrightarrow \Phi((k{\restriction}V_\gamma)(\vec{x}))),$$

hence, applying j,

$$(\forall \vec{x} \in V_{j(\gamma)})(\Phi(\vec{x}) \Leftrightarrow \Phi(j(k{\restriction}V_\gamma)(\vec{x}))),$$

so $j[k]$ is an elementary embedding of V_λ into itself.

The equality $\mathrm{crit}(j[k]) = j(\mathrm{crit}(k))$ follows from the fact that $k(\mathrm{crit}(k)) > \mathrm{crit}(k)$ implies $j[k](j(\mathrm{crit}(k)) > j(\mathrm{crit}(k))$, while $(\forall\gamma < \mathrm{crit}(k))(k(\gamma) = \gamma)$ implies $(\forall \gamma < j(\mathrm{crit}(k)))(j[k](\gamma) = \gamma)$. ⊣

Notice that, for j, k in $\mathcal{E}_\lambda$ and $\gamma < \lambda$, the equality

$$j[k]{\restriction}V_{j(\gamma)} = j(k{\restriction}V_\gamma) \tag{I.1}$$

is true by construction, as well as the formula

$$j[k](x) = jkj^{-1}(x) \tag{I.2}$$

whenever x belongs to the image of j.

Besides the application operation, composition is another binary operation on $\mathcal{E}_\lambda$. Let us emphasize that application is *not* composition. As should be clear from (I.2), application can be viewed as a sort of conjugation with respect to composition.

Let us turn to the algebraic study of the application and composition operations. The former is neither commutative nor associative. The operations satisfy the following identities.

1.7 Lemma (folklore). *For $j, k, \ell \in \mathcal{E}_\lambda \cup \{\mathrm{id}_{V_\lambda}\}$, we have*

$$j[k[\ell]] = j[k][j[\ell]],\ j{\circ}k = j[k]{\circ}j,\ (j{\circ}k)[\ell] = j[k[\ell]],\ j[k{\circ}\ell] = j[k]{\circ}j[\ell]. \tag{I.3}$$

Proof. Let $\gamma < \lambda$. Then $\ell{\restriction}V_\gamma$ belongs to V_β for some $\beta < \lambda$. From the definition, we have $k[\ell]{\restriction}V_{k(\gamma)} = (k{\restriction}V_\beta)(\ell{\restriction}V_\gamma)$. Applying j we get

$$j(k[\ell]{\restriction}V_{k(\gamma)}) = j(k{\restriction}V_\beta)[j(\ell{\restriction}V_\gamma)].$$

By (I.1), the left factor is $j[k[\ell]]{\restriction}V_{j(k(\gamma))}$, and $j(k(\gamma)) = j[k](j(\gamma))$ implies that the right factor is $j[k][j[\ell]]{\restriction}V_{j(k(\gamma))}$. As γ is arbitrary, we deduce $j[k[\ell]] = j[k][j[\ell]]$.

Let $x \in V_\lambda$. For γ sufficiently large, we have $x \in \mathrm{Dom}(k{\restriction}V_\gamma)$, hence

$$j(k(x)) = j((k{\restriction}V_\gamma)(x)) \ = j(k{\restriction}V_\gamma)(j(x)) = j[k](j(x)),$$

which establishes the equality $j{\circ}k = j[k]{\circ}j$. Applying the latter to $x = \ell{\restriction}V_\gamma$, one easily deduces $(j{\circ}k)[\ell] = j[k[\ell]]$.

Finally, using the fact that j preserves composition, we obtain

$$\begin{aligned} j[k{\circ}\ell]{\restriction}V_{j(\gamma)} &= j((k{\circ}\ell){\restriction}V_\gamma) = j((k{\restriction}V_{\ell(\gamma)}){\circ}(\ell{\restriction}V_\gamma)) \\ &= (j[k]{\restriction}V_{j\ell(\gamma)}){\circ}(j[\ell]{\restriction}V_{j(\gamma)}) = (j[k]{\circ}j[\ell]){\restriction}V_{j(\gamma)}, \end{aligned}$$

for every γ, and hence $j[k{\circ}\ell] = j[k]{\circ}j[\ell]$. ⊣

Also $j[\mathrm{id}_{V_\lambda}] = \mathrm{id}_{V_\lambda}$ and $\mathrm{id}_{V_\lambda}[j] = j$ hold for every j in $\mathcal{E}_\lambda \cup \{\mathrm{id}_{V_\lambda}\}$. In order to fix the vocabulary for the sequel, we put the following definitions:

1.8 Definition. (i) We say that $(S, *)$ is a *left self-distributive system*, or *LD-system*, if $*$ is a binary operation on S satisfying

$$x{*}(y{*}z) = (x{*}y){*}(x{*}z). \tag{LD}$$

(ii) We say that $(M, *, \cdot, 1)$ is a *left self-distributive monoid*, or *LD-monoid*, if $(M, \cdot, 1)$ is a monoid and $*$ is a binary operation on M satisfying

$$x{\cdot}y = (x{*}y){\cdot}x, \ \ (x{\cdot}y){*}z = x{*}(y{*}z), \ \ x{*}(y{\cdot}z) = (x{*}y){\cdot}(x{*}z), \ \ x{*}1 = 1. \tag{I.4}$$

Observe that an LD-monoid is an LD-system and $1{*}x = x$ always holds, as (I.4) implies $x{*}(y{*}z) = (x{\cdot}y){*}z = ((x{*}y){\cdot}x){*}z = (x{*}y){*}(x{*}z)$, and, similarly, $1{*}x = (1{*}x){\cdot}1 = 1{\cdot}x = x$. With these definitions (various other names have been used in literature), we can restate 1.7 as

1.9 Proposition. *Let λ be a limit ordinal. Then $\mathcal{E}_\lambda$ equipped with application is an LD-system, and $\mathcal{E}_\lambda \cup \{\mathrm{id}_{V_\lambda}\}$ equipped with application and composition is an LD-monoid.*

Before developing our study further, let us conclude this section with an independent result which we shall see in Section 3 leads to interesting consequences.

1.10 Proposition. *Assume $j : V_\lambda \prec V_\lambda$. Then we have $j[j](\alpha) \leqslant j(\alpha)$ for every ordinal $\alpha < \lambda$.*

Proof. Let β satisfy $j(\beta) > \alpha$ and $(\forall \xi < \beta)(j(\xi) \leqslant \alpha)$. As j is elementary, we deduce $j[j](j(\beta)) > j(\alpha)$ and $(\forall \xi < j(\beta))(j[j](\xi) \leqslant j(\alpha))$—we can make things rigorous by replacing the parameter j with some approximation of the form $j{\restriction}V_\gamma$ with γ sufficiently large. As $\alpha < j(\beta)$ holds, we can take $\xi = \alpha$ in the second formula, which gives $j[j](\alpha) \leqslant j(\alpha)$. ⊣

1.3. Iterations of an elementary embedding

We shall now turn to the specific study of the iterations of a fixed elementary embedding $j : V_\lambda \prec V_\lambda$, as developed by R. Laver. This means that we concentrate on the countable subfamily of $\mathcal{E}_\lambda$ consisting of those embeddings that can be obtained from j using application (or both application and composition).

1.11 Definition. For $j \in \mathcal{E}_\lambda$, $\mathrm{Iter}(j)$ denotes the sub-LD-system of $\mathcal{E}_\lambda$ generated by j, while $\mathrm{Iter}^*(j)$ denotes the sub-LD-monoid of $\mathcal{E}_\lambda \cup \{\mathrm{id}_{V_\lambda}\}$ generated by j. The elements of $\mathrm{Iter}^*(j)$ are called the *iterates* of j, while the elements of $\mathrm{Iter}(j)$ are called the *pure iterates* of j.

By definition, the pure iterates of j are those elementary embeddings that can be obtained from j using the application operation repeatedly, so they comprise $j, j[j], j[j[j]], j[j][j]$, etc. As application is a non-associative operation, the iterates of j do not reduce to powers of j; however, even the notion of a power has to be made precise. We shall use the following notation:

1.12 Definition. For j in $\mathcal{E}_\lambda$—or, more generally, in any binary system—we recursively define the nth *right power* $j^{[n]}$ of j and the nth *left power* $j_{[n]}$ of j by $j^{[1]} = j_{[1]} = j$, $j^{[n+1]} = j[j^{[n]}]$, and $j_{[n+1]} = j_{[n]}[j]$.

For future use, let us mention some relations between the powers in an arbitrary LD-system:

1.13 Lemma. *The following identities are satisfied in every LD-system*

$$x^{[p+1]} = x^{[q]}[x^{[p]}] \text{ for } 1 \leqslant q \leqslant p, \quad (x^{[p]})^{[q]} = x^{[p+q-1]} \text{ for } 1 \leqslant p, q. \tag{I.5}$$

The sequel of the study aims at determining some possible quotients of the algebraic structures $\mathrm{Iter}(j)$ and $\mathrm{Iter}^*(j)$, *i.e.*, to look for equivalence relations that are compatible with the involved algebraic operation(s). A simple idea could be to concentrate on the critical ordinals, or, more generally, on the values at particular fixed ordinals, but this naive approach is not relevant beyond the first levels. Another natural idea would be to consider the restrictions of the embeddings to a fixed rank, *i.e.*, to consider equivalence relations of the form $j{\restriction}V_\gamma = j'{\restriction}V_\gamma$, but such relations are not compatible with the application operation in general, and we are led to the following slightly different relations.

1.14 Definition. (Laver) Assume $j, j' \in \mathcal{E}_\lambda \cup \{\mathrm{id}_{V_\lambda}\}$. For γ limit below λ, we say that j and j' are *γ-equivalent*, denoted $j \overset{\gamma}{\equiv} j'$, if, for every x in V_γ, we have $j(x) \cap V_\gamma = j'(x) \cap V_\gamma$.

By definition, $\stackrel{\gamma}{\equiv}$ is an equivalence relation on $\mathcal{E}_\lambda \cup \{\mathrm{id}_{V_\lambda}\}$. Note that $j \stackrel{\gamma}{\equiv} j'$ implies $j(x) \cap V_\gamma = j'(x) \cap V_\gamma$ for every x in V_λ—not only in V_γ—since, for $y \in V_\beta$ with $\beta < \gamma$, the relation $y \in j(x) \cap V_\gamma$ is equivalent to $y \in j(x \cap V_\beta) \cap V_\gamma$, and $x \cap V_\beta$ belongs to $V_{\beta+1}$, hence to V_γ since γ is limit.

Let us begin with easy observations.

1.15 Lemma. *Assume $j \stackrel{\gamma}{\equiv} j'$ and $\alpha < \gamma$. Then we have either $j(\alpha) < \gamma$, whence $j'(\alpha) = j(\alpha)$, or $j(\alpha) \geqslant \gamma$, whence $j'(\alpha) \geqslant \gamma$. So, in particular, we have either* $\mathrm{crit}(j) = \mathrm{crit}(j') < \gamma$, *or both* $\mathrm{crit}(j) \geqslant \gamma$ *and* $\mathrm{crit}(j') \geqslant \gamma$.

Proof. Assume $j' \stackrel{\gamma}{\equiv} j$ and $\alpha, \beta < \gamma$. Then, by definition, $j(\alpha) > \beta$ is equivalent to $j'(\alpha) > \beta$. ⊣

1.16 Lemma. *Assume $j, k \in \mathcal{E}_\lambda$. Then $j[k]$ and k are* $\mathrm{crit}(j)$*-equivalent.*

Proof. Let $\gamma = \mathrm{crit}(j)$. An induction on the rank shows that $j{\restriction}V_\gamma$ is the identity mapping. Then $y \in k(x)$ is equivalent to $j(y) \in j[k](j(x))$, hence to $y \in j[k](x)$ for x, y in V_γ. ⊣

1.17 Proposition. *For limit $\gamma < \lambda$, γ-equivalence is compatible with composition.*

Proof. Assume $j \stackrel{\gamma}{\equiv} j'$ and $k \stackrel{\gamma}{\equiv} k'$. Let $x, y \in V_\gamma$, and $y \in j(k(x))$. As γ is limit, we have $x, y \in V_\beta$ for some $\beta < \gamma$, so $y \in j(k(x))$ implies $y \in j(k(x) \cap V_\beta) \cap V_\gamma$. By hypothesis, we have $k(x) \cap V_\beta = k'(x) \cap V_\beta \in V_{\beta+1} \subseteq V_\gamma$, hence

$$j'(k'(x) \cap V_\beta) \cap V_\gamma = j(k'(x) \cap V_\beta) \cap V_\gamma = j(k(x) \cap V_\beta) \cap V_\gamma.$$

We deduce $y \in j'(k'(x))$, hence $j(k(x)) \cap V_\gamma \subseteq j'(k'(x)) \cap V_\gamma$. By symmetry, we obtain $j(k(x)) \cap V_\gamma = j'(k'(x)) \cap V_\gamma$, so $j \circ k$ and $j' \circ k'$ are γ-equivalent. ⊣

1.18 Lemma. *Let $j : V_\lambda \prec V_\lambda$. Then, for each γ satisfying* $\mathrm{crit}(j) < \gamma < \lambda$, *there exists δ satisfying $\delta < \gamma \leqslant j(\delta) < j(\gamma)$.*

Proof. Let $\kappa = \mathrm{crit}(j)$. Let δ be the least ordinal satisfying $\gamma \leqslant j(\delta)$: since $\gamma \leqslant j(\gamma)$ is always true, δ exists, and we have $\delta \leqslant \gamma$. Assume $\delta = \gamma$. This means that $\xi < \gamma$ implies $j(\xi) < \gamma$, hence $j^n(\xi) < \gamma$ for each n. This contradicts $\gamma < \lambda$ and (remark after 1.3) $\lambda = \sup_n j^n(\kappa)$. ⊣

1.19 Proposition. *Assume $j \stackrel{\gamma}{\equiv} j'$ and $k \stackrel{\delta}{\equiv} k'$ with $j(\delta) \geqslant \gamma$. Then we have $j[k] \stackrel{\gamma}{\equiv} j'[k']$.*

Proof. Assume first $\mathrm{crit}(j) \geqslant \gamma$. By 1.15, we have also $\mathrm{crit}(j') \geqslant \gamma$. Moreover, $\delta \geqslant \gamma$ holds, for $\delta < \gamma$ would imply $j(\delta) = \delta < \gamma$. Hence, $k \stackrel{\delta}{\equiv} k'$ implies $k \stackrel{\gamma}{\equiv} k'$. Then, by 1.16, we find $j[k] \stackrel{\gamma}{\equiv} k \stackrel{\gamma}{\equiv} k' \stackrel{\gamma}{\equiv} j'[k']$.

Assume now $\mathrm{crit}(j) < \gamma$, and, therefore, $\mathrm{crit}(j') = \mathrm{crit}(j)$. Since $k \overset{\delta}{\equiv} k'$ implies $k \overset{\delta'}{\equiv} k'$ for $\delta' \leqslant \delta$, we may assume without loss of generality that δ is minimal satisfying $j(\delta) \geqslant \gamma$, which, by 1.18, implies $\gamma > \delta$. Let $j \overset{*}{\cap} V_\alpha$ denote the set $\{(x,y) \in V_\alpha^2 \ ; \ y \in j(x)\}$. By definition, $j \overset{\alpha}{\equiv} j'$ is equivalent to $j \overset{*}{\cap} V_\alpha = j' \overset{*}{\cap} V_\alpha$. We have

$$j[k] \overset{*}{\cap} V_\gamma = (j[k] \overset{*}{\cap} V_{j(\delta)}) \cap V_\gamma^2 = j(k \overset{*}{\cap} V_\delta) \cap V_\gamma^2.$$

By construction, $k \overset{*}{\cap} V_\delta$ is a set of ordered pairs of elements of V_δ, hence an element of V_γ. The hypotheses $k \overset{*}{\cap} V_\delta = k' \overset{*}{\cap} V_\delta$ and $j(x) \cap V_\gamma = j'(x) \cap V_\gamma$ for $x \in V_\gamma$ imply

$$j[k] \overset{*}{\cap} V_\gamma = j(k \overset{*}{\cap} V_\delta) \cap V_\gamma = j'(k \overset{*}{\cap} V_\delta) \cap V_\gamma = j'(k' \overset{*}{\cap} V_\delta) \cap V_\gamma = j'[k'] \overset{*}{\cap} V_\gamma,$$

so $j[k]$ and $j'[k']$ are γ-equivalent. ⊣

Let $j, k, \ell \in \mathcal{E}_\lambda$. Left self-distributivity gives $j[k[\ell]] = j[k][j[\ell]]$, but these embeddings need not be equal to $j[k][\ell]$, unless $j[\ell] = \ell$ holds. Now, by 1.16, $j[\ell]$ and ℓ are $\mathrm{crit}(j)$-equivalent, which implies that $j[k[\ell]]$ and $j[k][\ell]$ are $j[k](\mathrm{crit}(j))$-equivalent. Generalizing the argument, we obtain the following technical lemma. The convention is that $j[k][\ldots]$ means $(j[k])[\ldots]$.

