Product-free Lambek calculus and context-free grammars

Mati Pentus*

Department of Mathematical Logic,
Faculty of Mechanics and Mathematics,
Moscow University,
119899, Moscow, Russia

December 29, 1993; December 12, 1995

Abstract

In this paper we prove the Chomsky Conjecture (all languages recognized by the Lambek calculus are context-free) for both the full Lambek calculus and its product-free fragment. For the latter case we present a construction of context-free grammars involving only product-free types.

Introduction

The notion of a basic categorial grammar was introduced in [1]. In the same paper it was proved that the languages recognized by basic categorial grammars are precisely the context-free ones.

Another kind of categorial grammar was introduced by J. Lambek [8]. These grammars are based on a syntactic calculus, presently known as the Lambek calculus (cf. [13] for its semantic interpretations). Chomsky [6] conjectured that these grammars are also equivalent to context-free ones. In [7] Cohen proved that every basic categorial grammar (and, thus, every context-free grammar) is equivalent to a Lambek grammar. He also proposed a proof of the converse. However, as pointed out in [2], this proof contains an error. Buszkowski proved that some special kinds of Lambek grammars are context-free [2, 3, 4]. These grammars use weakly unidirectional types or types of order at most two.

The main result of this paper (Theorem 2) says that Lambek grammars generate only context-free languages. Thus they are equivalent to context-free grammars and also to

*This research was supported in part by the International Science Foundation under Grant No. NFQ300.

basic categorial grammars. This fact (sometimes called *the Chomsky Conjecture*) was proved in [10] and [11]. Here we present an improved version of the proof. Some details of this improvement were independently obtained by W. Buszkowski [5].

1 Preliminaries

For any set $\mathcal{M}$ we denote by $\mathcal{M}^*$ (resp. $\mathcal{M}^+$) the set of all finite (resp. finite non-empty) strings consisting of elements of $\mathcal{M}$. The set of all subsets of $\mathcal{M}$ will be denoted by $\mathrm{P}(\mathcal{M})$.

1.1 Lambek calculus

We consider the syntactic calculus introduced in [8]. The types of the Lambek calculus L are built of primitive types $p_1, p_2, \ldots$ and three binary connectives $\bullet, \backslash, /$. We shall denote the set of all types by Tp and the set of all types that do not contain $\bullet$ (product-free types) by $\mathrm{Tp}(\backslash, /)$.

Capital letters A, B, ... range over types. Capital Greek letters range over finite (possibly empty) sequences of types. The empty sequence is denoted by $\emptyset$.

Sequents of the Lambek calculus are of the form $\Gamma \to A$, where A is a type and Γ is a *non-empty* sequence of types.

Axioms: $p_i \to p_i$

Rules:

$$\frac{\Gamma \to A \quad \Delta \to B}{\Gamma\Delta \to A{\bullet}B}\ (\to \bullet) \qquad\qquad \frac{\Gamma A B \Delta \to C}{\Gamma(A{\bullet}B)\Delta \to C}\ (\bullet \to)$$

$$\frac{A\Pi \to B}{\Pi \to A\backslash B}\ (\to \backslash) \text{ where } \Pi \neq \emptyset \qquad\qquad \frac{\Pi \to A \quad \Gamma B \Delta \to C}{\Gamma\Pi(A\backslash B)\Delta \to C}\ (\backslash \to)$$

$$\frac{\Pi A \to B}{\Pi \to B/A}\ (\to /) \text{ where } \Pi \neq \emptyset \qquad\qquad \frac{\Pi \to A \quad \Gamma B \Delta \to C}{\Gamma(B/A)\Pi\Delta \to C}\ (/ \to)$$

$$\frac{\Pi \to B \quad \Gamma B \Delta \to A}{\Gamma\Pi\Delta \to A}\ (CUT)$$

We write $L \vdash \Gamma \to A$ if the sequent $\Gamma \to A$ is derivable in the Lambek calculus.

The cut-elimination theorem for this calculus is proved in [8].

Definition. The *length* $\|A\|$ of a type A is defined as the total number of primitive type occurreces in A.

$$\|p_i\| \rightleftharpoons 1 \qquad \|A{\bullet}B\| = \|A\backslash B\| = \|A/B\| \rightleftharpoons \|A\| + \|B\|$$

The length of a sequence of types is defined in the natural way.

$$\|A_1 \ldots A_n\| \rightleftharpoons \|A_1\| + \ldots + \|A_n\|$$

Definition. We denote the set of all primitive types occurring in a type A by $\mathrm{Var}(A)$.

For any two natural numbers m and q we introduce a finite set of types $\mathrm{Tp}(m,q)$ and a finite set of sequences of types $\mathrm{Ls}(m,q)$.

$$\begin{aligned}\mathrm{Tp}(m,q) &\rightleftharpoons \{A \in \mathrm{Tp}(\backslash,/) \mid \mathrm{Var}(A) \subseteq \{p_1,\ldots,p_q\} \text{ and } \|A\| \leq m\}\\ \mathrm{Ls}(m,q) &\rightleftharpoons \{\Pi \in \mathrm{Tp}(m,q)^+ \mid \|\Pi\| \leq 2m\}\end{aligned}$$

1.2 Lambek grammars and context-free grammars

Definition. A *Lambek grammar* is a triplet $\langle \mathcal{T}, D, f\rangle$, where $\mathcal{T}$ is a finite set (the alphabet), $D \in \mathrm{Tp}(\backslash,/)$, and f is a function $f\colon \mathcal{T} \to \mathrm{P}(\mathrm{Tp}(\backslash,/))$ such that for any $t \in \mathcal{T}$ the set $f(t)$ is finite.

The *language generated by the Lambek grammar* $\langle \mathcal{T}, D, f\rangle$ is defined as the set of all strings $t_1 \ldots t_n$ over the alphabet $\mathcal{T}$ for which there exists a derivable sequent $B_1 \ldots B_n \to D$ such that $B_i \in f(t_i)$ for all $i \leq n$. We shall denote this language by $\mathcal{L}(\mathcal{T}, D, f)$.

Definition. A *context-free grammar* is a quadruple $\langle \mathcal{T}, \mathcal{W}, S, \mathcal{R}\rangle$, where $\mathcal{T}$ and $\mathcal{W}$ are two disjoint finite sets (the alphabet of *terminal symbols* and the set of *non-terminal symbols*), $S \in \mathcal{W}$, and $\mathcal{R}$ is a finite set of *context-free rewrite rules* of the form $X \Rightarrow e$, where $X \in \mathcal{W}$ and $e \in (\mathcal{T} \cup \mathcal{W})^+$.

