

Topological Complexity of Context-Free omega-Languages: A Survey

Olivier Finkel

► **To cite this version:**

Olivier Finkel. Topological Complexity of Context-Free omega-Languages: A Survey. Nachum Der-showitz and Ephraim Nissan. Language, Culture, Computation. Computing - Theory and Technology - Essays Dedicated to Yaacov Choueka on the Occasion of His 75th Birthday, Part I, 8001, Springer, pp.50–77, 2014, Lecture Notes in Computer Science, 978-3-642-45320-5. ensl-00286373v2

HAL Id: ensl-00286373

<https://hal-ens-lyon.archives-ouvertes.fr/ensl-00286373v2>

Submitted on 12 Mar 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Topological Complexity of Context-Free ω -Languages : A Survey

Olivier Finkel

Equipe de Logique Mathématique
Institut de Mathématiques de Jussieu - Paris Rive Gauche UMR7586
CNRS et Université Paris Diderot Paris 7
Bâtiment Sophie Germain Case 7012
75205 Paris Cedex 13, France.
finkel@math.univ-paris-diderot.fr

Abstract. We survey recent results on the topological complexity of context-free ω -languages which form the second level of the Chomsky hierarchy of languages of infinite words. In particular, we consider the Borel hierarchy and the Wadge hierarchy of non-deterministic or deterministic context-free ω -languages. We study also decision problems, the links with the notions of ambiguity and of degrees of ambiguity, and the special case of ω -powers.

Keywords: Infinite words; pushdown automata; context-free (ω)-languages; ω -powers; Cantor topology; topological complexity; Borel hierarchy; Wadge hierarchy; complete sets; decision problems.

1 Introduction

The Chomsky hierarchy of formal languages of finite words over a finite alphabet is now well known, [49]. The class of regular languages accepted by finite automata forms the first level of this hierarchy and the class of context-free languages accepted by pushdown automata or generated by context-free grammars forms its second level [3]. The third and the fourth levels are formed by the class of context-sensitive languages accepted by linear-bounded automata or generated by Type-1 grammars and the class of recursively enumerable languages accepted by Turing machines or generated by Type-0 grammars [15]. In particular, context-free languages, firstly introduced by Chomsky to analyse the syntax of natural languages, have been very useful in Computer Science, in particular in the domain of programming languages, for the construction of compilers used to verify correctness of programs, [48].

There is a hierarchy of languages of infinite words which is analogous to the Chomsky hierarchy but where the languages are formed by infinite words over a finite alphabet. The first level of this hierarchy is formed by the class of regular ω -languages accepted by finite automata. They were first studied by Büchi in order to study decision problems for logical theories. In particular, Büchi proved that the monadic second order theory of one successor over the integers is decidable, using finite automata equipped

with a certain acceptance condition for infinite words, now called the Büchi acceptance condition. Well known pioneers in this research area are named Muller, Mc Naughton, Rabin, Landweber, Choueka, [61, 62, 68, 52, 16]. The theory of regular ω -languages is now well established and has found many applications for specification and verification of non-terminating systems; see [81, 78, 67] for many results and references. The second level of the hierarchy is formed by the class of context-free ω -languages. As in the case of languages of finite words it turned out that an ω -language is accepted by a (non-deterministic) pushdown automaton (with Büchi acceptance condition) if and only if it is generated by a context-free grammar where infinite derivations are considered. Context-free languages of infinite words were first studied by Cohen and Gold, [19, 20], Linna, [56–58], Boasson, Nivat, [64, 63, 7, 8], Beauquier, [4], see the survey [78]. Notice that in the case of infinite words Type-1 grammars and Type-0 grammars accept the same ω -languages which are also the ω -languages accepted by Turing machines with a Büchi acceptance condition [21, 78], see also the fundamental study of Engelfriet and Hoogetboom on \mathbf{X} -automata, i.e. finite automata equipped with a storage type \mathbf{X} , accepting infinite words, [29].