1.20 Lemma. *Assume $j, j_1, \ldots, j_p \in \mathcal{E}_\lambda$, and let $\gamma = \mathrm{crit}(j)$.*
(i) *Assume $j[j_1[j_2]\ldots[j_\ell]](\gamma) \geqslant \gamma'$ for $1 \leqslant \ell \leqslant p-1$. Then we have*

$$j[j_1][j_2]\ldots[j_p] \overset{\gamma'}{\equiv} j[j_1[j_2]\ldots[j_p]]. \tag{I.6}$$

(ii) *Assume $\mathrm{crit}(j_1[j_2]\ldots[j_\ell]) < \gamma$ for $1 \leqslant \ell \leqslant p-1$ and $\mathrm{crit}(j_1[j_2]\ldots[j_p]) \leqslant \gamma$. Then we have*

$$\mathrm{crit}(j[j_1][j_2]\ldots[j_p]) = j(\mathrm{crit}(j_1[j_2]\ldots[j_p])). \tag{I.7}$$

Proof. (i) Use induction on p. For $p = 1$, (I.6) is an equality. Otherwise, we have, by induction hypothesis, $j[j_1][j_2]\ldots[j_{p-1}] \overset{\gamma'}{\equiv} j[j_1[j_2]\ldots[j_{p-1}]]$, and, therefore,

$$j[j_1][j_2]\ldots[j_{p-1}][j_p] \overset{\gamma'}{\equiv} j[j_1[j_2]\ldots[j_{p-1}]][j_p]. \tag{I.8}$$

Lemma 1.16 gives $j_p \overset{\gamma}{\equiv} j[j_p]$, which implies

$$j[j_1[j_2]\ldots[j_{p-1}]][j_p] \overset{\gamma'}{\equiv} j[j_1[j_2]\ldots[j_{p-1}]][j[j_p]], \tag{I.9}$$

since $j[j_1[j_2]\ldots[j_{p-1}]](\gamma) \geqslant \gamma'$ holds by hypothesis. The right factor of (I.9) is also $j[j_1[j_2]\ldots[j_p]]$ by left self-distributivity, and combining (I.8) and (I.9) gives (I.6).

(ii) The case $p = 1$ is trivial. Assume $p \geqslant 2$, and let γ' be the smallest of $j[j_1](\gamma)$, $j[j_1[j_2]](\gamma)$, $\dots$, $j[j_1[j_2]\dots[j_{p-1}]](\gamma)$. Applying (i), we find

$$j[j_1][j_2]\dots[j_p] \overset{\gamma'}{\equiv} j[j_1[j_2]\dots[j_p]]. \tag{I.10}$$

Let q be minimal satisfying $\gamma' = j[j_1[j_2]\dots[j_q]](\gamma)$, and $j' = j_1[j_2]\dots[j_q]$. Then we have $\gamma' = j[j'](\gamma)$. By hypothesis, we have $\mathrm{crit}(j') < \gamma$, so there exists δ satisfying $\delta < \gamma \leqslant j'(\delta)$. From (I.10) we deduce

$$j(\gamma) \leqslant j(j'(\delta)) = j[j'](j(\delta)) = j[j'](\delta) < j[j'](\gamma) = \gamma'.$$

Hence $\mathrm{crit}(j_1[j_2]\dots[j_p]) \leqslant \gamma$ implies $\mathrm{crit}(j[j_1[j_2]\dots[j_p]]) \leqslant j(\gamma) < \gamma'$. Therefore the right embedding in (I.10) has its critical ordinal below γ', and, by 1.15, so has the left-hand embedding, and the two critical ordinals are equal. ⊣

1.4. Finite quotients

By 1.19, γ-equivalence is compatible with the application operation, so quotienting under $\overset{\gamma}{\equiv}$ leads to a well-defined LD-system. We shall describe this quotient LD-system completely when γ happens to be the critical ordinal of some iteration of the embedding j we are studying.

By construction, for $j : V_\lambda \prec V_\lambda$, the sets $\mathrm{Iter}(j)$ and $\mathrm{Iter}^*(j)$ consist of countably many elementary embeddings, each of which except the identity has a critical ordinal. So, we can associate with j the countable family of all critical ordinals of iterates of j.

1.21 Definition. The ordinal $\mathrm{crit}_n(j)$ is defined to be the $(n{+}1)$th element in the increasing enumeration of the critical ordinals of iterations of j.

The formulas $\mathrm{crit}(j[k]) = j(\mathrm{crit}(k))$, $\mathrm{crit}(j{\circ}k) = \inf(\mathrm{crit}(j), \mathrm{crit}(k))$ and an obvious induction show that $\mathrm{crit}(i) \geqslant \mathrm{crit}(j)$ holds for every iterate i of j. Hence $\mathrm{crit}_0(j)$ is always $\mathrm{crit}(j)$. We shall prove below the values $\mathrm{crit}_1(j) = j(\mathrm{crit}(j))$ and $\mathrm{crit}_2(j) = j^2(\mathrm{crit}(j))$. Things become complicated subsequently. At this point, we do not know (yet) that the sequence of the ordinals $\mathrm{crit}_n(j)$ exhaust all critical ordinals in $\mathrm{Iter}^*(j)$: it could happen that some nontrivial iterate i of j has its critical ordinal beyond all $\mathrm{crit}_n(j)$'s.

1.22 Theorem (Laver). *Assume $j : V_\lambda \prec V_\lambda$. Then $\mathrm{crit}_n(j)$-equivalence is a congruence on the LD-monoid $\mathrm{Iter}^*(j)$, and the quotient LD-monoid has 2^n elements, namely the classes of j, $j_{[2]}$, $\dots$, $j_{[2^n]}$, the latter also being the class of the identity.*

The proof requires several preliminary results.

1.23 Lemma. *Assume that i_1, i_2, $\dots$, i_{2^n} are iterates of j. Then we have $\mathrm{crit}(i_1[i_2]\dots[i_p]) \geqslant \mathrm{crit}_n(j)$ for some p with $p \leqslant 2^n$.*

Proof. We use induction on n. For $n = 0$, the result is the inequality $\mathrm{crit}(i_1) \geqslant \mathrm{crit}(j)$, which we have seen holds for every iterate i_1 of j. Otherwise, we apply the induction hypothesis twice. First, we find $q \leqslant 2^{n-1}$ satisfying

$$\mathrm{crit}(i_1[i_2]\dots[i_q]) \geqslant \mathrm{crit}_{n-1}(j). \tag{I.11}$$

If the inequality is strict, we have $\mathrm{crit}(i_1[i_2]\dots[i_q]) \geqslant \mathrm{crit}_n(j)$, and we are done. So, we can assume from now on that (I.11) is an equality. By applying the induction hypothesis again, we find $r \leqslant 2^{n-1}$ satisfying

$$\mathrm{crit}(i_{q+1}[i_{q+2}]\dots[i_{q+r}]) \geqslant \mathrm{crit}_{n-1}(j).$$

If r is chosen to be minimal, we can apply 1.20(i) with $p = r$, $j = i_1[i_2]\dots[i_q]$, $j_1 = i_{q+1}, \dots, j_p = i_{q+r}$, $\gamma = \mathrm{crit}_{n-1}(j)$, and $\gamma' = \mathrm{crit}_n(j)$. Indeed, with these notations, we have $\mathrm{crit}(j_1[j_2]\dots[j_s]) < \gamma$ for $s < r$, hence

$$\mathrm{crit}(j[j_1[j_2]\dots[j_s]]) = j(\mathrm{crit}(j_1[j_2]\dots[j_s])) = \mathrm{crit}(j_1[j_2]\dots[j_s]) < \gamma,$$

and, therefore, $j[j_1[j_2]\dots[j_s]](\gamma) > \gamma$, which gives $j[j_1[j_2]\dots[j_s]](\gamma) \geqslant \gamma'$ by definition. So we have

$$j[j_1][j_2]\dots[j_p] \overset{\gamma'}{\equiv} j[j_1[j_2]\dots[j_p]].$$

We have $\mathrm{crit}(j[j_1[j_2]\dots[j_p]]) \geqslant j(\gamma) \geqslant \gamma' = \mathrm{crit}_n(j)$, so, by 1.15, we deduce

$$\mathrm{crit}(j[j_1][j_2]\dots[j_p]) \geqslant \mathrm{crit}_n(j)$$

i.e., $\mathrm{crit}(i_1[i_2]\dots[i_q]\dots[i_{q+r}]) \geqslant \mathrm{crit}_n(j)$, as was expected. ⊣

The main task is now to show that all iterates of j can be approximated by left powers of j up to $\mathrm{crit}_n(j)$-equivalence. We begin with approximating arbitrary iterates by pure iterates.

1.24 Lemma. *Assume that n is a fixed integer, and i is an iterate of j. Then there exists a pure iterate i' of j that is $\mathrm{crit}_n(j)$-equivalent to i.*

Proof. Let $\gamma = \mathrm{crit}_n(j)$, and let A be the set of those iterates of j that are γ-equivalent to some pure iterate of j. The set A contains j, and it is obviously closed under application. So, in order to show that A is all of $\mathrm{Iter}^*(j)$, it suffices to show that A is closed under composition, and, because γ-equivalence is compatible with composition, it suffices to show that, if i_1, i_2 are pure iterates of j, then some pure iterate of j is γ-equivalent to $i_2 \circ i_1$. To this end, we define recursively a sequence of pure iterates of j, say $i_3, i_4, \dots$ by the recursive clause $i_{p+2} = i_{p+1}[i_p]$. Then we have

$$i_3 \circ i_2 = i_2[i_1] \circ i_2 = i_2 \circ i_1,$$

and, recursively, $i_{p+1}{\circ}i_p = i_2{\circ}i_1$ for every p. We claim that $\mathrm{crit}(i_p) \geqslant \gamma$ holds for at least one of the values $p = 2n$ or $p = 2n+1$. If this is known, we find $i_2{\circ}i_1 = i_p{\circ}i_{p-1} = i_p[i_{p-1}]{\circ}i_p \overset{\gamma}{\equiv} i_{p-1}$, and we are done.

In order to prove the claim, we separate the cases $\mathrm{crit}(i_2) > \mathrm{crit}(i_1)$ and $\mathrm{crit}(i_2) \leqslant \mathrm{crit}(i_1)$. In this first case, we have

$$\mathrm{crit}(i_3) = i_2(\mathrm{crit}(i_1)) = \mathrm{crit}(i_1), \quad \text{and} \quad \mathrm{crit}(i_4) = i_3(\mathrm{crit}(i_2)) > \mathrm{crit}(i_2).$$

An immediate induction gives

$$\mathrm{crit}(i_1) = \mathrm{crit}(i_3) = \mathrm{crit}(i_5) = \ldots, \ \mathrm{crit}(i_2) < \mathrm{crit}(i_4) < \mathrm{crit}(i_6) < \ldots.$$

By definition, we have $\mathrm{crit}(i_1) \geqslant \mathrm{crit}_0(j)$, and, therefore, $\mathrm{crit}(i_2) \geqslant \mathrm{crit}_1(j)$, and, inductively, $\mathrm{crit}(i_{2n}) \geqslant \gamma$, as was claimed.

Assume now $\mathrm{crit}(i_2) \leqslant \mathrm{crit}(i_1)$. Similar computations give

$$\mathrm{crit}(i_1) < \mathrm{crit}(i_3) < \mathrm{crit}(i_5) < \ldots, \ \mathrm{crit}(i_2) = \mathrm{crit}(i_4) = \mathrm{crit}(i_6) = \ldots,$$

and we find now $\mathrm{crit}(i_{2n+1}) \geqslant \gamma$. So the claim is established, and the proof is complete. ⊣

Let us e.g. consider $i = j{\circ}j$. We are in the case "$\mathrm{crit}(i_2) \leqslant \mathrm{crit}(i_1)$", and we know that the pure iterate i_{2n+1} as above is a $\mathrm{crit}_n(j)$-approximation of i. For instance, we find $i_3 = j_{[2]}$, $i_4 = j_{[3]}$, $i_5 = j_{[3]}[j_{[2]}] = {j_{[4]}}^{[2]}$, so $j{\circ}j$ and $(j_{[4]})^{[2]}$ are $\mathrm{crit}_2(j)$-equivalent. It can be seen that the critical ordinal of $(j_{[4]})^{[2]}$, *i.e.*, $j_{[4]}(\mathrm{crit}(j_{[4]}))$, is larger than $\mathrm{crit}_2(j)$, namely it is $\mathrm{crit}_3(j)$, so the previous equivalence is actually a $\mathrm{crit}_3(j)$-equivalence.

1.25 Proposition. *Assume $j : V_\lambda \prec V_\lambda$, $i \in \mathrm{Iter}^*(j)$, and $n \geqslant 0$. Then i is $\mathrm{crit}_n(j)$-equivalent to $j_{[p]}$ for some p with $p \leqslant 2^n$.*

Proof. By 1.24, we may assume that i is a pure iterate of j. The principle is to iteratively divide by j on the right, *i.e.*, we construct pure iterates of j, say i_0, i_1, ... such that i_0 is i, and i_p is $\mathrm{crit}_n(j)$-equivalent to $i_{p+1}[j]$ for every p. So, i is $\mathrm{crit}_n(j)$-equivalent to $i_p[j]\ldots[j]$ (p times j) for every p. We stop the process when we have either $i_p = j$, in which case i is $\mathrm{crit}_n(j)$-equivalent to $j_{[p+1]}$, or $p = 2^n$: in this case, we have obtained a sequence of 2^n iterates of j, and 1.23 completes the proof.

Let us go into details. In order to see that the construction is possible, let us assume that i_p has been obtained. If $i_p = j$ holds, we are done. Otherwise, i_p has the form $i'_1[i'_2[\ldots[i'_r[j]]\ldots]]$, where $i'_1, \ldots, i'_r$ are some uniquely defined pure iterates of j. Applying the identity $j[k[\ell]] = (j{\circ}k)[\ell]$ $r-2$ times, we find $i_p = (i'_1{\circ}\ldots{\circ}i'_r)[j]$, and we define i_{p+1} to be a pure iterate of j that is $\mathrm{crit}_n(j)$-equivalent to $i'_1{\circ}\ldots{\circ}i'_r$.

Assume that the construction continues for at least 2^n steps, and let us consider the 2^n embeddings $i_{2^n}[j]$, $i_{2^n}[j][j]$, ..., $i_{2^n}[j][j]\ldots[j]$ (2^n times j).

By 1.23, there must exist $p \leqslant 2^n$ such that the critical ordinal of $i_{2^n}[j][j]\ldots[j]$ (p times j) is at least $\mathrm{crit}_n(j)$. Let i' be the latter elementary embedding. Then i is $\mathrm{crit}_n(j)$-equivalent to $i_{2^n}[j][j]\ldots[j]$ (2^n times j), which is also $i'[j][j]\ldots[j]$ ($2^n - p$ times j), and, therefore, i is $\mathrm{crit}_n(j)$-equivalent to $\mathrm{id}[j][j]\ldots[j]$ ($2^n - p$ times j), *i.e.*, to $j_{[2^n-p]}$, and we are done as well. ⊣

The previous argument is effective. Starting with an arbitrary iteration i of j and a fixed level of approximation $\mathrm{crit}_n(j)$, we can find some left power of j that is $\mathrm{crit}_n[j]$-equivalent to i in a finite number of steps. However, the computation becomes quickly very intricate, and there is no uniform way to know how many steps are needed. For instance, let $i = j^{[3]}$, the simplest iterate of j that is not a left power. We write $i = (j{\circ}j)[j]$, and have to find an approximation of $j{\circ}j$. Now, $j{\circ}j$ is $\mathrm{crit}_3(j)$-equivalent to $j_{[3]}$, and, in this particular case, we obtain directly that $j^{[3]}$ is $\mathrm{crit}_3(j)$-equivalent to $j_{[3]}[j]$, *i.e.*, to $j_{[4]}$. If we look for $\mathrm{crit}_4(j)$-equivalence, the computation is much more complicated. The results below will show that, if i is $\mathrm{crit}_3(j)$-equivalent to $j_{[4]}$, then it is $\mathrm{crit}_4(j)$-equivalent either to $j_{[4]}$ or to $j_{[12]}$. By determining the critical ordinal of $i[j][j][j][j]$, we could find that the final result is $j^{[3]}$ being $\mathrm{crit}_4(j)$-equivalent to $j_{[12]}$. We shall see an easier alternative way for proving such statements in Subsection 3.2 below.