By $\bar{\mathcal{G}}(\mathcal{T}, \mathcal{W}, S, \mathcal{R})$ we denote the set of all expressions over the alphabet $\mathcal{T} \cup \mathcal{W}$ that arise through some finite sequence of rewritings of the start symbol S via the rules of $\mathcal{R}$.

The *language generated by the context-free grammar* $\langle \mathcal{T}, \mathcal{W}, S, \mathcal{R}\rangle$ is defined as $\mathcal{G}(\mathcal{T}, \mathcal{W}, S, \mathcal{R}) \rightleftharpoons \bar{\mathcal{G}}(\mathcal{T}, \mathcal{W}, S, \mathcal{R}) \cap \mathcal{T}^+$.

2 Construction of the context-free grammar corresponding to a given Lambek grammar

Our main aim is to prove that for any Lambek grammar there exists a context-free grammar such that the languages generated by these grammars coincide. Here we deduce this result from the hypothesis that every sequent $\Gamma \to A$ derivable in the Lambek calculus can be derived from some short derivable sequents by means of the cut rule only. The hypothesis will be proved later.

In order to formalize the notion of derivability by means of the cut rule only, we introduce for every pair of positive integers m and q a calculus $Lcut(m,q)$.

Definition. A sequent $\Gamma \to A$ is an axiom of $Lcut(m,q)$ iff $A \in \mathrm{Tp}(m,q)$, $\Gamma \in \mathrm{Ls}(m,q)$, and the sequent $\Gamma \to A$ is derivable in the Lambek calculus. The only rule of $Lcut(m,q)$ is (CUT).

Theorem 1 *Let $A_1,\ldots,A_n, B \in \mathrm{Tp}(m,q)$.*
If $L \vdash A_1 \ldots A_n \to B$ then $Lcut(m,q) \vdash A_1 \ldots A_n \to B$.

Theorem 1 will be proved in Section 6.2.

Theorem 2 *For any Lambek grammar there exists a context-free grammar such that the languages generated by these grammars coincide.*

Proof. Take an arbitrary Lambek grammar $\langle \mathcal{T}, D, f\rangle$. Evidently there are positive integers m and q such that $D \in \mathrm{Tp}(m,q)$ and for any $t \in \mathcal{T}$, $f(t) \subset \mathrm{Tp}(m,q)$ (since only a finite number of types are relevant in the definition of a language generated by a Lambek grammar).

Assume for convenience that $\mathcal{T}$ and $\mathrm{Tp}(m,q)$ do not intersect. Now we construct the desired context-free grammar $\langle \mathcal{T}, \mathcal{W}, S, \mathcal{R}\rangle$.

$$
\begin{aligned}
\mathcal{W} &\rightleftharpoons \mathrm{Tp}(m,q) \\
S &\rightleftharpoons D \\
\mathcal{R} &\rightleftharpoons \{B \Rightarrow t \mid t \in \mathcal{T} \text{ and } B \in f(t)\} \\
&\qquad \cup \{A \Rightarrow \Gamma \mid A \in \mathrm{Tp}(m,q), \Gamma \in \mathrm{Ls}(m,q), \text{ and } L \vdash \Gamma \to A\}
\end{aligned}
$$

The set $\mathcal{R}$ consists of obvious rules describing the function f and of $Lcut(m,q)$-axioms with their sequent arrows reversed.

We must prove that $\mathcal{L}(\mathcal{T}, D, f) = \mathcal{G}(\mathcal{T}, \mathcal{W}, S, \mathcal{R})$. First we establish $\mathcal{L}(\mathcal{T}, D, f) \subseteq \mathcal{G}(\mathcal{T}, \mathcal{W}, S, \mathcal{R})$.

Suppose that $t_1 \dots t_n \in \mathcal{L}(\mathcal{T}, D, f)$. By the definition of $\mathcal{L}(\mathcal{T}, D, f)$, there are types $B_1 \dots B_n$ such that $L \vdash B_1 \dots B_n \to D$ and $B_i \in f(t_i)$ for any $i \leq n$. Thus $(B_i \Rightarrow t_i) \in \mathcal{R}$ for any $i \leq n$. Therefore it suffices to prove that $B_1 \dots B_n \in \bar{\mathcal{G}}(\mathcal{T}, \mathcal{W}, S, \mathcal{R})$. By Theorem 1 $Lcut(m,q) \vdash B_1 \dots B_n \to D$.

Lemma 2.1 *If $Lcut(m,q) \vdash \Theta \to D$ then there is a $Lcut(m,q)$-derivation of $\Theta \to D$ using only cuts whose left premise is an axiom.*

Proof. Easy induction involving the following reduction.

$$
\dfrac{\dfrac{\Pi \to B \quad \Gamma B \Delta \to A}{\Gamma \Pi \Delta \to A} \qquad \Phi A \Psi \to D}{\Phi \Gamma \Pi \Delta \Psi \to D}
$$

$$\Downarrow$$

$$
\dfrac{\Pi \to B \qquad \dfrac{\Gamma B \Delta \to A \quad \Phi A \Psi \to D}{\Phi \Gamma B \Delta \Psi \to D}}{\Phi \Gamma \Pi \Delta \Psi \to D}
$$

■

It is easy to see that if there is a $Lcut(m,q)$-derivation of $\Theta \to D$ using only cuts whose left premise is an axiom, then $\Theta \in \bar{\mathcal{G}}(\mathcal{T}, \mathcal{W}, S, \mathcal{R})$. This completes the proof of $\mathcal{L}(\mathcal{T}, D, f) \subseteq \mathcal{G}(\mathcal{T}, \mathcal{W}, S, \mathcal{R})$.

Now we verify that $\mathcal{G}(\mathcal{T}, \mathcal{W}, S, \mathcal{R}) \subseteq \mathcal{L}(\mathcal{T}, D, f)$. The following lemma is well-known.

Lemma 2.2 *Let $\langle \mathcal{T}, \mathcal{W}, S, \mathcal{R}\rangle$ be a context-free grammar. Let $\mathcal{R} = \mathcal{R}_1 \cup \mathcal{R}_2$, where $\mathcal{R}_1 \subseteq \{B \Rightarrow t \mid B \in \mathcal{W}$ and $t \in \mathcal{T}\}$ and $\mathcal{R}_2 \subseteq \{A \Rightarrow \Gamma \mid A \in \mathcal{W}$ and $\Gamma \in \mathcal{W}^+\}$. If $t_1 \dots t_n \in \mathcal{G}(\mathcal{T}, \mathcal{W}, S, \mathcal{R})$, then there exists a sequence $B_1 \dots B_n \in \mathcal{W}^+$ such that*

(i) *the word $B_1 \dots B_n$ can be obtained from S using only rules from $\mathcal{R}_2$;*

(ii) *for each $i \leq n$ the set $\mathcal{R}_1$ contains the rule $B_i \Rightarrow t_i$.*

PROOF. Induction on the number of rewritings. ■

This completes the proof of Theorem 2. ■

3 Thin sequents

In this section we introduce the notion of "thin" sequents and reduce Theorem 1 to the "thin" case.

3.1 Definitions

Definition. For each natural number i we define a function σ_i: Tp $\to$ N as follows. (Here N stands for the set of all natural numbers.)

$$\sigma_i p_i \rightleftharpoons 1 \qquad \sigma_i p_j \rightleftharpoons 0, \text{ if } i \neq j$$
$$\sigma_i(A{\bullet}B) = \sigma_i(A\backslash B) = \sigma_i(A/B) \rightleftharpoons \sigma_i A + \sigma_i B$$

In other words, $\sigma_i(A)$ counts occurrences of the primitive type p_i in A. We extend this definition to finite sequences of types.

$$\sigma_i(A_1 \dots A_n) \rightleftharpoons \sigma_i A_1 + \dots + \sigma_i A_n$$

We make two observations concerning σ_i.

Lemma 3 *If a sequent $\Pi \to A$ is derivable in the Lambek calculus, then for any i, $\sigma_i(\Pi A)$ is an even number.*

PROOF. Straightforward induction on derivations. ■

Lemma 4 *For any type A, $\|A\| = \sum\limits_i \sigma_i A$.*

Definition. A sequent $\Pi \to A$ is *thin* iff $\sigma_i(\Pi A) \leq 2$ for any i (i.e., no primitive type occurs in the sequent more than twice).

3.2 Reduction to the case of thin sequents

Lemma 5 *If Theorem 1 holds for thin sequents, then it holds for all sequents of the Lambek calculus.*

Proof. Assume that Theorem 1 holds for thin sequents. We prove that for any types $B_1, \ldots, B_n, D \in \mathrm{Tp}(m, q)$, if $L \vdash B_1 \ldots B_n \to D$, then $Lcut(m, q) \vdash B_1 \ldots B_n \to D$.

The sequent $B_1 \ldots B_n \to D$ may have several derivations in the Lambek calculus. We consider only one of these derivations and introduce a new primitive type for each axiom instance in this derivation. Let these new primitive types be p_{q+1}, p_{q+2}, $\ldots p_{q+k}$, where k is the number of axiom instances in the derivation (i.e., $k = \frac{1}{2}\|B_1 \ldots B_n D\|$).

We define a function $\phi\colon \{p_{q+1}, \ldots, p_{q+k}\} \to \{p_1, \ldots, p_q\}$ associating with p_{q+i} the old primitive type that occurs in the i-th axiom instance in the derivation of $B_1 \ldots B_n \to D$. We extend the definition to complex types and sequents stipulating

$$\begin{aligned} \phi(E\backslash F) &\rightleftharpoons \phi(E)\backslash\phi(F) \\ \phi(E/F) &\rightleftharpoons \phi(E)/\phi(F) \\ \phi(E_1 \ldots E_m \to F) &\rightleftharpoons \phi(E_1) \ldots \phi(E_m) \to \phi(F) \end{aligned}$$

for any types E, E_j, F.

Now we are going to replace everywhere in the derivation of $B_1 \ldots B_n \to D$ old primitive types by new ones. In axioms this is done in the obvious way (using the new primitive type that corresponds to a given axiom instance as a substitute for the only primitive type occurring in this axiom instance). Conclusions of rules inherit the replacement from premises. (In the Lambek calculus, as well as in the multiplicative fragment of the linear logic, every primitive type in the conclusion of a rule has a unique prototype in one of the premises of the rule.)

Thus we obtain a derivation of a sequent $A_1 \ldots A_n \to C$ such that $\phi(A_1 \ldots A_n \to C) = B_1 \ldots B_n \to D$. Note that $A_1, \ldots, A_n, C \in \mathrm{Tp}(m, q{+}k)$ and the sequent $A_1 \ldots A_n \to C$ is thin.

By our assumption, $Lcut(m, q{+}k) \vdash A_1 \ldots A_n \to C$. It remains to replace in this $Lcut(m, q{+}k)$-derivation every sequent $\Pi \to E$ by its image $\phi(\Pi \to E)$. We obtain the desired $Lcut(m, q)$-derivation of $B_1 \ldots B_n \to D$. ■

4 Free group interpretation

Let FG stand for the free group generated by the enumerable set of all primitive types $\{p_1, p_2, p_3, \ldots\}$. The identity element will be denoted by ε. For any element $u \in FG$, we define $|u|$ as the length of u written as a reduced word, i.e., a word that does not contain any subwords of the form $p_i p_i^{-1}$ or $p_i^{-1} p_i$.

Definition. The *free group interpretation* (written as $[\![\ ]\!]$) is the following mapping of types and finite sequences of types into FG.

$$[\![p_i]\!] \rightleftharpoons p_i$$

$$\begin{array}{rcl} [\![A{\bullet}B]\!] & \rightleftharpoons & [\![A]\!][\![B]\!] \\ [\![A\backslash B]\!] & \rightleftharpoons & [\![A]\!]^{-1}[\![B]\!] \\ [\![A/B]\!] & \rightleftharpoons & [\![A]\!][\![B]\!]^{-1} \\ [\![A_1 \ldots A_n]\!] & \rightleftharpoons & [\![A_1]\!] \ldots [\![A_n]\!] \end{array}$$

Remark. For any type A, $|[\![A]\!]| \leq \|A\|$.

Lemma 6 *If a sequent $\Gamma \to C$ is derivable in the Lambek calculus, then $[\![\Gamma]\!] = [\![C]\!]$.*

D. Roorda obtained this result in terms of atomic markings and balance. The lemma has also an immediate proof in the free group environment [9].