1.26 Proposition. *Assume that j is a nontrivial elementary embedding of a limit rank into itself. Then, for every p, we have* $\mathrm{crit}(j_{[p]}) = \mathrm{crit}_m(j)$, *where m is the largest integer such that 2^m divides p.*

Proof. We establish inductively on $n \geqslant 0$ that $\mathrm{crit}(j_{[2^n]}) \geqslant \mathrm{crit}_n(j)$ holds, and that $\mathrm{crit}(j_{[p]}) = \mathrm{crit}(j_{[2^m]})$ holds for $p < 2^n$ with m the largest integer such that 2^m divides p. For $n = 0$, we already know $\mathrm{crit}(j) = \mathrm{crit}_0(j)$. Otherwise, let us consider the embeddings $j_{[2^n+p]}$ for $1 \leqslant p \leqslant 2^n$. By definition, we have $j_{[2^n+p]} = j_{[2^n]}[j]\ldots[j]$ (p times j), and, by induction hypothesis, we have $\mathrm{crit}(j_{[s]}) < \mathrm{crit}(j_{[2^n]})$ for $s < 2^n$, and $\mathrm{crit}(j_{[2^n]}) \geqslant \mathrm{crit}_n(j)$. By 1.20(ii) applied with $j = j_{[2^n]}$ and $j_1 = \ldots = j_p = j$, we have

$$\mathrm{crit}(j_{[2^n+p]}) = j_{[2^n]}(\mathrm{crit}(j_{[p]})).$$

For $p < 2^n$, we deduce $\mathrm{crit}(j_{[2^n+p]}) = \mathrm{crit}(j_{[p]}) = \mathrm{crit}_m(j)$ where m is the largest integer such that 2^m divides p, which is also the largest integer such that 2^m divides $2^n + p$. For $p = 2^n$, we obtain

$$\mathrm{crit}(j_{[2^{n+1}]}) = j_{[2^n]}(\mathrm{crit}(j_{[2^n]}) > \mathrm{crit}(j_{[2^n]}) = \mathrm{crit}_n(j),$$

and we deduce $\mathrm{crit}(j_{[2^{n+1}]}) \geqslant \mathrm{crit}_{n+1}(j)$. So the induction is complete. Now, it follows from 1.25 that the critical ordinal of any iterate of j is either equal to the critical ordinal of some left power of j, or is larger than all ordinals $\mathrm{crit}_m(j)$. Since the sequence of all ordinals $\mathrm{crit}(j_{[2^n]})$ is increasing, the only possibility is $\mathrm{crit}(j_{[2^n]}) = \mathrm{crit}_n(j)$. ⊣

1.27 Lemma. *The left powers* $j_{[p]}$ *and* $j_{[p']}$ *are* $\mathrm{crit}_n(j)$*-equivalent if and only if* $p = p' \bmod 2^n$ *holds.*

Proof. We have $\mathrm{crit}(j_{[2^n]}) = \mathrm{crit}_n(j)$, so $j_{[2^n]}$ is $\mathrm{crit}_n(j)$-equivalent to the identity mapping, which, by Prop 1.19, inductively implies that $j_{[p]}$ and $j_{[2^n+p]}$ are $\mathrm{crit}_n(j)$-equivalent for every p. Hence the condition of the lemma is sufficient. On the other hand, we prove using induction on $n \geqslant 0$ that $1 \leqslant p < p' \leqslant 2^n$ implies that $j_{[p]}$ and $j_{[p']}$ are not $\mathrm{crit}_n(j)$-equivalent. The result is vacuously true for $n = 0$. Otherwise, for $p' \neq 2^{n-1} + p$, the induction hypothesis implies that $j_{[p]}$ and $j_{[p']}$ are not $\mathrm{crit}_{n-1}(j)$-equivalent, and *a fortiori* they are not $\mathrm{crit}_n(j)$-equivalent. Now, assume $p' = 2^{n-1} + p$ and $j_{[p]}$ and $j_{[p']}$ are $\mathrm{crit}_n(j)$-equivalent. By applying $2^{n-1} - p$ times 1.19, we deduce that $j_{[2^{n-1}]}$ and $j_{[2^n]}$ are $\mathrm{crit}_n(j)$-equivalent, which is impossible as we have $\mathrm{crit}(j_{[2^{n-1}]}) < \mathrm{crit}_n(j)$ and $\mathrm{crit}(j_{[2^n]}) \geqslant \mathrm{crit}_n(j)$. ⊣

We are now ready to complete the proof of 1.22.

Proof. The result is clear from 1.25 and 1.27. That $j_{[2^n]}$ and the identity mapping are $\mathrm{crit}_n(j)$-equivalent follows from $\mathrm{crit}_n(j)$ being the critical ordinal of $j_{[2^n]}$. ⊣

1.5. The Laver–Steel theorem

Assume $j : V_\lambda \prec V_\lambda$. By 1.6, $j^n(\mathrm{crit}(j))$ is the critical ordinal of $j^{[n+1]}$, which is also, by 1.13, $j^{[n]}[j^{[n]}]$: so, in the sequence of right powers $j, j^{[2]}, j^{[3]}, \dots$, every term is a left divisor of the next one. Kunen's bound asserts that the supremum of the critical ordinals in the previous sequence is λ. Actually, this property has nothing to do with the particular choice of the elementary embeddings $j^{[n]}$, and it is an instance of a much stronger statement, which is itself a special case of a general result of Steel about the Mitchell ordering [22]:

1.28 Theorem (Steel). *Assume that* $j_1, j_2, \dots$ *is a sequence in* $\mathcal{E}_\lambda$ *that is increasing with respect to divisibility, i.e., for every* n*, we have* $j_{n+1} = j_n[k_n]$ *for some* k_n *in* $\mathcal{E}_\lambda$*. Then we have* $\sup_n \mathrm{crit}(j_n) = \lambda$.

Here we shall give a simple proof of the considered specific statement, which is due to R. Dougherty.

1.29 Definition. Assume $j \in \mathcal{E}_\lambda$, and $\gamma < \lambda$. We say that the ordinal α is *γ-representable by j* if it can be expressed as $j(f)(x)$ where f and x belong to V_γ and f is a mapping with ordinal values; The set of all ordinals that are γ-representable by j is denoted $S_\gamma(j)$.

1.30 Lemma. *Assume* $j' = j[k]$ *in* $\mathcal{E}_\lambda$*, and let* γ *be an inaccessible cardinal satisfying* $\mathrm{crit}(j) < \gamma < \lambda$*. Then the order type of* $S_\gamma(j)$ *is larger than the order type of* $S_\gamma(j')$*.*

Proof. The point is to construct an increasing mapping of $S_\gamma(j')$ into some proper initial segment of $S_\gamma(j)$. The idea is that $S_\gamma(j')$ is (more or less) the image under j of some set $S_\delta(k)$ with $\delta < \gamma$, which we can expect to be smaller than $S_\gamma(j)$ because $\delta < \gamma$ holds and γ is inaccessible.

By 1.18, there exists an ordinal δ satisfying $\delta < \gamma \leqslant j(\delta)$. Let G be the function that maps every pair (f, x) in V_δ^2 such that f is a function with ordinal values and x lies in the domain of $k(f)$ to $k(f)(x)$. By construction, the image of G is the set $S_\delta(k)$. The cardinality of this set is at most that of V_δ^2, hence it is strictly less than γ since γ is inaccessible. So the order type of the set $S_\delta(k)$ is less than γ, and, by ordinal recursion, we construct an order-preserving mapping H of $S_\delta(k)$ onto some ordinal β below γ. Let us apply now j: the mapping $j(H)$ is also order-preserving, and it maps $j(S_\delta(k))$, which is $S_{j(\delta)}(j')$, onto $j(\beta)$. By hypothesis, $j(\delta) \geqslant \gamma$ holds, so $S_{j(\delta)}(j')$ includes $S_\gamma(j')$. Let α be an ordinal in the latter set: by definition, there exist f, x in V_γ, f a mapping with ordinal values, x an element in the domain of $j'(f)$, satisfying $\alpha = j'(f)(x)$, and we have

$$j(H)(\alpha) = j(H)(j'(f)(x)) = j(H)(j(G)((f,x))) = j(H{\circ}G)((f,x)). \quad \text{(I.12)}$$

Now H and G belong to V_γ, and therefore both $H{\circ}G$ and (f,x) are elements of V_γ. Thus (I.12) shows that the ordinal $j(H)(\alpha)$ is γ-representable by j, and the mapping $j(H)$ is an order-preserving mapping of $S_\gamma(j')$ into $S_\gamma(j)$. Moreover, the image of the mapping H is, by definition, the ordinal β, so the image of $j(H)$ is the ordinal $j(\beta)$, and, therefore, $j(H)$ is an order-preserving mapping of $S_\gamma(j')$ into $\{\xi \in S_\gamma(j); \xi < j(\beta)\}$. Now we have $j(\beta) = j(f)(0)$, where f is the mapping $\{(0,\beta)\}$. Since $\beta < \gamma$ holds, we deduce that $j(\beta)$ is itself γ-representable by j, and that the above set $\{\xi \in S_\gamma(j); \xi < j(\beta)\}$ is a proper subset of $S_\gamma(j)$. So the order type of $S_\gamma(j')$, which is that of $\{\xi \in S_\gamma(j); \xi < j(\beta)\}$, is strictly smaller than the order type of $S_\gamma(j)$. ⊣

We can now prove the Steel theorem easily.

Proof. Assume for a contradiction that there exists an ordinal γ satisfying $\gamma < \lambda$ and $\gamma > \mathrm{crit}(j_n)$ for every n. We may assume that γ is an inaccessible cardinal: indeed, by Kunen's bound, there exists an integer m such that $j_1^m(\mathrm{crit}(j_1)) \geqslant \gamma$ holds, and we know that $j_1^m(\mathrm{crit}(j_1))$ is inaccessible. Now 1.30 applies to each pair (j_n, j_{n+1}), showing that the order types of the sets $S_\gamma(j_n)$ make a decreasing sequence, which is impossible. ⊣

1.31 Theorem (Laver). *Assume $j : V_\lambda \prec V_\lambda$.*

(i) *The ordinals $\mathrm{crit}_n(j)$ are cofinal in λ, i.e., there exists no θ with $\theta < \lambda$ such that $\mathrm{crit}_n(j) < \theta$ holds for every n.*

(ii) *For every iterate i of j, we have $\mathrm{crit}(i) = \mathrm{crit}_m(j)$ for some integer m, and, therefore, i is not $\mathrm{crit}_m(j)$-equivalent to the identity.*

Proof. (i) By definition, every entry in the sequence j, $j_{[2]}$, $j_{[3]}$, ... is a left divisor of the next one, hence the Laver-Steel theorem implies that the critical ordinals of j, $j_{[2]}$, ... are cofinal in λ. By definition, these critical ordinals are exactly the ordinals $\mathrm{crit}_n(j)$.

(ii) Proposition 1.25 implies that either $\mathrm{crit}(i) > \mathrm{crit}_m(j)$ holds for every m, or there exists m satisfying $\mathrm{crit}(i) = \mathrm{crit}_m(j)$. By (i), the first case is impossible. $\dashv$

Observe that the point in the previous argument is really the Steel theorem, because 1.25 or 1.23 alone do not preclude the critical ordinal of some iterate i lying above all $\mathrm{crit}_m(j)$'s.

If follows from the previous result that, for every m, the image under j of the critical ordinal $\mathrm{crit}_m(j)$ is again an ordinal of the form $\mathrm{crit}_n(j)$. Indeed, $\mathrm{crit}_m(j)$ is the critical ordinal of $j_{[2^m]}$, and, therefore, $j(\mathrm{crit}_m(j))$ is the critical ordinal of $j[j_{[2^m]}]$, hence the critical ordinal of some iterate of j and, therefore, an ordinal of the form $\mathrm{crit}_n(j)$ for some finite n.

1.6. Counting the critical ordinals

As we already observed, the definition of an elementary embedding implies that the critical ordinal of $j[k]$ is the image under j of the critical ordinal of k, and it follows that every embedding in $\mathcal{E}_\lambda$ induces an increasing injection on the critical ordinals of $\mathcal{E}_\lambda$. In particular, every iterate of an embedding j acts on the critical ordinals of the iterates of j, which we have seen in the previous section consists of an ω-indexed sequence $(\mathrm{crit}_n(j))_{n<\omega}$. Let us introduce, for $j : V_\lambda \prec V_\lambda$, two mappings $\hat{\jmath}, \tilde{\jmath} : \omega \to \omega$ by

$$\hat{\jmath}(m) = p \quad \text{if and only if} \quad j(\mathrm{crit}_m(j)) = \mathrm{crit}_p(j),$$

and $\tilde{\jmath}(n) = \hat{\jmath}^n(0)$. By definition, $\mathrm{crit}_{\tilde{\jmath}(n)}$ is $j^n(\mathrm{crit}_0(j))$, so, if we use κ for $\mathrm{crit}(j)$ and κ_n for $j^n(\kappa)$, we simply have $\mathrm{crit}_{\tilde{\jmath}(n)} = \kappa_n$: thus $\tilde{\jmath}(n)$ is the number of critical ordinals of iterates of j below κ_n.

The aim of this section is to prove the following result:

1.32 Theorem (Dougherty [7]). *For $j : V_\lambda \prec V_\lambda$, the function $\tilde{\jmath}$ grows faster than any primitive recursive function.*

For the rest of the section, we fix $j : V_\lambda \prec V_\lambda$, and write γ_m for $\mathrm{crit}_m(j)$. Thus $\hat{\jmath}$ is determined by $\gamma_{\hat{\jmath}(m)} = j(\gamma_m)$ and $\tilde{\jmath}$ by $\gamma_{\tilde{\jmath}(n)} = j^n(\gamma_0)$. We are going to establish lower bounds for the values of the function $\tilde{\jmath}$. The first values of the function $\tilde{\jmath}$ can be computed exactly by determining sequences of iterated values for $j_{[p]}$. We use the notation

$$i : \mapsto \theta_0 \mapsto \theta_1 \mapsto \dots$$

to mean that we have $\theta_0 = \mathrm{crit}(i)$, $\theta_1 = i(\theta_0)$ ($= \mathrm{crit}(i^{[2]})$), etc. For instance, by definition of $\tilde{\jmath}$, we have

$$j : \gamma_0 \mapsto \gamma_{\tilde{\jmath}(1)} \mapsto \gamma_{\tilde{\jmath}(2)} \mapsto \gamma_{\tilde{\jmath}(3)} \mapsto \dots .$$

Now, for each sequence of the form

$$i : \Vdash\!\!\to \ \theta_0 \mapsto \theta_1 \mapsto \theta_2 \mapsto \dots ,$$

we deduce for each elementary embedding j_0 a new sequence

$$j_0[i] : \Vdash\!\!\to \ j_0(\theta_0) \mapsto j_0(\theta_1) \mapsto j_0(\theta_2) \mapsto \dots .$$

Applying the previous principle to the above sequence with $j_0 = j$, and using $\tilde{\jmath}(1) = 1$, we obtain the sequence

$$j_{[2]} : \Vdash\!\!\to \gamma_1 \mapsto \gamma_{\tilde{\jmath}(2)} \mapsto \gamma_{\tilde{\jmath}(3)} \mapsto \dots .$$

Applying the same principle with $j_0 = j_{[2]}$, we obtain

$$j_{[3]} : \Vdash\!\!\to \gamma_0 \mapsto \gamma_{\tilde{\jmath}(2)} \mapsto \gamma_{\tilde{\jmath}(3)} \mapsto \dots .$$

Then $\gamma_2 = \mathrm{crit}(j_{[4]})$ implies $\gamma_2 = j_{[3]}(\gamma_0)$, so the previous sequence shows that the latter ordinal is $\gamma_{\tilde{\jmath}(2)}$, *i.e.*, we have proved $\gamma_{\tilde{\jmath}(2)} = \gamma_2$, and, therefore we have $\hat{\jmath}(1) = 2$. Similar (but more tricky) arguments give $\hat{\jmath}(2) = 4$. Equivalently, we have $\tilde{\jmath}(1) = 1$, $\tilde{\jmath}(2) = 2$, $\tilde{\jmath}(3) = 4$, which means that the critical ordinals of the right powers j, $j^{[2]}$, and $j^{[3]}$ are γ_1, γ_2, and γ_4 respectively.