PROOF. Induction on derivations.

CASE 1: Axiom.

Trivial.

CASE 2: $(\to \bullet)$

$$\frac{\Gamma \to A \quad \Delta \to B}{\Gamma\Delta \to A{\bullet}B}\,(\to \bullet)$$

By the induction hypothesis $[\![\Gamma]\!] = [\![A]\!]$ and $[\![\Delta]\!] = [\![B]\!]$. Consequently $[\![\Gamma\Delta]\!] = [\![A]\!][\![B]\!] = [\![A{\bullet}B]\!]$.

CASE 3: $(\bullet \to)$

$$\frac{\Gamma A B \Delta \to C}{\Gamma(A{\bullet}B)\Delta \to C}\,(\bullet \to)$$

Obvious.

CASE 4: $(\to \backslash)$

$$\frac{A\Pi \to B}{\Pi \to A\backslash B}\,(\to \backslash)$$

Multiplying the equality $[\![A]\!][\![\Pi]\!] = [\![B]\!]$ by $[\![A]\!]^{-1}$ on the left, one obtains $[\![\Pi]\!] = [\![A]\!]^{-1}[\![B]\!]$. Thus $[\![\Pi]\!] = [\![A\backslash B]\!]$.

CASE 5: $(\to /)$

Similar to the previous case.

CASE 6: $(\backslash \to)$

$$\frac{\Pi \to A \quad \Gamma B \Delta \to C}{\Gamma\Pi(A\backslash B)\Delta \to C}\,(\backslash \to)$$

If $[\![\Pi]\!] = [\![A]\!]$ then $[\![\Pi]\!][\![A]\!]^{-1} = \varepsilon$.

In turn, $[\![\Gamma]\!][\![B]\!][\![\Delta]\!] = [\![C]\!]$ entails $[\![\Gamma]\!][\![\Pi]\!][\![A]\!]^{-1}[\![B]\!][\![\Delta]\!] = [\![C]\!]$. Thus

$$[\![\Gamma]\!][\![\Pi]\!][\![A\backslash B]\!][\![\Delta]\!] = [\![C]\!].$$

CASE 7: $(/ \to)$

Similar to the previous case. ■

5 Interpolation

In this section we prove the interpolation theorem for the product-free fragment of the Lambek calculus and obtain a corollary for interpolation of thin sequents.

Interpolation in the product-free fragment of the Lambek calculus is more complicated than in the full Lambek calculus. (See [12] for the proof of the interpolation theorem in the full Lambek calculus allowing empty antecedents.) Namely, in the product-free fragment we must allow not only single types, but also finite sequences of types to appear as interpolants.

Lemma 7 *Let $\Phi \in \mathrm{Tp}(\backslash,/)^*$, $\Theta \in \mathrm{Tp}(\backslash,/)^*$, $\Psi \in \mathrm{Tp}(\backslash,/)^*$, $C \in \mathrm{Tp}(\backslash,/)$, and $L \vdash \Phi\Theta\Psi \to C$. (Some, but not all, of Φ, Θ, and Ψ can be empty.) Then there is a natural number $r{\geq}0$, there are sequences of types $\Theta_1,\ldots,\Theta_r \in \mathrm{Tp}(\backslash,/)^+$, and there are types $E_1,\ldots,E_r \in \mathrm{Tp}(\backslash,/)$ such that*

(i) $\Theta_1 \ldots \Theta_r = \Theta$, *i.e., the sequence Θ is divided into r nonempty continuous subsequences (if $\Theta = \emptyset$ then $r{=}0$);*

(ii) $L \vdash \Theta_j \to E_j$ *for any $j{\leq}r$;*

(iii) $L \vdash \Phi E_1 \ldots E_r \Psi \to C$*;*

(iv) $\sigma_i(E_1 \ldots E_r) \leq \min(\sigma_i(\Theta), \sigma_i(\Phi\Psi C))$ *for any i.*

We shall write $\Phi[\Theta]\Psi \to C$ instead of $\Phi\Theta\Psi \to C$ in order to show the selected part of the antecedent.

We say that the sequence $E_1 \ldots E_r$ is an *interpolant* for Θ in $\Phi\Theta\Psi \to C$.

Example 1 Consider the derivable sequent $[p_1(p_1\backslash p_2)p_3](p_3\backslash(p_2\backslash p_4)) \to p_4$. Applying Lemma 7 we obtain a division of the selected subsequence $p_1(p_1\backslash p_2)p_3 \equiv \Theta_1\Theta_2$, where $\Theta_1 = p_1(p_1\backslash p_2)$ and $\Theta_2 = p_3$ (here $r{=}2$). The corresponding interpolant is p_2p_3, i.e., $E_1 = p_2$ and $E_2 = p_3$. Really, $L \vdash p_1(p_1\backslash p_2) \to p_2$, $L \vdash p_3 \to p_3$, and $L \vdash p_2p_3(p_3\backslash(p_2\backslash p_4)) \to p_4$. Note that no single product-free formula is an interpolant for $p_1(p_1\backslash p_2)p_3$ in $[p_1(p_1\backslash p_2)p_3](p_3\backslash(p_2\backslash p_4)) \to p_4$.

Proof of Lemma 7. Induction on the length of a cut-free derivation.

Case 1: $\Phi\Theta\Psi \to C$ is an axiom, i.e., there exists i such that $C \equiv p_i \equiv \Phi\Theta\Psi$. Actually, the proof also works for axioms of the form $C \to C$ for arbitrary (not necessarily primitive) types C.

Case 1a: $[C] \to C$

We put $r = 1$, $\Theta_1 = C$, $E_1 = C$.

Case 1b: $[\,]C \to C$

We put $r = 0$.

Case 1c: $C[\,] \to C$

We put $r = 0$.

In all the following cases we shall consider the partition of premises induced by the given partition of the conclusion of a rule. By induction hypothesis there exist interpolants for the premises.

Case 2: $(\to \backslash)$

$$\frac{A\Phi[\Theta]\Psi \to B}{\Phi[\Theta]\Psi \to A\backslash B}\ (\to \backslash)$$

By the induction hypothesis we find $\Theta_1, \ldots, \Theta_r, E_1, \ldots, E_r$ such that $\Theta_1 \ldots \Theta_r = \Theta$, $L \vdash \Theta_j \to E_j$ for any $j \leq r$, $A\Phi E_1 \ldots E_r\Psi \to B$, and $\sigma_i(E_1 \ldots E_r) \leq \min(\sigma_i(\Theta), \sigma_i(\Phi\Psi C))$ for any i.

We verify that (i), (ii), (iii), and (iv) hold for the conclusion of the rule $(\to \backslash)$ with the same $\Theta_1, \ldots, \Theta_r, E_1, \ldots, E_r$ as for the premise. The clauses (i) and (ii) are evident from the induction hypothesis. The derivation

$$\frac{A\Phi E_1 \ldots E_r\Psi \to B}{\Phi E_1 \ldots E_r\Psi \to A\backslash B}\ (\to \backslash)$$

establishes (iii). The clause (iv) follows from $\sigma_i(\Gamma\Delta(A\backslash B)) = \sigma_i(A\Gamma\Delta B)$ and the induction hypothesis.

Case 3: $(\to /)$

Similar.

Case 4: $(\backslash \to)$

Case 4a:

$$\frac{\Pi'[\Pi'']\Pi''' \to A \quad \Gamma B\Delta \to C}{\Gamma\Pi'[\Pi'']\Pi'''(A\backslash B)\Delta \to C}\ (\backslash \to)$$

Similar to case 2.

Case 4b:

$$\frac{\Pi \to A \quad \Gamma'[\Gamma'']\Gamma''' B\Delta \to C}{\Gamma'[\Gamma'']\Gamma'''\Pi(A\backslash B)\Delta \to C}\ (\backslash \to)$$

Similar to case 2.

Case 4c:

$$\frac{\Pi \to A \quad \Gamma B\Delta'[\Delta'']\Delta''' \to C}{\Gamma\Pi(A\backslash B)\Delta'[\Delta'']\Delta''' \to C}\ (\backslash \to)$$

Similar to case 2.

Case 4d:

$$\frac{[\Pi']\Pi'' \to A \quad \Gamma'[\Gamma'']B\Delta \to C}{\Gamma'[\Gamma''\Pi']\Pi''(A\backslash B)\Delta \to C}\ (\backslash \to)$$

Let $E_1 \ldots E_r$ and $F_1 \ldots F_m$ be the interpolants of the left and right premises respectively. It is easy to verify that $F_1 \ldots F_m E_1 \ldots E_r$ is an interpolant for the conclusion of the rule $(\backslash \to)$.

Case 4e:

$$\frac{\Pi \to A \quad \Gamma'[\Gamma'' B\Delta']\Delta'' \to C}{\Gamma'[\Gamma''\Pi(A\backslash B)\Delta']\Delta'' \to C}\ (\backslash \to)$$

Let $E_1 \dots E_r$ denote the interpolant for the right premise. We establish that it is also an interpolant for the conclusion. The clause (iii) is obvious.

By the induction hypothesis, $\Gamma'' B \Delta' \equiv \Theta_1 \dots \Theta_r$. Let the particular occurrence of formula B belong to the part Θ_k. Then $\Theta_k \equiv \Xi B \Upsilon$ for some sequences Ξ and Υ.

We put $\tilde{\Theta_k} = \Xi\Pi(A\backslash B)\Upsilon$ and $\tilde{\Theta_j} = \Theta_j$ for any $j \neq k$. Evidently $\Gamma''\Pi(A\backslash B)\Delta' \equiv \tilde{\Theta_1} \dots \tilde{\Theta_r}$. Using the induction hypothesis (ii) we obtain

$$\frac{\Pi \to A \quad \Xi B \Upsilon \to E_k}{\Xi\Pi(A\backslash B)\Upsilon \to E_k} \, (\backslash \to)$$

and $\tilde{\Theta_j} \to E_j$ for any $j \neq k$. This proves (ii). To prove (iv), it is sufficient to observe that $\sigma_i(\Gamma'' B \Delta') \leq \sigma_i(\Gamma''\Pi(A\backslash B)\Delta')$.

Case 4f:

$$\frac{[\Pi']\Pi'' \to A \quad \Gamma[B\Delta']\Delta'' \to C}{\Gamma\Pi'[\Pi''(A\backslash B)\Delta']\Delta'' \to C} \, (\backslash \to)$$

Let $E_1, \dots, E_r$ be an interpolant for the right premise, corresponding to the partition $B\Delta' \equiv \Theta_1 \dots \Theta_r$. Let $F_1, \dots, F_m$ be an interpolant for the left premise, corresponding to the partition $\Pi' \equiv \Xi_1 \dots \Xi_m$.

Then, for a suitable sequence Υ,

(1) $\Theta_1 \equiv B\Upsilon$,

(2) $\Delta' \equiv \Upsilon\Theta_2 \dots \Theta_r$,

(3) $B\Upsilon \to E_1$,

(4) $\Theta_j \to E_j$ for any $j \neq 1$,

(5) $\Gamma E_1 \dots E_r \Delta'' \to C$,

(6) $\sigma_i(E_1 \dots E_r) \leq \min(\sigma_i(B\Delta'), \sigma_i(\Gamma\Delta'' C))$ for any i,

(7) $\Pi' \equiv \Xi_1 \dots \Xi_m$,

(8) $\Xi_j \to F_j$ for any $j \leq m$,

(9) $F_1 \dots F_m \Pi'' \to A$,

(10) $\sigma_i(F_1 \dots F_m) \leq \min(\sigma_i(\Pi'), \sigma_i(\Pi'' A))$ for any i.

We show that $(F_m\backslash(\dots\backslash(F_1\backslash E_1)\dots))\, E_2 \dots E_r$ is an interpolant of the conclusion, corresponding to the partition $\Pi''(A\backslash B)\Delta' \equiv \tilde{\Theta}_1 \dots \tilde{\Theta}_r$, where $\tilde{\Theta}_1 = \Pi''(A\backslash B)\Upsilon$ and $\tilde{\Theta}_j = \Theta_j$ for any $j \neq 1$.

Clause (iv) is obvious from $\sigma_i(E_1 \dots E_r) + \sigma_i(F_1 \dots F_m) \leq \min(\sigma_i(B\Delta'), \sigma_i(\Gamma\Delta'' C)) + \min(\sigma_i(\Pi'' A), \sigma_i(\Pi')) \leq \min(\sigma_i(\Pi''(A\backslash B)\Delta'), \sigma_i(\Gamma\Pi'\Delta'' C))$.