We turn now to the proof of 1.32. The basic argument is the following simple observation.

1.33 Lemma. *Assume that some iterate i of j satisfies $i : \gamma_p \mapsto \gamma_q \mapsto \gamma_r$. Then we have $r - q \geqslant q - p$.*

Proof. As the restricion of i to ordinals is increasing, $\gamma_p < \alpha < \alpha' < \gamma_q$ implies $\gamma_q < i(\alpha) < i(\alpha') < \gamma_r$. Moreover, if α is the critical ordinal of i_1, $i(\alpha)$ is that of $i[i_1]$, and, if i_1 is an iterate of j, so is $i[i_1]$. Hence the number of critical ordinals of iterates of j between γ_q and γ_r, which is $r - q - 1$, is at least the number of critical ordinals of iterates of j between γ_p and γ_q, which is $q - p - 1$. ⊣

1.34 Definition. A sequence of ordinals $(\alpha_0, \dots, \alpha_p)$ is said to be *realizable* (with respect to j) if we have $i : \Vdash\!\!\to \alpha_0 \mapsto \dots \mapsto \alpha_p$ for some iterate i of j. We say that the sequence $(\alpha_0, \dots, \alpha_p)$ is a *base* for the sequence $\vec{\theta} = (\theta_0, \dots, \theta_n)$ if, for each $m < n$, the sequence $(\alpha_0, \dots, \alpha_p, \theta_m, \theta_{m+1})$ is realizable.

Observe that the existence of a base for a sequence $\vec\theta$ implies that $\vec\theta$ is increasing, and that, if $(a_0, \ldots, a_p)$ is a base for $\vec\theta$, so is every final subsequence of the form $(a_m, \ldots, \alpha_p)$: if i admits the critical sequence $\Vdash\!\!\to \alpha_0 \mapsto \ldots \mapsto \theta_m \mapsto \theta_{m+1}$, then $i^{[2]}$ admits the critical sequence $\Vdash\!\!\to \alpha_1 \mapsto \ldots \mapsto \theta_m \mapsto \theta_{m+1}$.

1.35 Lemma. *Assume that the sequence $(\theta_0, \theta_1, \ldots)$ admits a base. Then $\theta_n \geqslant \gamma_{2^n}$ holds for every n.*

Proof. Assume that (γ_p) is a base for $(\theta_0, \theta_1, \ldots)$. Define f by $\theta_n = \gamma_{f(n)}$. Lemma 1.33 gives $f(n+1) - f(n) \geqslant f(n) - p$ for every n. As $f(0) > p$ holds by definition, we deduce $f(n) \geqslant 2^n + p$ inductively. ⊣

For instance, the embedding $j_{[2]}$ leaves γ_0 fixed, and it maps $\gamma_{\tilde\jmath(r)}$ to $\gamma_{\tilde\jmath(r+1)}$ for $r \geqslant 1$. So its $(r-1)$-th power with respect to composition satisfies

$$(j_{[2]})^{r-1} : \Vdash\!\!\to \gamma_1 \mapsto \gamma_{\tilde\jmath(r)}, \quad \gamma_2 \mapsto \gamma_{\tilde\jmath(r+1)}.$$

Applying these values to the critical sequence of j, we obtain

$$(j_{[2]})^{r-1}[j] : \Vdash\!\!\to \gamma_0 \mapsto \gamma_{\tilde\jmath(r)} \mapsto \gamma_{\tilde\jmath(r+1)}.$$

Hence (γ_0) is a base for the sequence $(\gamma_{\tilde\jmath(1)}, \gamma_{\tilde\jmath(2)}, \ldots)$. Lemma 1.35 gives $\tilde\jmath(n) \geqslant 2^{n-1}$. In particular, we find $\tilde\jmath(4) \geqslant 8$. This bound destroys any hope of computing an exact value by applying the scheme used for the first values: indeed this would entail computing values until at least $j_{[255]}$. We shall see below that the value of $\tilde\jmath(4)$ is actually much larger than 8.

In order to improve the previous results, we use the following trick to expand the sequences admitting a base by inserting many intermediate new critical ordinals.

1.36 Lemma. *Assume that $(\alpha_0, \ldots, \alpha_p, \beta, \gamma)$ is realizable, $\vec\theta$ is based on (β) and it goes from γ to δ in n steps. Then there exists a sequence based on $(\alpha_0, \ldots, \alpha_p)$ that goes from β to δ in 2^n steps.*

Proof. We use induction on $n \geqslant 0$. For $n = 0$, the sequence (β, γ) works, since $(\alpha_0, \ldots, \alpha_p, \beta, \gamma)$ being realizable means that (β, γ) is based on $(\alpha_0, \ldots, \alpha_p)$. For $n > 0$, let δ' be the next to last term of $\vec\theta$. The induction hypothesis gives a sequence $\vec\tau\,'$ based on $(\alpha_0, \ldots, \alpha_p)$ that goes from γ to δ' in 2^{n-1} steps. As (δ', δ) is based on (β), there exists an embedding i satisfying

$$i : \Vdash\!\!\to \beta \mapsto \delta' \mapsto \delta.$$

We define the sequence $\vec\tau$ by extending $\vec\tau\,'$ with 2^{n-1} additional terms

$$\tau_{2^{n-1}+m} = i(\tau'_m) \qquad \text{for } 1 \leqslant m \leqslant 2^{n-1}.$$

By hypothesis, we have $\tau'_{2^{n-1}} = \delta'$, hence $\tau_{2^n} = i(\delta') = \delta$. So $\vec{\tau}$ goes from β to δ in 2^n steps. Moreover, $(\alpha_0, \ldots, \alpha_p)$ is a base for $\vec{\tau}\,'$, so, for $0 \leqslant m < 2^{n-1}$, there exists i'_m satisfying

$$i'_m : \Vdash\!\!\rightarrow \alpha_0 \mapsto \ldots \mapsto \alpha_p \mapsto \tau'_m \mapsto \tau'_{m+1}.$$

As β is the critical ordinal of i and $\alpha_p < \beta$ holds, this implies

$$i[i'_m] : \Vdash\!\!\rightarrow \alpha_0 \mapsto \ldots \mapsto \alpha_p \mapsto i(\tau'_m) \mapsto i(\tau'_{m+1}),$$

which shows that $(\alpha_0, \ldots, \alpha_p)$ is a base for $\vec{\tau}$. Note that the case $m = 0$ works because $\tau'_0 = \beta$ implies $i(\tau'_0) = i(\beta) = \delta' = \tau_{2^{n-1}}$, as is needed ⊣

By playing with the above construction one more time, we can obtain still longer sequences. In order to specify them, we use an *ad hoc* iteration of the exponential function, namely g_p inductively defined by $g_0(n) = n$, $g_{p+1}(0) = 0$, and $g_{p+1}(n) = g_{p+1}(n-1) + g_p(2^{g_{p+1}(n-1)})$. Thus, g_1 is an iterated exponential. Observe that $g_p(1) = 1$ holds for every p.

1.37 Lemma. *Assume that $(\beta_0, \ldots, \beta_{p+1}, \gamma)$ is realizable, $\vec{\theta}$ is based on (β_p, β_{p+1}) and it goes from γ to δ in n steps. Then there exists a sequence based on (β_{p+1}) that goes from γ to δ in $g_{p+1}(n)$ steps.*

Proof. We use induction on $p \geqslant 0$, and, for each p, on $n \geqslant 1$. For $n = 1$, the sequence (γ, δ) works, since, if i satisfies $\Vdash\!\!\rightarrow \beta_p \mapsto \beta_{p+1} \mapsto \gamma \mapsto \delta$, then $i^{[2]}$ satisfies $\Vdash\!\!\rightarrow \beta_{p+1} \mapsto \gamma \mapsto \delta$. Assume $n \geqslant 2$. Let δ' be the next to last term of $\vec{\theta}$. By induction hypothesis, there exists a sequence $\vec{\tau}\,'$ based on (β_{p+1}) that goes from γ to δ' in $g_{p+1}(n-1)$ steps. As in 1.36, we complete the sequence by appending new terms, but, before translating it, we still fatten it one or two more times. First, we apply 1.36 to construct a new sequence $\vec{\tau}\,''$ based on (β_p, β_{p+1}) that goes from β_{p+1} to δ' in $2^{g_{p+1}(n-1)}$ steps and is based on (β_{p-1}, β_p) for $p \neq 0$ (*resp.* on (β_p) for $p = 0$). For $p \neq 0$, we are in position for applying the current lemma with $p - 1$ to the sequence of $\vec{\tau}\,''$. So we obtain a new sequence $\vec{\tau}\,'''$ based on (α_p), and going from β_{p+1} to δ' in $g_p(2^{g_{p+1}(n-1)})$ steps. For $p = 0$, we simply take $\vec{\tau}\,''' = \vec{\tau}\,''$: as $g_0(N) = N$ holds, this remains consistent with our notations. Now we make the translated copy: we choose i satisfying $\Vdash\!\!\rightarrow \beta_p \mapsto \beta_{p+1} \mapsto \delta' \mapsto \delta$, and complete $\vec{\tau}\,'$ with the new terms

$$\tau_{g_{p+1}(n-1)+m} = i(\tau'''_m) \qquad \text{for} \qquad 0 < m \leqslant g_p(2^{g_{p+1}(n-1)}).$$

The sequence $\vec{\tau}$ has length $g_{p+1}(n-1) + g_p(2^{g_{p+1}(n-1)}) = g_{p+1}(n)$, and it goes from γ to $i(\delta')$, which is δ. It remains to verify the base condition for the new terms. Now assume that i'''_m satisfies $\Vdash\!\!\rightarrow \beta_p \mapsto \tau'''_m \mapsto \tau'''_{m+1}$. As in the proof of 1.36, we see that $i[i'''_m]$ satisfies $\Vdash\!\!\rightarrow \beta_{p+1} \mapsto i(\tau'''_m) \mapsto i(\tau'''_{m+1})$, which completes the proof, as $i(\tau'''_0) = \delta'$ guarantees continuity. ⊣

By combining 1.36 and 1.37, we obtain:

1.38 Lemma. *Assume that* $(\beta_0, \ldots, \beta_{p+1}, \gamma)$ *is realizable,* $\vec{\theta}$ *is based on* (β_p, β_{p+1}) *and it goes from* γ *to* δ *in* n *steps. Then there exists a sequence based on* (β_0) *that goes from* β_1 *to* δ *in* $h_1(h_2(\ldots(h_{p+1}(n))\ldots))$ *steps, where* $h_q(m)$ *is defined to be* $2^{g_q(m)}$.

Proof. We use induction on $p \geqslant 0$. In every case, 1.37 constructs from $\vec{\theta}$ a new sequence $\vec{\theta}'$ based on (β_{p+1}) going from β_{p+1} to δ in $g_{p+1}(n) + 1$ steps. Then, 1.36 constructs from $\vec{\theta}'$ a new sequence $\vec{\theta}''$ that goes from (β_{p+1}) to δ in $2^{g_{p+1}(n)} + 1 = h_{p+1}(n) + 1$ steps, a sequence based on (α_{p-1}, α_p) for $p \neq 0$, and on (α_p) for $p = 0$. For $p = 0$, the sequence $\vec{\theta}''$ works. Otherwise, we are in position for applying the induction hypothesis to $\vec{\theta}''$. ⊣

We deduce the following lower bound for the function $\tilde{\jmath}$.

1.39 Proposition. *Assume* $j : V_\lambda \prec V_\lambda$. *Then, for* $n \geqslant 3$, *we have*

$$\tilde{\jmath}(r) \geqslant 2^{h_1(h_2(\ldots(h_{n-2}(1))\ldots))}. \tag{I.13}$$

Proof. By definition, $(\gamma_{\tilde{\jmath}(n-1)}, \gamma_{\tilde{\jmath}(n)})$ is based on $(\gamma_{\tilde{\jmath}(n-3)}, \gamma_{\tilde{\jmath}(n-2)})$, and the auxiliary sequence $(\gamma_0, \ldots, \gamma_{\tilde{\jmath}(n-2)})$ is realizable. Indeed, j satisfies

$$j : \Vdash\!\!\to \gamma_{\tilde{\jmath}(0)} \mapsto \gamma_{\tilde{\jmath}(1)} \mapsto \gamma_{\tilde{\jmath}(2)} \mapsto \gamma_{\tilde{\jmath}(3)},$$

and, therefore, we have

$$j^{[n+1]} : \Vdash\!\!\to \gamma_{\tilde{\jmath}(n)} \mapsto \gamma_{\tilde{\jmath}(n+1)} \mapsto \gamma_{\tilde{\jmath}(n+2)} \mapsto \gamma_{\tilde{\jmath}(n+3)}$$

for every n. By applying 1.38, we find a new sequence based on (γ_0) that goes from γ_1 to $\gamma_{\tilde{\jmath}(n)}$ in $h_1(h_2(\ldots(h_{n-2}(1))\ldots))$ steps. We conclude using 1.38. ⊣

We thus proved $\tilde{\jmath}(4) \geqslant 2^8 = 256$, and $\tilde{\jmath}(5) \geqslant 2^{h_1(h_2(h_3(1)))} = 2^{2^{g_1(16)}}$. It follows that $\tilde{\jmath}(5)$ is more than a tower of base 2 exponentials of height 17.

Let us recall that the Ackermann function f_p^{Ack} is defined inductively by $f_0^{\mathrm{Ack}}(n) = n+1$, $f_{p+1}^{\mathrm{Ack}}(0) = f_p^{\mathrm{Ack}}(1)$, and $f_{p+1}^{\mathrm{Ack}}(n+1) = f_p^{\mathrm{Ack}}(f_{p+1}^{\mathrm{Ack}}(n))$. We put $f_\omega^{\mathrm{Ack}}(n) = f_n^{\mathrm{Ack}}(n)$. Using the similarity between the definitions of f_p^{Ack} and g_p, it is easy to complete the proof of 1.32.

Proof. The function f_ω^{Ack} is known to grow faster than every primitive recursive function, so it is enough to show $2^{h_1(h_2(\ldots(h_{n-2}(1))\ldots))} \geqslant f_\omega^{\mathrm{Ack}}(n-1)$ for $n \geqslant 5$. First, we have $g_p(n+3) > f_p^{\mathrm{Ack}}(n)$ for all p, n. This is obvious for $p = 0$. Otherwise, for $n = 0$, using $g_p(2) \geqslant 3$, we find

$$g_p(3) > g_{p-1}(2^{g_p(2)}) > f_{p-1}^{\mathrm{Ack}}(6) > f_{p-1}^{\mathrm{Ack}}(1) = f_p^{\mathrm{Ack}}(0).$$

Then, for $n > 0$, we obtain

$$g_p(n+3) > g_{p-1}(2^{g_p(n+2)}) > g_{p-1}(f_p^{\text{Ack}}(n-1)+3)$$
$$> f_{p-1}^{\text{Ack}}(f_p^{\text{Ack}}(n-1)) = f_p^{\text{Ack}}(n).$$

Finally, we have $g_2(n) = n+2$ for every n, and therefore

$$2^{h_1(h_2(\dots(h_{n+2}(1))\dots))} = 2^{h_1(h_2(\dots(h_{n+1}(2))\dots))} = 2^{h_1(h_2(\dots(h_n(2^{n+3}))\dots))}$$
$$> g_n(2^{n+3})) \geqslant g_n(n+3) > f_n^{\text{Ack}}(n),$$

hence $2^{h_1(h_2(\dots(h_{n+2}(1))\dots))} \geqslant f_\omega^{\text{Ack}}(n-1)$. ⊣

Let us finally mention without proof the following strengthening of the lower bound for $\tilde{\jmath}(4)$:

1.40 Proposition (Dougherty). *For $j : V_\lambda \prec V_\lambda$, we have*

$$\tilde{\jmath}(4) \geqslant f_9^{\text{Ack}}(f_8^{\text{Ack}}(f_8^{\text{Ack}}(254))).$$

In other words, there are at least the above huge number of critical ordinals below κ_4 in $\text{Iter}(j)$.