Evidently, (i) holds, since $\Pi''(A\backslash B)\Delta' \equiv \tilde{\Theta}_1 \dots \tilde{\Theta}_r$.

Next we prove (ii). We only need to verify that $L \vdash \tilde{\Theta}_1 \to (F_m\backslash(\ldots\backslash(F_1\backslash E_1)\ldots))$.

$$\dfrac{\dfrac{\dfrac{\dfrac{F_1\ldots F_m\Pi'' \to A \quad B\Upsilon \to E_1}{F_1\ldots F_m\Pi''(A\backslash B)\Upsilon \to E_1}(\backslash\to)}{F_2\ldots F_m\Pi''(A\backslash B)\Upsilon \to F_1\backslash E_1}(\to\backslash)}{\vdots}(\to\backslash)}{\Pi''(A\backslash B)\Upsilon \to (F_m\backslash(\ldots\backslash(F_1\backslash E_1)\ldots))}(\to\backslash)$$

Finally, we prove (iii).

$$\dfrac{\Xi_m \to F_m \quad \dfrac{\dfrac{\dfrac{\Xi_1 \to F_1 \quad \Gamma E_1 E_2\ldots E_r\Delta'' \to C}{\Gamma\Xi_1(F_1\backslash E_1)E_2\ldots E_r\Delta'' \to C}(\backslash\to)}{\vdots}(\backslash\to)}{\Gamma\Xi_1\ldots\Xi_{m-1}(F_{m-1}\backslash(\ldots\backslash(F_1\backslash E_1)\ldots))E_2\ldots E_r\Delta'' \to C}(\backslash\to)}{\Gamma\Xi_1\ldots\Xi_{m-1}\Xi_m(F_m\backslash(F_{m-1}\backslash(\ldots\backslash(F_1\backslash E_1)\ldots)))E_2\ldots E_r\Delta'' \to C}(\backslash\to)$$

CASE 5: $(/\to)$
Similar to case 4. ■

Lemma 8 *Let the sequent $\Phi\Theta\Psi \to C$ be thin and $E_1\ldots E_r$ be an interpolant for Θ in $\Phi[\Theta]\Psi \to C$, corresponding to the partition $\Theta \equiv \Theta_1\ldots\Theta_r$. Then*

(i) *for any $i\leq r$, the sequent $\Theta_i \to E_i$ is thin;*

(ii) *the sequent $\Phi E_1\ldots E_r\Psi \to C$ is thin;*

(iii) $\|E_1\ldots E_r\| = |[\![\Theta]\!]|$.

PROOF. To prove (i), we note that $\sigma_i(E_j) \leq \sigma_i(E_1\ldots E_r) \leq \sigma_i(\Phi\Psi C)$ and thus $\sigma_i(\Theta_j E_j) = \sigma_i(\Theta_j) + \sigma_i(E_j) \leq \sigma_i(\Theta) + \sigma_i(\Phi\Psi C) \leq 2$.

To establish (ii), we observe that $\sigma_i(\Phi E_1\ldots E_r\Psi C) = \sigma_i(\Phi\Psi C) + \sigma_i(E_1\ldots E_r) \leq \sigma_i(\Phi\Psi C) + \sigma_i(\Theta) \leq 2$.

It remains to prove (iii). According to Lemma 6, $[\![\Phi]\!][\![\Theta]\!][\![\Psi]\!] = [\![C]\!]$, whence $[\![\Theta]\!] = [\![\Phi]\!]^{-1}[\![C]\!][\![\Psi]\!]^{-1}$. Thus the reduced words for $[\![\Theta]\!]$ and $[\![\Phi]\!]^{-1}[\![C]\!][\![\Psi]\!]^{-1}$ coincide. Next we verify that the reduced word for $[\![\Theta]\!]$ contains exactly those primitive types that occur in $E_1\ldots E_r$. Take an arbitrary positive integer i.
CASE 1: $\sigma_i(\Theta) = 0$
In this case p_i occurs neither in $[\![\Theta]\!]$ nor in $E_1\ldots E_r$.
CASE 2: $\sigma_i(\Theta) = 1$
Now there is exactly one occurrence of p_i in $[\![\Theta]\!]$. Obviously, $\sigma_i(E_1\ldots E_r) \leq 1$. On the other hand, $\sigma_i(E_1\ldots E_r) \neq 0$, since both, $\sigma_i(\Phi\Theta\Psi C)$ and $\sigma_i(\Phi E_1\ldots E_r\Psi C)$, are even.
CASE 3: $\sigma_i(\Theta) = 2$
In this case $\sigma_i(\Phi\Psi C) = 0$, whence p_i does not occur in $[\![\Phi]\!]^{-1}[\![C]\!][\![\Psi]\!]^{-1}$ and consequently it has no occurrences in the reduced word for $[\![\Theta]\!]$. Evidently $\sigma_i(E_1\ldots E_r) = 0$.

We have also seen that no primitive type has more than one occurrence in the reduced word for $[\![\Theta]\!]$ and no primitive type has more than one occurence in $E_1\ldots E_r$. Thus $\|E_1\ldots E_r\| = |[\![\Theta]\!]|$. ■

6 Completeness of $Lcut(m, q)$ in the case of thin sequents

6.1 A property of the free group

Lemma 9 *If $u_1, \dots, u_n \in FG$, $n>1$, and $u_1 \dots u_n = \varepsilon$, then there is a number $k<n$ such that $|u_k u_{k+1}| \leq \max(|u_k|, |u_{k+1}|)$.*

Proof. Let k be the least positive integer less than n such that $|u_1 \dots u_{k+1}| \leq |u_1 \dots u_k|$. If $k = 1$, then the proof is obvious. Let $1 < k < n$. Then $|u_1 \dots u_k| > |u_1 \dots u_{k-1}|$. We put $u \rightleftharpoons u_1 \dots u_{k-1}$, $v \rightleftharpoons u_k$, and $w \rightleftharpoons u_{k+1}$. Now we can apply the following lemma.

Lemma 9.1 *If $u, v, w \in FG$, $|u| < |uv|$, and $|uv| \geq |uvw|$, then $|vw| \leq \max(|v|, |w|)$.*

Proof. Assume for the contrary that $|vw| > |v|$ and $|vw| > |w|$.