2. The word problem for self-distributivity

The previous results about iterations of elementary embeddings have led to several applications outside Set Theory. The first application deals with free LD-systems and the word problem for the self-distributivity law $x(yz) = (xy)(xz)$. In 1989, Laver deduced from 1.20 that the LD-system $\text{Iter}(j)$ has a specific algebraic property, namely that left division has no cycle in this LD-system, and he derived a solution for the word problem of (LD). Here we shall describe these results, following the independent and technically more simple approach of [4].

2.1. Iterated left division in LD-systems

For $(S, *)$ a (nonassociative) algebraic system, and x, y in S, we say that x is a *left divisor* of y if $y = x*z$ holds for some z in S; we say that x is an *iterated left divisor* of y, and write $x \sqsubset y$ if, for some positive k, there exist $z_1, \dots, z_k$ satisfying $y = (\dots((x*z_1)*z_2)\dots)*z_k$. So $\sqsubset$ is the transitive closure of left divisibility. In the sequel, we shall be interested in LD-systems where left division (or, equivalently, iterated left division) has no cycle.

We write T_n for the set of all terms constructed using the variables $x_1, \dots, x_n$ and a binary operator $*$, and T_∞ for the union of all T_n's. We denote by $=_{LD}$ the congruence on T_∞ generated by all pairs of the form

$(t_1{*}(t_2{*}t_3)), (t_1{*}t_2){*}(t_1{*}t_3))$. Then, by standard arguments, $T_n/=_{LD}$ is a free LD-system with n generators, which we shall denote by F_n. The word problem of (LD) is the question of algorithmically deciding the relation $=_{LD}$.

2.1 Theorem (Dehornoy [4]; also Laver [18] for an independent approach). *Assume that there exists at least one LD-system where left division has no cycle.*

(i) *Iterated left division in a free LD-system with one generator is a linear ordering.*

(ii) *The word problem of* (LD) *is decidable.*

The rest of this subsection is an outline of the proof of this statement, which can be skipped by a reader exclusively interested in Set Theory.

2.2 Definition. For t, t' terms in T_∞, we say that t' is an *LD-expansion* of t if we can go from t to t' by applying finitely many transformations consisting of replacing a subterm of the form $t_1{*}(t_2{*}t_3)$ with the corresponding term $(t_1{*}t_2){*}(t_1{*}t_3)$.

By definition, t' being LD-equivalent to t means that we can transform t to t' by applying the law (LD) in either direction, *i.e.*, from $x{*}(y{*}z)$ to $(x{*}y){*}(x{*}z)$ or *vice versa*, while t' being an LD-expansion of t means that we transform t to t' by applying (LD), but only in the expanding direction, *i.e.*, from $x{*}(y{*}z)$ to $(x{*}y){*}(x{*}z)$, but not in the converse, contracting direction.

2.3 Definition. For t a term and k small enough, we denote by $\mathrm{left}^k(t)$ the kth iterated left subterm of t: we have $\mathrm{left}^0(t) = t$ for every t, and $\mathrm{left}^k(t) = \mathrm{left}^{k-1}(t_1)$ for $t = t_1{*}t_2$ and $k \geqslant 1$. For t_1, t_2 in T_∞, we say that $t_1 \sqsubset_{LD} t_2$ is true if we have $t'_1 = \mathrm{left}^k(t'_2)$ for some k, t'_1, t'_2 satisfying $k \geqslant 1$, $t'_1 =_{LD} t_1$, and $t'_2 =_{LD} t_2$.

By construction, saying that $t_1 \sqsubset_{LD} t_2$ is true in T_1 is equivalent to saying that the class of t_1 in the free LD-system F_1 is an iterated left divisor of the class of t_2. The core of the argument is:

2.4 Proposition. *Let t_1, t_2 be one variable terms in T_1. Then at least one of $t_1 \sqsubset_{LD} t_2$, $t_1 =_{LD} t_2$, $t_2 \sqsubset_{LD} t_1$ holds.*

2.5 Corollary. *If $(S, *)$ is an LD-system with one generator, then any two elements of S are comparable with respect to iterated left division.*

Proving 2.4 relies on three specific properties of left self-distributivity. As in Section 1, we use the notation $x^{[n]}$ for the nth right power of x.

2.6 Lemma. *For every term t in T_1, we have $x^{[n+1]} =_{LD} t{*}x^{[n]}$ for n sufficiently large.*

Proof. We use induction on t. For $t = x$, we have $x^{[n+1]} = x*x^{[n]}$ for every n, by definition. Assume now $t = t_1*t_2$. Assuming that the result is true for t_1 and t_2, we obtain for n sufficiently large

$$x^{[n+1]} =_{LD} t_1*x^{[n]} =_{LD} t_1*(t_2*x^{[n-1]}) \\ =_{LD} (t_1*t_2)*(t_1*x^{[n-1]}) =_{LD} (t_1*t_2)*x^{[n]} = t*x^{[n]},$$

which is the result for t. ⊣

2.7 Lemma. *Assume that* $\mathrm{left}^n(t)$ *is defined, and* t' *is an LD-expansion of* t*. Then* $\mathrm{left}^{n'}(t')$ *is an LD-expansion of* $\mathrm{left}^n(t)$ *for some* $n' \geqslant n$*.*

Proof. It suffices to prove the result when t' is obtained by replacing exactly one subterm t_0 of t of the form $t_1*(t_2*t_3)$ with the corresponding $(t_1*t_2)*(t_1*t_3)$. If t_0 is $\mathrm{left}^j(t)$ with $j < n$, then $\mathrm{left}^{n+1}(t')$ is equal to $\mathrm{left}^n(t)$; if t_0 is $\mathrm{left}^j(t)$ with $j \geqslant n$, then $\mathrm{left}^n(t')$ is an LD-expansion of $\mathrm{left}^n(t)$; otherwise, we have $\mathrm{left}^n(t') = \mathrm{left}^n(t)$. ⊣

2.8 Lemma. *Any two LD-equivalent terms admit a common LD-expansion.*

Proof (sketch). The point is to prove that, if t' and t'' are any two LD-expansions of some term t, then t' and t'' admit a common LD-expansion. Now, let us say that t' is a p-expansion of t if t' is obtained from t by applying (LD) at most p times (in the expanding direction). Then, for every term t, one can explicitly define a certain LD-expansion ∂t of t that is a common LD-expansion of all 1-expansions of t, then check that, if t' is an LD-expansion of t, then $\partial t'$ is an LD-expansion of ∂t, and deduce using an induction that, for every p, the term $\partial^p t$ is an LD-expansion of all p-expansions of t. It follows that, if t' and t'' are any two LD-expansions of some term t, then t' and t'' admit common LD-expansions, namely all terms $\partial^p t$ with p sufficiently large. ⊣

It is now easy to complete the proof of 2.4.

Proof. Let t_1, t_2 be arbitrary terms in T_1. By 2.6, we have $t_1*x^{[n]} =_{LD} x^{[n+1]} =_{LD} t_2*x^{[n]}$ for n sufficiently large. Fix such a n. By 2.8, the terms $t_1*x^{[n]}$ and $t_2*x^{[n]}$ admit a common LD-expansion, say t. By 2.7, there exist nonnegative integers n_1, n_2 such that, for $i = 1, 2$, the term $\mathrm{left}^{n_i}(t)$ is an LD-expansion of $\mathrm{left}(t_i*x^{[n]})$, *i.e.*, of t_i. Thus we have $t_1 =_{LD} \mathrm{left}^{n_1}(t)$, and $t_2 =_{LD} \mathrm{left}^{n_2}(t)$. Three cases may occur: for $n_1 > n_2$, $\mathrm{left}^{n_1}(t)$ is an iterated left subterm of $\mathrm{left}^{n_2}(t)$, and, therefore, $t_1 \sqsubset_{LD} t_2$ holds; for $n_1 = n_2$, t_1 and t_2 both are LD-equivalent to $\mathrm{left}^{n_1}(t)$, and $t_1 =_{LD} t_2$ is true; finally, for $n_1 < n_2$, $\mathrm{left}^{n_2}(t)$ is an iterated left subterm of $\mathrm{left}^{n_1}(t)$, and, therefore, $t_2 \sqsubset_{LD} t_1$ holds. ⊣

Finally, we can complete the proof of 2.1.

Proof. (i) Proposition 2.4 tells us that any two elements of the free LD-system F_1 are comparable with respect to the iterated left divisibility relation. Assume that S is any LD-system. The universal property of free LD-systems guarantees that there exists a homomorphism π of F_1 into S. If $(a_1, \dots, a_n)$ is a cycle for left division in F_1, then $(\pi(a_1), \dots, \pi(a_n))$ is a cycle for left division in S. So, if there exists at least one LD-system S where left division has no cycle, the same must be true for F_1, which means that the iterated left divisibility relation of F_1 is irreflexive. As it is always transitive, it is a (strict) linear ordering.

(ii) Let us consider the case of one variable terms first. When we are given two terms t_1, t_2 in T_1, we can decide wheher $t_1 =_{LD} t_2$ is true as follows: we systematically enumerate all pairs (t'_1, t'_2) such that t'_1 is LD-equivalent to t_1 and t'_2 is LD-equivalent to t_2. By 2.4, there will eventually appear some pair (t'_1, t'_2) such that either t'_1 and t'_2 are equal, or t'_1 is a proper iterated left subterm of t'_2, or t'_2 is a proper iterated left subterm of t'_1. In the first case, we conclude that $t_1 =_{LD} t_2$ is true, in the other cases, we can conclude that $t_1 =_{LD} t_2$ is false whenever we know that $t \sqsubset_{LD} t'$ excludes $t =_{LD} t'$, *i.e.*, whenever we know that left division has no cycle in F_1.

The case of terms with several variables is not really more difficult. For t a general term, let $t^\dagger$ denote the term obtained from t by replacing all variables with x_1. Assume we are given t_1, t_2 in T_n. We can decide whether $t_1 =_{LD} t_2$ is true as follows. First we compare $t_1^\dagger$ and $t_2^\dagger$ as above. If the latter terms are not LD-equivalent, then t_1 and t_2 are not LD-equivalent either (as $t \mapsto t^\dagger$ trivially preserves LD-equivalence). Otherwise, we can effectively find a common LD-expansion t of $t_1^\dagger$ and $t_2^\dagger$. Then we consider the LD-expansion t'_1 of t_1 obtained in the same way as t is obtained from $t_1^\dagger$, *i.e.*, by applying (LD) at the same successive positions, and, similarly, we consider t'_2 obtained from t_2 as t is obtained from $t_2^\dagger$. By constuction, we have $(t'_1)^\dagger = (t'_2)^\dagger = t$, *i.e.*, the terms t'_1, t'_2, and t coincide up to the name of the variables. Two cases may occur. Either t'_1 and t'_2 are equal, in which case we conclude that $t_1 =_{LD} t_2$ is true, or t'_1 and t'_2 have some variable clash, in which case we can conclude that $t_1 =_{LD} t_2$ is false. Indeed, using the techniques of 2.8, it is not hard to prove that $t'_1 =_{LD} t'_2$ would imply $t_0 =_{LD} \mathrm{left}^n(t_0)$ for some term t_0 effectively constructed from t'_1 and t'_2, thus would contradict the hypothesis that left divisibility in F_1 has no cycle. ⊣

2.2. Using elementary embeddings

In the mid 1980's, R. Laver showed the following:

2.9 Proposition (Laver). *Left division in the LD-system $\mathcal{E}_\lambda$ has no cycle.*

Proof. Assume that $j_1, \dots, j_n$ is a cycle for left division in $\mathcal{E}_\lambda$. Consider the infinite periodic sequence $j_1, \dots, j_n, j_1, \dots, j_n, j_1, \dots$. The Laver–Steel theorem applies, and it asserts that the supremum of the critical ordinals in this sequence is λ. But, on the other hand, there are only n different embeddings in the sequence, and the supremum of finitely many ordinals below λ cannot be λ, a contradiction. $\dashv$

The original proof of the previous result in [18] did not use the Laver-Steel theorem, but instead a direct computation based on 1.20.

Using the results of Subsection 2.1, we immediately deduce:

2.10 Theorem (Laver, 1989). *Assume Axiom (I3). Then:*

(i) *Iterated left division in a free LD-system with one generator is a linear ordering.*

(ii) *The word problem of* (LD) *is decidable.*

Another application of 2.9 is a complete algebraic characterization of the LD-system made by the iterations of an elementary embedding.

2.11 Lemma ("Laver's criterion"). *A sufficient condition for an LD-system S with one generator to be free is that left division in S has no cycle.*

Proof. Assume that left division in S has no cycle. Let π be a surjective homomorphism of F_1 onto S, which exists by the universal property of F_1. Let x, y be distinct elements of F_1. By 2.5, at least one of $x \sqsubset y$, $y \sqsubset x$ is true in F_1, which implies that at least one of $\pi(x) \sqsubset \pi(y)$, $\pi(y) \sqsubset \pi(x)$ is true in S. The hypothesis that left division has no cycle in S implies that, in S, the relation $a \sqsubset b$ excludes $a = b$. So, here, we deduce that $\pi(x) \neq \pi(y)$ is true in every case, which means that π is injective, and, therefore, it is an isomophism, *i.e.*, S is free. $\dashv$

We deduce the first part of the following result

2.12 Theorem (Laver). *Assume $j : V_\lambda \prec V_\lambda$. Then* $\mathrm{Iter}(j)$ *equipped with the application operation is a free LD-system, and* $\mathrm{Iter}^*(j)$ *equipped with application and composition is a free LD-monoid.*

We skip the details for the LD-monoid structure, which are easy. The general philosophy is that, in an LD-monoid, most of the nontrivial information is concentrated in the self-distributive operation. In particular, if X is any set and F_X is the free LD-system based on X, then the free LD-monoid based on X is the free monoid generated by F_X, quotiented under the congruence generated by the pairs $(x{\cdot}y, (x{*}y){\cdot}x)$. It easily follows that there exists a realizaton of the free monoid based on X inside the free LD-system based on X. So, in particular, every solution for the word problem of (LD) gives a solution for the word problem of the laws that define LD-monoids.

2.3. Avoiding elementary embeddings

The situation created by 2.10 was strange, as one would expect no link between large cardinals and such a simple combinatorial property as the word problem of (LD). Therefore, finding an alternative proof not relying on a large cardinal axiom—or proving that some set-theoretical axiom is needed here—was a natural challenge.