There exist three reduced words x_1, y_1, and z_1 in FG such that $u = x_1 y_1^{-1}$, $v = y_1 z_1$, $uv = x_1 z_1$, and the words $x_1 y_1^{-1}$, $y_1 z_1$, $x_1 z_1$ are reduced. From $|u| < |uv|$ we obtain $|x_1|+|y_1| < |x_1|+|z_1|$, whence $|y_1| < |z_1|$, and finally $|y_1| < \frac{1}{2}|v|$.

Similarly, there exist three reduced words x_2, y_2, and z_2 in FG such that $v = x_2 y_2$, $w = y_2^{-1} z_2$, $vw = x_2 z_2$, and the words $x_2 y_2$, $y_2^{-1} z_2$, $x_2 z_2$ are reduced. From $|w| < |vw|$ we obtain $|y_2|+|z_2| < |x_2|+|z_2|$, whence $|y_2| < |x_2|$, and finally $|y_2| < \frac{1}{2}|v|$.

The reduced words $y_1 z_1$ and $x_2 y_2$ coincide. In view of $|y_1| < \frac{1}{2}|v|$ and $|y_2| < \frac{1}{2}|v|$ there exists $v_0 \in FG$ such that $z_1 = v_0 y_2$, $x_2 = y_1 v_0$, $y_1 v_0 y_2$ is a reduced word, and $v_0 \neq \varepsilon$.

Now we can represent uvw as $x_1 y_1^{-1} \cdot y_1 v_0 y_2 \cdot y_2^{-1} z_2 = x_1 v_0 z_2$. Note that $x_1 v_0 z_2$ is reduced, since $v_0 \neq \varepsilon$ and both $x_1 v_0$ and $v_0 z_2$ are reduced. Thus $|x_1 v_0 z_2| = |uvw| \leq |uv| = |x_1 z_1| = |x_1 v_0 y_2|$ and therefore $|z_2| \leq |y_2|$.

On the other hand, from $|vw| > |v|$ we obtain $|x_2 z_2| > |x_2 y_2|$, whence $|z_2| > |y_2|$. Contradiction. ■ ■

6.2 Proof of Theorem 1

We shall prove Theorem 1 for the case of thin sequents.

Lemma 10 *Let $A_1, \dots, A_n, C \in \mathrm{Tp}(m, q)$. If the sequent $A_1 \dots A_n \to C$ is derivable and thin, then $Lcut(m, q) \vdash A_1 \dots A_n \to C$.*

Proof. Induction on $\|A_1 \dots A_n\|$. If $\|A_1 \dots A_n\| \leq 2m$, then the sequence $A_1 \dots A_n$ belongs to $\mathrm{Ls}(m, q)$ and thus $A_1 \dots A_n \to C$ is an axiom of $Lcut(m, q)$.

Assume that $\|A_1 \dots A_n\| > 2m$. We are going to divide the sequence $A_1 \dots A_n$ into continuous subsequences $\Pi_1 = A_1 \dots A_{j_1}$, $\Pi_2 = A_{j_1+1} \dots A_{j_2}$, $\dots$, $\Pi_l = A_{j_{l-1}+1} \dots A_{j_l}$ such that $0 < j_1 < j_2 < \dots < j_l = n$, $\|\Pi_i\| \leq m$ for any $i \leq l$, and $\|\Pi_i \Pi_{i+1}\| > m$ for any $i<l$. This is done as follows.

First we choose for j_1 the maximal value satisfying $\|A_1 \dots A_{j_1}\| \leq m$. Next, for each i we put $j_i = \max\{j \mid j_{i-1}<j\leq n \text{ and } \|A_{j_{i-1}+1} \dots A_j\| \leq m\}$ until we obtain $j_i = n$. Obviously, for any i, $\|\Pi_i\| \leq m$ and $\|\Pi_i \Pi_{i+1}\| > m$.

Note that $[\![\Pi_1]\!] \dots [\![\Pi_l]\!] = [\![A_1]\!] \dots [\![A_n]\!] = [\![C]\!]$ according to Lemma 6. Thus $[\![\Pi_1]\!] \dots [\![\Pi_l]\!][\![C]\!]^{-1} = \varepsilon$. Now let $u_1 \rightleftharpoons [\![\Pi_1]\!]$, $\dots u_l \rightleftharpoons [\![\Pi_l]\!]$, and $u_{l+1} \rightleftharpoons [\![C]\!]^{-1}$. Evidently $u_1 \dots u_l u_{l+1} = \varepsilon$ and $|u_i| \leq m$ for any i (recall that $|[\![\Pi_i]\!]| \leq \|\Pi_i\|$ and $|[\![C]\!]^{-1}| = |[\![C]\!]| \leq \|C\|$).

Applying Lemma 9 we find a positive integer $k \leq l$ such that $|u_k u_{k+1}| \leq m$. The following two cases arise.

Case 1: $k < l$

We have $|[\![\Pi_k \Pi_{k+1}]\!]| \leq m$ for that particular k. Applying Lemma 8 for

$$\underbrace{\Pi_1 \dots \Pi_{k-1}}_{\Phi} \underbrace{\Pi_k \Pi_{k+1}}_{\Theta} \underbrace{\Pi_{k+2} \dots \Pi_l}_{\Psi} \to C$$

we find an interpolant $E_1 \dots E_r$ for $\Pi_k \Pi_{k+1}$ in $\Pi_1 \dots \Pi_l \to C$. This means that $\Pi_k \Pi_{k+1}$ is divided into r continuous subsequences $\Theta_1, \dots, \Theta_r$ such that $L \vdash \Theta_i \to E_i$ for every $i \leq r$, $L \vdash \Pi_1 \dots \Pi_{k-1} E_1 \dots E_r \Pi_{k+2} \dots \Pi_l \to C$, and $\|E_1 \dots E_r\| = |[\![\Pi_k \Pi_{k+1}]\!]| \leq m$.

Note that $\|E_1 \dots E_r\| \leq m$, but $\|\Pi_k \Pi_{k+1}\| > m$. Thus

$$\|\Pi_1 \dots \Pi_{k-1} E_1 \dots E_r \Pi_{k+2} \dots \Pi_l\| < \|\Pi_1 \dots \Pi_l\|$$

and we can apply the induction hypothesis for the thin derivable sequent $\Pi_1 \dots \Pi_{k-1} E_1 \dots E_r \Pi_{k+2} \dots \Pi_l \to C$.