2.13 Theorem (Dehornoy [5]). *That left division in the free LD-system with one generator has no cycle is a theorem of ZFC.*

Outline of proof. The argument of [5] consists of studying the law (LD) by introducing a certain monoid $\mathcal{G}_{LD}$ that captures its specific geometry. Viewing terms as binary trees, one considers, for each possible address α of a subterm, the partial operator Ω_α on terms corresponding to applying (LD) at position α in the expanding direction, *i.e.*, expanding the subterm rooted at the vertex specified by α. If $\mathcal{G}_{LD}$ is the monoid generated by all operators $\Omega_\alpha^{\pm 1}$ using composition, then two terms t, t' are LD-equivalent if and only if some element of $\mathcal{G}_{LD}$ maps t to t'. Because the operators Ω_α are partial in an essential way, the monoid $\mathcal{G}_{LD}$ is not a group. However, one can guess a presentation of $\mathcal{G}_{LD}$ and work with the group G_{LD} admitting that presentation. Then the key step is to construct a realization of the free LD-system with one generator in some quotient of G_{LD}, a construction that is reminiscent of Henkin's proof of the completeness theorem. The problem is to associate with each term t in T_1 a distinguished operator in $\mathcal{G}_{LD}$ (or its copy in the group G_{LD}) in such a way that the obstruction to satisfying (LD) can be controlled. The solution is given by 2.6: the latter asserts that, for each term t, the term $x^{[n+1]}$ is LD-equivalent to $t{*}x^{[n]}$ for n sufficiently large, so some operator χ_t in $\mathcal{G}_{LD}$ must map $x^{[n+1]}$ to $t{*}x^{[n]}$, *i.e.*, in some sense, construct the term t. Moreover 2.6 gives an explicit inductive definition of χ_t in terms of χ_{t_1} and χ_{t_2} when t is $t_1{*}t_2$. Translating this definition into G_{LD} yields a self-distributive operation on some quotient of G_{LD}, and proving that left division has no cycle in the LD-system so obtained is then easy—even if a number of verifications are in order. ⊣

2.14 Remark. A relevant geometry group can be constructed for every algebraic law (or family of algebraic laws). When the self-distributivity law is replaced with the associativity law, the corresponding group is Richard Thompson's group F [2]. So G_{LD} is a sort of higher analog to F.

Theorem 2.13 allows one to eliminate any set-theoretical assumption from the statements of 2.10. Actually, it gives more. Indeed, the quotient of G_{LD} appearing in the above proof turns out to be Artin's braid group B_∞, and the results about G_{LD} led to unexpected braid applications.

Being a rather ubiquitous object, Artin's braid group B_n admits many equivalent definitions. Usually, B_n is introduced for $2 \leqslant n \leqslant \infty$ as the group generated by elements σ_i, $1 \leqslant i < n$, subject to the relations

$$\sigma_i\sigma_j = \sigma_j\sigma_i \text{ for } |i-j| \geqslant 2, \quad \sigma_i\sigma_j\sigma_i = \sigma_j\sigma_i\sigma_j \text{ for } |i-j| = 1. \tag{I.14}$$

The connection with braid diagrams comes when σ_i is associated with an n-strand diagram where the $(i+1)$st strand crosses over the ith strand; then the relations in (I.14) correspond to ambient isotopy.

2.15 Theorem (Dehornoy [5]). *For b_1, b_2 in B_∞, say that $b_1 < b_2$ holds if, among all possible expressions of $b_1^{-1}b_2$ in terms of the $\sigma_i^{\pm1}$, there is at least one where the generator σ_i of minimal index i occurs only positively (i.e., no σ_i^{-1}). Then the relation $<$ is a left-invariant linear ordering on B_∞.*

The result is a consequence of 2.13. Indeed, one can prove that there exists a (partial) action of the group B_n on the nth power of every left cancellative LD-system; one then obtains a linear ordering on B_n by defining, for b_1, b_2 in B_n and $\vec{a}$ in F_1^n, the relation $b_1 <_{\vec{a}} b_2$ to mean that $\vec{a} \cdot b_1$ is lexicographically smaller than $\vec{a} \cdot b_2$. One then checks that $<_{\vec{a}}$ does not depend on the choice of $\vec{a}$ and it coincides with the relation $<$ of 2.15. In this way, one obtains the previously unknown result that the braid groups are orderable. A number of alternative characterizations of the braid ordering have been found subsequently, in particular in terms of homeomorphisms of a punctured disk, and of hyperbolic geometry [6]. Various results have been derived, in particular new efficient solutions for the word problem of braids with possible cryptographic applications.

The following result, first discovered by Laver (well foundedness), was then made more explicit by Burckel (computation of the order type):

2.16 Theorem (Laver [20], Burckel [1]). *For each n, the restriction of the braid ordering to the braids that can be expressed without any σ_i^{-1} is a well-ordering of type $\omega^{\omega^{n-2}}$.*

Returning to self-distributivity, we can mention as a last application a simple solution to the word problem of (LD) involving the braid group B_∞. Indeed, translating 2.6 to B_∞ leads to the explicit operation

$$x{*}y = x\,\mathrm{sh}(y)\,\sigma_1\,\mathrm{sh}(x)^{-1}, \tag{I.15}$$

where sh is the endomorphism that maps σ_i to σ_{i+1} for every i. Laver's criterion 2.11 implies that every sub-LD-system of $(B_\infty, *)$ with one generator is free. Then, in order to decide whether two terms on one variable are LD-equivalent, it suffices to compare their evaluations in B_∞ when x is mapped to 1 and (I.15) is used, which is easy. Note that, once (I.15) has been guessed, it is trivial to check that it defines a self-distributive operation on B_∞, and, therefore, any argument proving that left division in $(B_\infty, *)$ has no cycle is sufficient for fulfilling the assumptions of 2.1 without resorting to the rather convoluted construction of G_{LD}. Several such arguments have been given now, in particular by D. Larue using automorphisms of a free group [16] and by I. Dynnikov using laminations [6].

The developments sketched above have no connection with Set Theory. As large cardinal axioms turned out to be unnecessary, one could argue that Set Theory is not involved here, and deny that any of these developments can be called an application of Set Theory. The author disagrees with such an opinion. Had not Set Theory given the first hint that the algebraic properties of LD-systems are a deep subject [17, 3], then it is not clear that anyone would have tried to really understand the law (LD). The production of an LD-system with acyclic division using large cardinals gave evidence that some other example might be found in ZFC, and hastened its discovery. Without Set Theory, it is likely that the braid ordering would not have been discovered, at least as soon[1]: could not this be accepted as a definition for the braid ordering to be considered an application of Set Theory? It is tempting to compare the role of Set Theory here with the role of physics when it gives evidence for some formulas that remain then to be proved in a standard mathematical framework.

3. Periods in the Laver tables

Here we describe another combinatorial application of the set theoretical results of Section 1. This application involves some finite LD-systems discovered by R. Laver in his study of iterations of elementary embeddings [19], and, in contrast to the results mentioned in Section 2, the results have not yet received any ZF proof.

3.1. Finite LD-systems

The results of Subsection 1.4 give, for each $j : V_\lambda \prec V_\lambda$, an infinite family of finite quotients of $\mathrm{Iter}(j)$, namely one with 2^n elements for each n. The finite LD-systems obtained in this way will be called the Laver tables here. In this section, we shall show how to construct the Laver tables directly, and list some of their properties.

Let us address the question of constructing a finite LD-system with one generator. We start with an incomplete table on the elements $1, \ldots, N$, and try to complete it by using the self-distributivity law. Here, we consider the case when the first column is assumed to be cyclic, *i.e.*, we have

$$a{*}1 = a+1, \text{ for } a = 1, \ldots, N-1, \quad N{*}1 = 1. \tag{I.16}$$

[1] *A posteriori*, it became clear that the orderability of braid groups could have been deduced from old work by Nielsen, but this was not noted until recently.

3.1 Lemma. (i) *For every N, there exists a unique operation $*$ on $\{1, \dots, N\}$ satisfying* (I.16) *and, for all a, b,*

$$a{*}(b{*}1) = (a{*}b){*}(a{*}1).$$

(ii) *The following relations hold in the resulting system:*

$$a{*}b \begin{cases} = b & \text{for } a = N, \\ = a+1 & \text{for } b = 1, \text{ and for } a{*}(b-1) = N, \\ > a{*}(b-1) & \text{otherwise.} \end{cases}$$

For $a < N$, there exists $p \leqslant N - a$ and $c_1 = a+1 < c_2 < \dots < c_p = N$ such that, for every b, we have $a{*}b = c_i$ with $i = b \pmod p$, hence, in particular, $a{*}b > a$.

Let us denote by S_N the system given by 3.1. At this point, the question is whether S_N is actually an LD-system: by construction, certain occurrences of (LD) hold in the table, but this does not guarantee that the law holds for all triples. Actually, it need not: for instance, the reader can check that, in S_5, one has $2{*}(2{*}2) = 3 \neq (2{*}2){*}(2{*}2) = 5$.

3.2 Proposition. (i) *If N is not a power of 2, there exists no LD-system satisfying* (I.16).

(ii) *For each n, there exists a unique LD-system with domain $\{1, \dots, 2^n\}$ that satisfies* (I.16)*, namely the system S_{2^n} of 3.1.*

The combinatorial proof relies on an intermediate result, namely that S_N is an LD-system if and only if the equality $a{*}N = N$ is true for every a. It is not hard to see that this is impossible when N is not a power of 2. On the other hand, the verification of the property when N is a power of 2 relies on the following connection between S_N and $S_{N'}$ when N' is a multiple of N:

3.3 Lemma. (i) *Assume that S is an LD-system and $g_{[N'+1]} = g$ holds in S. Then mapping a to $g_{[a]}$ defines a homomorphism of $S_{N'}$ into S.*

(ii) *In particular, if S_N is an LD-system and N divides N', then mapping a to $a \bmod N$ defines a homomorphism of $S_{N'}$ onto S_N.*

(Here $a \bmod N$ denotes the unique integer equal to a *modulo* N lying in the interval $\{1, \dots, N\}$.)

3.4 Definition. For $n \geqslant 0$, the nth Laver table, denoted A_n, is defined to be the LD-system S_{2^n}, *i.e.*, the unique LD-system with domain $\{1, 2, \dots, 2^n\}$ that satisfies (I.16).

The first Laver tables are

A_0	1
1	1

A_1	1	2
1	2	2
2	1	2

A_2	1	2	3	4
1	2	4	2	4
2	3	4	3	4
3	4	4	4	4
4	1	2	3	4

A_3	1	2	3	4	5	6	7	8
1	2	4	6	8	2	4	6	8
2	3	4	7	8	3	4	7	8
3	4	8	4	8	4	8	4	8
4	5	6	7	8	5	6	7	8
5	6	8	6	8	6	8	6	8
6	7	8	7	8	7	8	7	8
7	8	8	8	8	8	8	8	8
8	1	2	3	4	5	6	7	8

The reader can compute that the first row in the table A_4 is $2, 12, 14, 16, 2, \ldots$ while that of A_5 is $2, 12, 14, 16, 28, 30, 32, 2, \ldots$

By 3.1, every row in the table A_n is periodic and it comes in the proof of 3.2 that the corresponding period is a power of 2. In the sequel, we write $o_n(a)$ for the number such that $2^{o_n(a)}$ is the period of a in A_n, *i.e.*, the number of distinct values in the row of a. The examples above show that the periods of 1 in $A_0, \ldots, A_5$ are $1, 1, 2, 4, 4$, and 8 respectively, corresponding to the equalities $o_0(1) = 0$, $o_1(1) = 0$, $o_2(1) = 1$, $o_3(1) = 2$, $o_4(1) = 2$, $o_5(1) = 3$. Observe that the above values are non-decreasing.

It is not hard to prove that, for each n, the unique generator of A_n is 1, its unique idempotent is 2^n, and we have $2^n *_n a = a$ and $a *_n 2^n = 2^n$ for every a.

An important point is the existence of a close connection between the tables A_n and A_{n+1} for every n (we write $*_n$ for the multiplication in A_n):

3.5 Lemma. (i) *For each n, the mapping $a \mapsto a \bmod 2^n$ is a surjective morphism of A_{n+1} onto A_n.*

(ii) *For every n, and every a with $1 \leqslant a \leqslant 2^n$, there exists a number $\theta_{n+1}(a)$ with $0 \leqslant \theta_{n+1}(a) \leqslant 2^{o_n(a)}$ and $\theta_{n+1}(2^n) = 0$ such that, for every b with $1 \leqslant b \leqslant 2^n$, we have*

$$a *_{n+1} b = a *_{n+1} (2^n + b) = \begin{cases} a *_n b & \text{for } b \leqslant \theta_{n+1}(a), \\ a *_n b + 2^n & \text{for } b > \theta_{n+1}(a), \end{cases}$$

$$(2^n + a) *_{n+1} b = (2^n + a) *_{n+1} (2^n + b) = a *_n b + 2^n.$$

For instance, the values of the mapping θ_4 are

a	1	2	3	4	5	6	7	8
$\theta_4(a)$	1	1	2	4	2	2	1	0

We obtain in this way a short description of A_n: the above 8 values contain all information needed for constructing the table of A_4 (16×16 elements) from that of A_3.

The LD-systems A_n play a fundamental role among finite LD-systems. In particular, it is shown in [12] how every LD-system with one generator can be obtained by various explicit operations (analogous to products) from a well-defined unique table A_n. Let us mention that as an LD-system A_n admits the presentation $\langle g \; ; \; g_{[m+1]} = g\rangle$ for every number m of the form $2^n(2p+1)$, and that the structure $(A_n, *)$ can be enriched with a second binary operation so as to become an LD-monoid:

3.6 Proposition. *There exists a unique associative product on A_n that turns $(A_n, *, \cdot)$ into an LD-monoid, namely the operation defined by*

$$a{\cdot}b = (a*(b+1)) - 1 \quad \text{for } b < 2^n, \quad a{\cdot}b = a \quad \text{for } b = 2^n. \tag{I.17}$$

3.2. Using elementary embeddings

In order to establish a connection between the tables A_n of the previous section and the finite quotients of $\mathrm{Iter}(j)$ described in Subsection 1.4, we shall use the following characterization:

3.7 Lemma. *Assume that S is an LD-system admitting a single generator g satisfying $g_{[2^n+1]} = g$ and $g_{[a]} \neq g$ for $a \leqslant 2^n$. Then S is isomorphic to A_n.*

Proof. Assume that S is an LD-system generated by an element g satisfying the above conditions. A double induction gives, for $a, b \leqslant 2^n$, the equality $g_{[a]} * g_{[b]} = g_{[a*b]}$, where $a*b$ refers to the product in A_n. So the set of all left powers of g is closed under product, and S, which has exactly 2^n elements, is isomorphic to A_n. ⊣

We immediately deduce from 1.22:

3.8 Proposition (Laver [19]). *For $j : V_\lambda \prec V_\lambda$, the quotient of* $\mathrm{Iter}(j)$ *under* $\mathrm{crit}_n(j)$*-equivalence is isomorphic to A_n.*

Under the previous isomorphism, the element a of A_n is the image of the class of the embedding $j_{[a]}$, and, in particular, 2^n is the image of the class of $j_{[2^n]}$, which is also the class of the identity map.

By construction, if S is an LD-system, and a is an element of S, there exists a well-defined evaluation for every term t in T_1 when the variable x is given the value a. We shall use $t(1)^{A_n}$, or simply $t(1)$, for the evaluation in A_n of a term $t(x)$ of T_1 at $x = 1$, and $t(j)$ for the evaluation of $t(x)$ in $\mathrm{Iter}(j)$ at $x = j$. With this notation, it should be clear that, for every term $t(x)$, the image of the $\mathrm{crit}_n(j)$-equivalence class of $t(j)$ in A_n under the isomorphism of 3.8 is $t(1)^{A_n}$.

The previous isomorphism can be used to obtain results about the iterations of an elementary embedding. For instance, let us consider the question of determining which left powers of j are $\mathrm{crit}_4(j)$-approximations of $j{\circ}j$ and of $j^{[3]}$. By looking at the table of the LD-monoid A_4, we obtain $A_4 \models 1{\circ}1 = 11$, and $A_4 \models 1^{[3]} = 12$. We deduce that $j{\circ}j$ is $\mathrm{crit}_4(j)$-equivalent to $j_{[11]}$ and $j^{[3]}$ is $\mathrm{crit}_n(j)$-equivalent to $j_{[12]}$.

The key to further results is the possibility of translating into the language of the finite tables A_n the values of the critical ordinals associated with the iterations of an elementary embedding.

3.9 Proposition. *Assume $j : V_\lambda \prec V_\lambda$. Then, for every term t and for $n \geqslant m \geqslant 0$ and $n \geqslant a \geqslant 1$,*

(i) $\mathrm{crit}(t(j)) \geqslant \mathrm{crit}_n(j)$ *is equivalent to* $A_n \models t(1) = 2^n$.

(ii) $\mathrm{crit}(t(j)) = \mathrm{crit}_n(j)$ *is equivalent to* $A_{n+1} \models t(1) = 2^n$.

(iii) $t(j)(\mathrm{crit}_m(j)) = \mathrm{crit}_n(j)$ *is equivalent to* $A_{n+1} \models t(1){*}2^m = 2^n$.

(iv) $j_{[a]}(\mathrm{crit}_m(j)) = \mathrm{crit}_n(j)$ *is equivalent to the period of a jumpimg from 2^m to 2^{m+1} between A_n and A_{n+1}.*

Proof. (i) By definition, $\mathrm{crit}(t(j)) \geqslant \mathrm{crit}_n(j)$ is equivalent to $t(j)$ being $\mathrm{crit}_n(j)$-equivalent to the identity mapping, hence to the image of $t(j)$ in A_n being the image of the identity, which is 2^n.

(ii) Assume $\mathrm{crit}(t(j)) = \mathrm{crit}_n(j)$. Then we have $\mathrm{crit}(t(j)) \geqslant \mathrm{crit}_n(j)$ and $\mathrm{crit}(t(j)) \not\geqslant \mathrm{crit}_{n+1}(j)$, so, by (i), $A_n \models t(1) = 2^n$ and $A_{n+1} \not\models t(1) = 2^{n+1}$. Now $A_n \models t(1) = 2^n$ implies $A_{n+1} \models t(1) = 2^n$ or 2^{n+1}, so 2^n is the only possible value here. Conversely, $A_{n+1} \models t(1) = 2^n$ implies $A_n \models t(1) = 2^n$ and $A_{n+1} \not\models t(1) = 2^{n+1}$, so, by (i), $\mathrm{crit}(t(j)) \geqslant \mathrm{crit}_n(j)$ and $\mathrm{crit}(t(j)) \not\geqslant \mathrm{crit}_{n+1}(j)$, hence $\mathrm{crit}(t(j)) = \mathrm{crit}_n(j)$.

(iii) As $\mathrm{crit}_m(j)$ is the critical ordinal of $j_{[2^m]}$, we have the equality $t(j)(\mathrm{crit}_m(j)) = \mathrm{crit}(t(j)[j_{[2^m]}])$. By (ii), $\mathrm{crit}(t(j)[j_{[2^m]}]) = \mathrm{crit}_n(j)$ is equivalent to $A_{n+1} \models t(1){*}1_{[2^m]} = 2^n$. Now we have $A_{n+1} \models 1_{[2^m]} = 2^m$ for $n \geqslant m$.

(iv) The image of $j_{[a]}$ is a both in A_n and A_{n+1}, hence (iii) tells us that $j_{[a]}(\mathrm{crit}_m(j)) = \mathrm{crit}_n(j)$ is equivalent to $A_{n+1} \models a{*}2^m = 2^n$. If the latter holds, the period p of a in A_{n+1} is 2^{m+1}: indeed, $A_{n+1} \models a{*}2^m < 2^{n+1}$ implies $p > 2^m$, while $2 \times 2^n = 2^{n+1}$ implies $p \leqslant 2 \times 2^m$. Conversely, assume that the period of a is 2^m in A_n and 2^{m+1} in A_{n+1}. We deduce $A_n \models a{*}2^m = 2^n$ and $A_{n+1} \not\models a{*}2^m = 2^{n+1}$, so the only possibility is $A_{n+1} \models a{*}2^m = 2^n$. ⊣

For instance, we can check $A_3 \models 1^{[3]} = 4$, and $A_5 \models 1^{[4]} = 16$. Using the dictionary, we deduce that the critical ordinal of $j^{[3]}$ is $\mathrm{crit}_2(j)$, while the critical ordinal of $j^{[4]}$ is $\mathrm{crit}_4(j)$. Also, we find $A_4 \models 4{*}4 = 8$, which implies that $j_{[4]}$ maps $\mathrm{crit}_2(j)$ to $\mathrm{crit}_3(j)$—as can be established directly. Similarly, we have $A_5 \models 1{*}4 = 16$, corresponding to $j(\mathrm{crit}_2(j)) = \mathrm{crit}_4(j)$. As for (iv), we see that the period of 1 jumps from 1 to 2 between A_1 and A_2, that it jumps from 2 to 4 between A_2 and A_3, and that it jumps from 4 to 8 between A_4 and A_5. We deduce that, if j is an elementary embedding of V_λ into itself, then j maps $\mathrm{crit}_0(j)$ to $\mathrm{crit}_1(j)$, $\mathrm{crit}_1(j)$ to $\mathrm{crit}_2(j)$, and $\mathrm{crit}_2(j)$

to $\mathrm{crit}_4(j)$, *i.e.*, we have $\kappa_2 = \gamma_4$ with the notations of Subsection 1.6. Similarly, the period of 3 jumps from 8 to 16 between A_5 and A_6: we deduce that $j_{[3]}$ maps $\mathrm{crit}_3(j)$ to $\mathrm{crit}_5(j)$.

By 3.9(iii): $\hat{\jmath}(m) = n$ is equivalent to $A_{n+1} \models 1{*}2^m = 2^n$. As the latter condition does not involve j, we deduce

3.10 Corollary. *For $j : V_\lambda \prec V_\lambda$, the mappings $\hat{\jmath}$ and $\tilde{\jmath}$ do not depend on j.*

In the previous examples, we used the connection between the iterates of an elementary embedding and the tables A_n to deduce information about elementary embeddings from explicit values in A_n. We can also use the correspondence in the other direction, and deduce results about the tables A_n from properties of the elementary embeddings.

Now, the existence of the function $\hat{\jmath}$ and, therefore, of its iterate $\tilde{\jmath}$, which we have seen is a direct consequence of the Laver–Steel theorem, translates into the following asymptotic result about the periods in the tables A_n. We recall that $o_n(a)$ denotes the integer such that the period of a in A_n is $2^{o_n(a)}$.

3.11 Proposition (Laver). *Assume Axiom (I3). Then, for every a, the period of a in A_n tends to infinity with n. More precisely, for $j : V_\lambda \prec V_\lambda$,*

$$o_n(a) \leqslant \tilde{\jmath}(r) \qquad \textit{if and only if} \qquad n \leqslant \tilde{\jmath}(r+1) \tag{I.18}$$

holds for $r \geqslant a$. In particular, (I.18) *holds for every r in the case $a = 1$.*

Proof. Assume first $a = 1$. Then j maps $\mathrm{crit}_{\tilde{\jmath}(r)}(j)$ to $\mathrm{crit}_{\tilde{\jmath}(r+1)}(j)$ for every r. Hence, by 3.9(iv), the period of 1 doubles from $2^{\tilde{\jmath}(r)}$ to $2^{\tilde{\jmath}(r)+1}$ between $A_{\tilde{\jmath}(r+1)}$ and $A_{\tilde{\jmath}(r+1)+1}$. So we have

$$o_{\tilde{\jmath}(r+1)}(1) = \tilde{\jmath}(r), \quad \text{and} \quad o_{\tilde{\jmath}(r+1)+1}(1) = \tilde{\jmath}(r) + 1,$$

which gives (I.18). Assume now $a \geqslant 2$. By 1.13, we have $(j_{[a]})^{[r]} = j^{[r+1]}$ for $r \geqslant a$, so the critical ordinal of $(j_{[a]})^{[r]}$ is $\mathrm{crit}_{\tilde{\jmath}(r)}(j)$. Hence, for $r \geqslant a$, the embedding $j_{[a]}$ maps $\mathrm{crit}_{\tilde{\jmath}(r)}(j)$ to $\mathrm{crit}_{\tilde{\jmath}(r+1)}(j)$, and the argument is as for $a = 1$. ⊣

We conclude with another result about the periods in the tables A_n.

3.12 Proposition (Laver). *Assume Axiom (I3). Then, for every n, the period of 2 in A_n is at least the period of 1.*

Proof. Assume that the period of 1 in A_n is 2^m. Let n' be the largest integer such that the period of 1 in $A_{n'}$ is 2^{m-1}. By construction, the period of 1 jumps from 2^{m-1} to 2^m between $A_{n'}$ and $A_{n'+1}$. Assume that j is a nontrivial elementary embedding of a rank into itself. By 3.9(iv), j maps $\mathrm{crit}_m(j)$ to $\mathrm{crit}_{n'}(j)$. Now, by 1.10, $j[j]$ maps $\mathrm{crit}_m(j)$ to some ordinal of the form $\mathrm{crit}_{n''}(j)$ with $n'' \leqslant n'$. This implies that the period of 2 jumps from 2^{m-1} to 2^m between $A_{n''}$ and $A_{n''+1}$. By construction, we have $n'' \leqslant n' < n$, hence the period of 2 in A_n is at least 2^m. ⊣

3.3. Avoiding elementary embeddings

Once again, the situation of 3.11 and 3.12 is strange, as it is not clear why any large cardinal hypothesis should be involved in the asymptotic behaviour of the periods in the finite LD-systems A_n. So we would either get rid of the large cardinal hypothesis, or prove that it is necessary.

We shall mention partial results in both directions. In the direction of eliminating the large cardinal assumption, *i.e.*, of getting arithmetic proofs, R. Dougherty and A. Drápal have proposed a scheme that essentially consists in computing the rows of (sufficiently many) elements $2^p - a$ in A_n using induction on a, which amounts to constructing convenient families of homomorphisms between the A_n's. Here we shall mention statements corresponding to the first two levels of the induction:

3.13 Theorem (Drápal [11]). (i) *For every d, and for $0 \leqslant m \leqslant 2^d + 1$, $b \mapsto 2^{2^d} b$ defines an injective homomorphism of A_m into A_{m+2^d}; it follows that, for $2^d \leqslant n \leqslant 2^{d+1} + 1$, the row of $2^{2^d} - 1$ in A_n is given by*

$$(2^{2^d} - 1) *_n b = 2^{2^d} b.$$

(ii) *For every d, and for $0 \leqslant m \leqslant 2^{2^{d+1}}$, the mapping f_d defined by $f_d : 2^i \mapsto 2^{(i+1)2^d} - 2^{i2^d}$ and $f_d(\sum b_i 2^i) = \sum b_i f_d(2^i)$ defines an injective homomorphism of A_m into A_{m2^d}; it follows that, for $0 \leqslant n \leqslant 2^{2^{d+1}+d}$ such that 2^d divides n, the row of $2^{2^d} - 2$ in A_n is given by*

$$(2^{2^d} - 2) *_n b = f_d(b).$$

So far, the steps $a \leqslant 4$ have been completed, but the complexity quickly increases, and whether the full proof can be completed remains open.

3.4. Not avoiding elementary embeddings?

We conclude with a result in the opposite direction:

3.14 Theorem (Dougherty–Jech [9]). *It is impossible to prove in PRA (Primitive Recursive Arithmetic) that the period of 1 in the table A_n goes to infinity with n.*

The idea is that enough of the computations of Subsection 1.6 can be performed in PRA to guarantee that, if the period of 1 in A_n tends to infinity with n, then some function growing faster than the Ackermann function provably exists.

Assume $j : V_\lambda \prec V_\lambda$. For every term t in T_1, the elementary embedding $t(j)$ acts on the family $\{\mathrm{crit}_n(j); n \in \omega\}$, and, as was done for j, we can associate with $t(j)$ an increasing injection $\widetilde{t(j)} : \omega \to \omega$ by

$$\widetilde{t(j)}(m) = n \quad \text{if and only if} \quad t(j)(\mathrm{crit}_m(j)) = \mathrm{crit}_n(j).$$

If t and t' are LD-equivalent terms, we have $t(j) = t'(j)$, hence $\widetilde{t(j)} = \widetilde{t'(j)}$, so, for a in the free LD-system F_1, we can define f^j_a to be the common value of $\widetilde{t(j)}$ for t representing a. We obtain in this way an F_1-indexed family of increasing injections of ω to itself, distinct from identity, and, by construction, the equality

$$\operatorname{crit}(f^j_{a*b}) = f^j_a(\operatorname{crit}(f^j_b)) \tag{I.19}$$

is satisfied for all a, b in F_1, where we define $\operatorname{crit}(f)$ to be the least m satisfying $f(m) > m$. The sequence $(f^j_a; a \in F_1)$ is the trace of the action of j on critical ordinals, and we shall see it captures enough of the combinatorics of elementary embeddings to deduce the results of Subsection 1.6.

Let us try to construct directly, without elementary embedding, some similar family of injections on ω satisfying (I.19). To this end, we can resort to the Laver tables. Indeed, by 3.9, the condition $t(j)(\operatorname{crit}_m(j)) = \operatorname{crit}_n(j)$ in the definition of $\widetilde{t(\jmath)}$ is equivalent to $A_{n+1} \models t(1)*2^m = 2^n$. So we are led to

3.15 Definition. (PRA) For a in F_1, we define f_a to be the partial mapping on ω such that $f_a(m) = n$ holds if, for some term t representing a, we have $A_{n+1} \models t(1)*2^m = 2^n$.

As A_{n+1} is an LD-system, the value of $t(1)*2^m$ computed in A_{n+1} depends on the LD-class of t only, so the previous definition is non-ambiguous. If there exists $j : V_\lambda \prec V_\lambda$, then, for each a in F_1, the mapping f_a coincides with f^j_a, and, therefore, each f_a is a total increasing injection of ω to ω, distinct from identity, and the f_a's satisfy the counterpart of (I.19). In particular, we can state

3.16 Proposition. (ZFC + I3) *For each a in F_1, the function f_a is total.*

Some of the previous results about the f_a's can be proved directly. Let us define a *partial increasing injection* on ω to be an increasing function of ω into itself whose domain is either ω, or a finite initial segment of ω. We shall say that a partial increasing injection f is *nontrivial* if $f(m) > m$ holds for at least one m, and that m is the *critical integer* of f, denoted $m = \operatorname{crit}(f)$, if we have $f(n) = n$ for $n < m$, and $f(m) \neq m$, *i.e.*, either $f(m) > m$ holds or $f(m)$ is not defined.

For f a partial increasing injection on ω, and m, n in ω, we write $f(m) \tilde{\geqslant} n$ if either $f(m)$ is defined and $f(m) \geqslant n$ holds, or $f(m)$ is not defined; we write $\operatorname{crit}(f) \tilde{\geqslant} m$ for $(\forall n < m)(f(n) = n)$. Then $f(m) = n$ is equivalent to the conjunction of $f(m) \tilde{\geqslant} n$ and $f(m) \tilde{\not\geqslant} n+1$, and $\operatorname{crit}(f) = m$ is equivalent to the conjunction of $\operatorname{crit}(f) \tilde{\geqslant} m$ and $\operatorname{crit}(f) \tilde{\not\geqslant} m+1$.

3.17 Lemma. (PRA) (i) *For every p, we have* $\operatorname{crit}(f_{x_{[2^p]}}) = p$.

(ii) *For t representing a, and for $n \geqslant m$, $f_a(m) \mathrel{\tilde{\geqslant}} n$ is equivalent to $A_n \models t(1){*}2^m = 2^n$;*

(iii) *The mapping f_a is a partial increasing injection;*

(iv) *The relation $\mathrm{crit}(f_a) \mathrel{\tilde{\geqslant}} n$ is equivalent to $A_n \models t(1) = 2^n$.*

(v) *If $\mathrm{crit}(f_b)$ and $f_a(\mathrm{crit}(f_b))$ are defined, so is $\mathrm{crit}(f_{a*b})$ and we have $\mathrm{crit}(f_{a*b}) = f_a(\mathrm{crit}(f_b))$.*

Proof. (i) First, $f_{x_{[2^p]}}(m) = m$ is equivalent to $A_{m+1} \models 1_{[2^p]}{*}2^m = 2^m$ by definition. This holds for $m < p$, as we have $A_{m+1} \models 1_{[2^p]} = 2^{m+1}$, and $A_{m+1} \models 2^{m+1}{*}x = x$ for every x. On the other hand, $A_{m+1} \models 1_{[2^p]} = 2^m$ holds, hence so does $A_{m+1} \models 1_{[2^p]}{*}2^m = 2^{m+1} \neq 2^m$. So $\mathrm{crit}(f_{x_{[2^p]}})$ exists, and it is p.

(ii) If $f_a(m) = p$ holds for some $p \geqslant n$, $A_{p+1} \models t(1){*}2^m = 2^{p+1}$, hence $A_n \models t(1){*}2^m = 2^n$ by projecting. And $f_a(m)$ not being defined means that there exists no p satisfying $A_{p+1} \models t(1){*}2^m < 2^{p+1}$: in other words $A_{p+1} \models t(1){*}2^m = 2^{p+1}$ for $p+1 \geqslant m$, and, in particular, for $p+1 = n$.