On the other hand, for any $i \leq r$, $\Theta_i \to E_i$ is an axiom of $Lcut(m,q)$, since $\|E_i\| \leq \|E_1 \dots E_r\| \leq m$ and $\|\Theta_i\| \leq \|\Pi_k \Pi_{k+1}\| \leq 2m$.

We have proved that $Lcut(m,q) \vdash \Pi_1 \dots \Pi_{k-1} E_1 \dots E_r \Pi_{k+2} \dots \Pi_l \to C$ and $Lcut(m,q) \vdash \Theta_i \to E_i$ for any $i \leq r$. Applying the cut rule r times we derive $Lcut(m,q) \vdash \Pi_1 \dots \Pi_{k-1} \Theta_1 \dots \Theta_r \Pi_{k+2} \dots \Pi_l \to C$, i.e., $Lcut(m,q) \vdash \Pi_1 \dots \Pi_l \to C$, i.e., $Lcut(m,q) \vdash A_1 \dots A_n \to C$.

Case 2: $k = l$

We have $|[\![\Pi_l]\!][\![C]\!]^{-1}| \leq m$. Applying Lemma 8 for

$$\underbrace{\Pi_1 \dots \Pi_{l-1}}_{\Theta} \underbrace{\Pi_l}_{\Psi} \to C$$

we find an interpolant $E_1 \dots E_r$ for $\Pi_1 \dots \Pi_{l-1}$ in $\Pi_1 \dots \Pi_l \to C$. This means that $\Pi_1 \dots \Pi_{l-1}$ is divided into r continuous subsequences $\Theta_1, \dots, \Theta_r$ such that $L \vdash \Theta_i \to E_i$ for every $i \leq r$, $L \vdash E_1 \dots E_r \Pi_l \to C$, and $\|E_1 \dots E_r\| = |[\![\Pi_1 \dots \Pi_{l-1}]\!]|$.

Recall that $[\![\Pi_1 \dots \Pi_{l-1} \Pi_l]\!] = [\![C]\!]$, whence $[\![\Pi_1 \dots \Pi_{l-1}]\!] = [\![C]\!][\![\Pi_l]\!]^{-1} = ([\![\Pi_l]\!][\![C]\!]^{-1})^{-1}$ and further, $|[\![\Pi_1 \dots \Pi_{l-1}]\!]| = |([\![\Pi_l]\!][\![C]\!]^{-1})^{-1}| = |[\![\Pi_l]\!][\![C]\!]^{-1}| \leq m$. Thus $\|E_1 \dots E_r\| = |[\![\Pi_1 \dots \Pi_{l-1}]\!]| \leq m$. It follows that $E_1 \dots E_r \Pi_l \in \mathrm{Ls}(m,q)$ and consequently, $E_1 \dots E_r \Pi_l \to C$ is an axiom of $Lcut(m,q)$.

On the other hand, for any $i \leq r$, $\|\Theta_i\| \leq \|\Theta_1 \dots \Theta_r\| = \|\Pi_1 \dots \Pi_{l-1}\| < \|\Pi_1 \dots \Pi_l\|$ and we can apply the induction hypothesis for the thin derivable sequents $\Theta_i \to E_i$.

The rest of the proof of case 2 is similar to that of case 1. ■

Proof of Theorem 1. Immediate from Lemma 10 and Lemma 5. ■

In [5] W. Buszkowski presented a similar proof of Theorem 2 for the case if the designated type D of a Lambek grammar is primitive.

Remark. All the results of this paper hold also in the full Lambek calculus (including product), in the multiplicative linear logic, and in the Lambek calculus with the unit. Thus, the grammars relying on the multiplicative fragment of the linear logic generate only context-free languages (and they generate all context-free languages). Here we assume that a language generated by a linear logic grammar contains, by definition, only non-empty words.

Acknowledgements

I am grateful to Prof. S. Artemov for guiding me into the subject and pointing out the most important problems, to Prof. M. Kanovich for suggesting useful improvements of exposition, to L. Beklemishev, V. Krupski, and N. Pankratiev for checking the proof. I also wish to thank the anonymous referee, who suggested significant improvements of the proofs of Theorem 2, Lemma 8, and Lemma 9.

References

[1] Y. Bar-Hillel, C. Gaifman, and E. Shamir. On categorial and phrase-structure grammars. *Bull. Res. Council Israel Sect. F*, 9F:1–16, 1960.

[2] W. Buszkowski. The equivalence of unidirectional Lambek categorial grammars and context-free grammars. *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, 31:369–384, 1985.

[3] W. Buszkowski. Generative power of categorial grammars. In R.T. Oehrle, E. Bach, and D. Wheeler, editors, *Categorial Grammars and Natural Language Structures*, pages 69–94, Reidel, Dordrecht, 1988.

[4] W. Buszkowski. On generative capacity of the Lambek calculus. In J. van Eijck, editor, *Logics in AI*, pages 139–152, Springer, Berlin, 1991.

[5] W. Buszkowski. On the equivalence of Lambek categorial grammars and basic categorial grammars. ILLC Prepublication Series, LP–93–07 Institute for Logic, Language and Computation, University of Amsterdam, 1993.

[6] N. Chomsky. Formal properties of grammars. In R.D. Luce et al., editors, *Handbook of Mathematical Psychology*, vol. 2, pages 323–418, Wiley, New York, 1963.

[7] J.M. Cohen. The equivalence of two concepts of categorial grammar. *Information and Control*, 10:475–484, 1967.

[8] J. Lambek. The mathematics of sentence structure. *American Mathematical Monthly*, 65(3):154–170, 1958.

[9] M. Pentus. *Equivalent Types in Lambek Calculus and Linear Logic*. Preprint No.2 of the Department of Math. Logic, Steklov Math. Institute, Series Logic and Computer Science, Moscow, 1992.

[10] M. Pentus. *Lambek Grammars Are Context Free*. Preprint No.8 of the Department of Math. Logic, Steklov Math. Institute, Series Logic and Computer Science, Moscow, 1992.

[11] M. Pentus. *Lambek grammars are context free*. Proceedings of the 8th Annual IEEE Symposium on Logic in Computer Science, LICS'93, 429–433, 1993.

[12] D. Roorda. *Resource Logics: Proof-theoretical Investigations*. PhD thesis, Fac. Math. and Comp. Sc., University of Amsterdam, 1991.

[13] J. van Benthem. *Language in Action*. North-Holland, Amsterdam, 1991.