(iii) Assume $f_a(m+1) = n+1$. Then $A_{n+2} \models t(1){*}2^{m+1} = 2^{n+1}$ holds, *i.e.*, $t(1)$ has period 2^{m+2} at least in A_{n+2}. By projecting from A_{n+2} to A_{n+1}, we deduce that $t(1)$ has period 2^{m+1} at least in A_{n+1}, hence $A_{n+1} \models t(1){*}2^m \leqslant 2^n$. If the latter relation is an equality, we deduce $f_a(m) = n$. Otherwise, by projecting, we find some integer $p < n$ for which $A_{p+1} \models t(1){*}2^m = 2^p$, and we deduce $f_a(m) = p$. In both cases, $f_a(m)$ exists, and its value is at most n. This shows that the domain of f_a is an initial segment of ω, and that f_a is increasing.

(iv) Assume $\mathrm{crit}(f_a) \mathrel{\tilde{\geqslant}} n$, *i.e.*, $f_a(m) = m$ holds for $m < n$. We have $f_a(n-1) \not\mathrel{\tilde{\geqslant}} n$, hence $A_n \models t(1){*}2^{n-1} \leqslant 2^{n-1}$, whence $A_n \models t(1) = 2^n$, as $A_n \models a{*}2^{n-1} = 2^n$ holds for $a < 2^n$. Conversely, assume $A_n \models t(1) = 2^n$, and $m < n$. By projecting from A_n to A_{m+1}, we obtain $A_{m+1} \models t(1) = 2^{m+1}$, hence $A_{m+1} \models t(1){*}2^m = 2^m < 2^{m+1}$, which gives $f_a(m) \not\mathrel{\tilde{\geqslant}} m+1$ by (ii). As $f_a(m) \mathrel{\tilde{\geqslant}} m$ holds by (ii), we deduce $f_a(m) = m$.

(v) Let $a, b \in F_1$ be represented by t_1 and t_2 respectively. Assume first $f_a(p) \mathrel{\tilde{\geqslant}} n$ and $\mathrm{crit}(f_b) \mathrel{\tilde{\geqslant}} p$. By (iv), the hypotheses are $A_n \models t_1(1){*}2^p = 2^n$, and $A_p \models t_2(1) = 2^p$. By projecting from A_n to A_p, we deduce that $t_2(1)^{A_n}$ is a multiple of 2^p. Hence, the hypothesis $A_n \models t_1(1){*}2^p = 2^n$ implies $A_n \models (t_1{*}t_2)(1) = t_1(1){*}t_2(1) = 2^n$, hence, by (iv), $\mathrm{crit}(f_{a*n}) \mathrel{\tilde{\geqslant}} n$.

Assume now $f_a(p) \not\mathrel{\tilde{\geqslant}} n+1$ and $\mathrm{crit}(f_b) \not\mathrel{\tilde{\geqslant}} p+1$. The hypotheses are $A_{n+1} \models t_1(1){*}2^p \neq 2^{n+1}$, *i.e.*, the period of $t_1(1)$ in A_{n+1} is 2^{p+1} at least, and $A_{p+1} \models t_2(1) \neq 2^{p+1}$, hence $A_{p+1} \models t_2(1) \leqslant 2^p$. We cannot have $A_{n+1} \models t_2(1) \geqslant 2^{p+1}$ because, by projecting from A_{n+1} to A_{p+1}, we would deduce $A_{p+1} \models t_2(1) = 2^{p+1}$, contradicting our hypothesis. Hence we have $A_{n+1} \models t_2(1) \leqslant 2^p$, and the hypothesis that the period of $t_1(1)$ in A_{n+1} is 2^{p+1} at least implies $A_{n+1} \models t_1(1){*}t_2(1) \leqslant 2^n$, hence $\mathrm{crit}(f_{a*b}) \not\mathrel{\tilde{\geqslant}} n+1$. So the conjunction of $f_a(p) = n$ and $\mathrm{crit}(f_b) = p$ implies $\mathrm{crit}(f_{a*b}) = n$. $\dashv$

The only point we have not proved so far is that the function f_a be total. Before going further, let us observe that the latter property is connected with the asymptotic behaviour of the periods in the tables A_n, as well as with several equivalent statements:

3.18 Proposition. (PRA) *The following statements are equivalent:*

(i) *For each a in F_1, the function f_a is total;*

(ii) *For every term t, the period of $t(1)$ in A_n goes to infinity with n—so, in particular, the period of every fixed a in A_n goes to infinity with n;*

(iii) *The period of 1 in A_n goes to infinity with n;*

(iv) *For every r, there exists an n satisfying $A_n \models 1^{[r]} < 2^n$;*

(v) *The subsystem of the inverse limit of all A_n's generated by $(1, 1, \dots)$ is free.*

Proof. Let t be an arbitrary term in T_1, and a be its class in F_1. Saying that the period of $t(1)$ in A_n goes to ∞ with n means that, for every m, there exists n with $A_n \models t(1){*}2^m < 2^n$, *i.e.*, $f_a(m) \tilde{\ngeqslant} n$. If the function f_a is total, such an n certainly exists, so (i) implies (ii). Conversely, if (ii) is satisfied, the existence of n satisfying $f_a(m) \tilde{\ngeqslant} n$ implies that $f_a(m)$ is defined, so (i) and (ii) are equivalent, and they imply (iii), which is the special case $t = x$ of (ii).

Assume now (iii). By the previous argument, the mapping f_x is total. If f_a and f_b are total, then, by 3.17(v), $\mathrm{crit}(f_{b^{[n]}})$ exists for every n, and so does $f_a(\mathrm{crit}(f_{b^{[n]}}))$, which is $\mathrm{crit}(f_{(a*b)^{[n]}})$. This proves that $f_{a*b}(m)$ exists for arbitrary large values of m, and this is enough to conclude that f_{a*b} is total. So, inductively, we deduce that f_a is total for every a, which is (i).

Then, we prove that (ii) implies (iv) using induction on $r \geqslant 1$. The result is obvious for $r = 1$. Let p be maximal satisfying $A_p \models 1^{[r-1]} = 2^p$, which exists by induction hypothesis. By (ii), we have $A_n \models 1{*}2^p < 2^n$ for some $n > p$, so the period of 1 in A_n is a multiple of 2^{p+1}. By hypothesis, we have $A_{p+1} \models 1^{[r-1]} = 2^p$, hence $A_n \models 1^{[r-1]} = 2^p \mod 2^{p+1}$, so 2^p is the largest power of 2 that divides $1^{[r-1]}$ computed in A_n. As the period of 1 in A_n is a multiple of 2^{p+1}, we obtain $A_n \not\models 1{*}1^{[r-1]} = 2^n$, so $A_n \models 1^{[r]} = 1{*}1^{[r-1]} < 2^n$.

Assume now (iv), and let t be an arbitrary term. By 2.6, there exist q, r satisfying $t^{[r]} =_{LD} x^{[q]}$. By (iv), $A_n \models 1^{[q]} = t(1)^{[r]} < 2^n$ for some n, hence $A_n \models t(1) < 2^n$, since every right power of 2^n in A_n is 2^n. Hence (iv) implies (iii).

Assume (i), and let $t, t_1, \dots, t_p$ be arbitrary terms. By (ii), we can find n such that none of the terms t, $t{*}t_1$, $(t{*}t_1){*}t_2$, ..., $(\dots(t{*}t_1)\dots){*}t_p$ evaluated at 1 in A_n is 2^n: this is possible since $A_n \models t(1) \neq 2^n$ implies $A_m \models t(1) \neq 2^m$ for $m \geqslant n$. So we have

$$A_n \models t(1) < (t{*}t_1)(1) < ((t{*}t_1){*}t_2)(1) < \dots$$

and, in particular, $A_n \models t(1) \neq (\ldots(t{*}t_1){*})\ldots{*}t_p)(1)$. This implies that left division in the sub-LD-system of the inverse limit of all A_n's generated by $(1, 1, \ldots)$ has no cycle, and, therefore, by Laver's criterion, this LD-system is free. Conversely, assume that (i) fails, *i.e.*, there exists $p \geqslant 1$ such that $A_n \models 1{*}2^p = 2^n$ for every n. Let α denote the sequence $(1, 1, \ldots)$ in the inverse limit. Then we have $\alpha_{[2^p]} = (1, 2, \ldots, 2^p, 2^p, \ldots)$ and

$$\alpha{*}\alpha_{[2^p+1]} = (\alpha{*}\alpha_{[2^p]}){*}(\alpha{*}\alpha) = \alpha{*}\alpha.$$

The sub-LD-system generated by α cannot be free, since $g{*}g = g{*}g_{[2^p+1]}$ does not hold in the free LD-system generated by g. So (v) is equivalent to (i)–(iv). ⊣

The status of the equivalent statements of 3.18 remains currently open. However, the results of Subsection 1.6 enables us to say more. We have seen that the function $\tilde{\jmath}$ associated with an elementary embedding j grows faster than any primitive recursive function. In terms of the functions f_a^j, we have $\tilde{\jmath}(n) = (f_x^j)^n(0)$. As the functions f_a^j and f_a coincide when the former exist, it is natural to look at the values $f_x^n(0)$. The point is that we can obtain for this function the same lower bound as for its counterpart f_x^j without using any set theoretical hypothesis:

3.19 Proposition. (PA) *Assume that, for each a, the function f_a is total. Then the function $n \mapsto f_a^n(0)$ grows faster than any primitive recursive function.*

Proof. We consider the proof of 1.32, and try to mimick it using f_a and critical integers instead of f_a^j and critical ordinals. This is possible, because the only properties used in Subsection 1.6 are the left self-distributivity law and Relation (I.19) about critical ordinals. First, the counterpart of 1.33 is true since every value of f_a is an increasing injection and its domain is an initial interval of ω. Then the definitions of a base and of a realizable sequence can be translated without any change. Let us consider 1.36. With our current notation, the point is to be able to deduce from the hypothesis

$$f_b : \Vdash\!\!\to m_0 \mapsto m_1 \mapsto \ldots \mapsto m_p \tag{I.20}$$

the conclusion

$$f_{a*b} : \Vdash\!\!\to f_a(m_0) \mapsto f_a(m_1) \mapsto \ldots \mapsto f_a(m_p). \tag{I.21}$$

An easy induction on r gives the equality $(f_a)^n(\mathrm{crit}(f_a)) = \mathrm{crit}(f_{a^{[n+1]}})$. Now (I.20) can be restated as

$$\mathrm{crit}(f_b) = m_0, \quad \mathrm{crit}(f_{b^{[2]}}) = m_1, \ \ldots, \quad \mathrm{crit}(f_{b^{[n+1]}}) = m_n.$$

By applying f_a and using 3.17(v), we obtain

$$\mathrm{crit}(f_{a*b}) = f_a(m_0), \;\; \ldots, \quad \mathrm{crit}(f_{a*b^{[n+1]}}) = f_a(m_n).$$

By (LD), we have $f_{a*b^{[n]}} = f_{(a*b)^{[n]}}$, and therefore (I.20) implies (I.21).

So the proof of 1.36 goes through in the framework of the f_a's, and so do those of the other results of Subsection 1.6. We deduce that, for $n > 3$, there are at least $2^{h_1(h_2(\ldots(h_{n-2}(1))\ldots))}$ critical integers below the number $f_x^n(0)$, where h_p are the fast growing function of Subsection 1.6, and, finally, we conclude that the function $n \mapsto f_x^n(0)$ grows at least as fast as the Ackermann function. ⊣

It is then easy to complete the proof of 3.14:

Proof. By 3.18, proving that the period of 1 in A_n goes to infinity with n is equivalent (in PRA) to proving that the functions f_a are total. By 3.19, such a proof would also give a proof of the existence of a function growing faster than the Ackermann function. The latter function is not primitive recursive, and, therefore, such a proof cannot exist in PRA. ⊣

As the gap between PRA and (I3) is large, there remains space for many developments here.

To conclude, let us observe that, in the proof of 3.19, the hypothesis that the injections are total is not really used. Indeed, we establish lower bounds for the values, and the precise result is an alternative: for each r, either the value of $f_x^n(0)$ is not defined, or this value is at least some explicit value. In particular, the result is local, and the lower bounds remain valid for small values of r even if $f_x^n(0)$ is not defined for some large n. So, for instance, we have seen in Subsection 1.6 that, for $j : V_\lambda \prec V_\lambda$, we have $\tilde{\jmath}(4) \geqslant 256$, which, when translated into the language of A_n, means that the period of 1 in A_n is 16 for every n between 9 and 256 at least. The above argument shows that this lower bound remains valid even if Axiom (I3) is not assumed. The same result is true with the stronger inequality of 1.40, so we obtain

3.20 Theorem (Dougherty). *If it exists, the first integer n such that the period of 1 in A_n reaches 32 is at least $f_9^{\mathrm{Ack}}(f_8^{\mathrm{Ack}}(f_8^{\mathrm{Ack}}(254)))$.*

We refer to [8, 10, 13] (and to unpublished work by Laver) for many more computations about the critical ordinals of iterated elementary embeddings.

Bibliography

[1] Serge Burckel. The wellordering on positive braids. *J. Pure Appl. Algebra*, 120(1):1–17, 1997.

[2] Jim Cannon, William Floyd, and Walter Parry. Introductory notes on Richard Thompson's groups; *Ens. Math.*, 42:215–257, 1996.

[3] Patrick Dehornoy. Π^1_1-complete families of elementary sequences. *Ann. P. Appl. Logic*, 38:257–287, 1988.

[4] Patrick Dehornoy. Sur la structure des gerbes libres. *C. R. Acad. Sci. Paris*, 309:143–148, 1989.

[5] Patrick Dehornoy. Braid groups and self-distributive operations. *Trans. Amer. Math. Soc.*, 345(1):115–151, 1994.

[6] Patrick Dehornoy, Ivan Dynnikov, Dale Rolfsen, and Bert Wiest. *Why are braids orderable?*, volume 14 of *Panoramas et Synthèses*. Soc. Math. de France, Paris, 2002.

[7] Randall Dougherty. Critical points in an algebra of elementary embeddings. *Ann. P. Appl. Logic*, 65:211–241, 1993.

[8] Randall Dougherty. Critical points in an algebra of elementary embeddings (ii). In W. Hodges, editor, *Logic: From Foundations to Applications*, pages 103–136. Oxford Acad. Press, 1996.

[9] Randall Dougherty and Thomas Jech. Finite left-distributive algebras and embedding algebras. *Advances in Math.*, 130:201–241, 1997.

[10] Randall Dougherty and Thomas Jech. Left-distributive embedding algebras. *Electr. Res. Announc. Amer. Math. Soc.*, 3:23–37, 1997.

[11] Aleš Drápal. Persistence of left distributive algebras. *J. Pure Appl. Algebra*, 105:137–165, 1995.

[12] Aleš Drápal. Finite left distributive groupoids with one generator. *Int. J. for Algebra Comput.*, 7(6):723–748, 1997.

[13] Thomas Jech. Large ordinals. *Advances in Math.*, 125:155–170, 1997.

[14] Akihiro Kanamori. *The higher infinite*. Springer-Verlag, 1994.

[15] Kenneth Kunen. Elementary embeddings and infinitary combinatorics. *J. Symb. Logic*, 36:407–413, 1971.

[16] David Larue. On braid words and irreflexivity. *Algebra Univ.*, 31:104–112, 1994.

[17] Richard Laver. Elementary embeddings of a rank into itself. *Abstracts Amer. Math. Soc.*, 7:6, 1986.

[18] Richard Laver. The left distributive law and the freeness of an algebra of elementary embeddings. *Advances in Math.*, 91(2):209–231, 1992.

[19] Richard Laver. On the algebra of elementary embeddings of a rank into itself. *Advances in Math.*, 110:334–346, 1995.

[20] Richard Laver. Braid group actions on left distributive structures and well-orderings in the braid group. *J. Pure Appl. Algebra*, 108(1):81–98, 1996.

[21] Robert Solovay, William Reinhardt, and Akihiro Kanamori. Strong axioms of infinity and elementary embeddings. *Ann. Math. Logic*, 13:73–116, 1978.

[22] John Steel. The well foundedness of the Mitchell ordering. *J. Symbolic Logic*, 58:931–940, 1993.