Context-free ω -languages have occurred recently in the works on games played on infinite pushdown graphs, following the fundamental study of Walukiewicz, [85, 82] [74, 40].

Since the set X^ω of infinite words over a finite alphabet X is naturally equipped with the Cantor topology, a way to study the complexity of ω -languages is to study their topological complexity. The first task is to locate ω -languages with regard to the Borel and the projective hierarchies, and next to the Wadge hierarchy which is a great refinement of the Borel hierarchy. It is then natural to ask for decidability properties and to study decision problems like : is there an effective procedure to determine the Borel rank or the Wadge degree of any context-free ω -language ? Such questions were asked by Lescow and Thomas in [55]. In this paper we survey some recent results on the topological complexity of context-free ω -languages. Some of them were very surprising as the two following ones:

1. there is a 1-counter finitary language L such that L^ω is analytic but not Borel, [35].
2. The Wadge hierarchy, hence also the Borel hierarchy, of ω -languages accepted by real time 1-counter Büchi automata is the same as the Wadge hierarchy of ω -languages accepted by Büchi Turing machines, [41].

The Borel and Wadge hierarchies of *non deterministic* context-free ω -languages are not effective. One can neither decide whether a given context-free ω -language is a Borel set nor whether it is in a given Borel class Σ_α^0 or Π_α^0 . On the other hand *deterministic* context-free ω -languages are located at a low level of the Borel hierarchy: they are all Δ_3^0 -sets. They enjoy some decidability properties although some important questions in this area are still open. We consider also the links with the notions of ambiguity and of degrees of ambiguity, and the special case of ω -powers, i.e. of ω -languages in the form V^ω , where V is a (context-free) finitary language. Finally we state some perspectives and give a list of some questions which remain open for further study.

The paper is organized as follows. In Section 2 we recall the notions of context-free ω -languages accepted by Büchi or Muller pushdown automata. Topological notions and Borel and Wadge hierarchies are recalled in Section 3. In Section 4 is studied the case of non-deterministic context-free ω -languages while deterministic context-free ω -languages are considered in Section 5. Links with notions of ambiguity in context free languages are studied in Section 6. Section 7 is devoted to the special case of ω -powers. Perspectives and some open questions are presented in last Section 8.

2 Context-free ω -languages

We assume the reader to be familiar with the theory of formal (ω)-languages [81, 78]. We shall use usual notations of formal language theory.

When X is a finite alphabet, a *non-empty finite word* over X is any sequence $x = a_1 \dots a_k$, where $a_i \in X$ for $i = 1, \dots, k$, and k is an integer ≥ 1 . The *length* of x is k , denoted by $|x|$. The *empty word* has no letters and is denoted by λ ; its length is 0. For $x = a_1 \dots a_k$, we write $x(i) = a_i$ and $x[i] = x(1) \dots x(i)$ for $i \leq k$ and $x[0] = \lambda$. X^* is the *set of finite words* (including the empty word) over X .

For $V \subseteq X^*$, the complement of V (in X^*) is $X^* - V$ denoted V^- .

The *first infinite ordinal* is ω . An ω -word over X is an ω -sequence $a_1 \dots a_n \dots$, where for all integers $i \geq 1$, $a_i \in X$. When σ is an ω -word over X , we write $\sigma = \sigma(1)\sigma(2) \dots \sigma(n) \dots$, where for all i , $\sigma(i) \in X$, and $\sigma[n] = \sigma(1)\sigma(2) \dots \sigma(n)$ for all $n \geq 1$ and $\sigma[0] = \lambda$.

The usual concatenation product of two finite words u and v is denoted $u.v$ (and sometimes just uv). This product is extended to the product of a finite word u and an ω -word v : the infinite word $u.v$ is then the ω -word such that:

$$(u.v)(k) = u(k) \text{ if } k \leq |u|, \text{ and } (u.v)(k) = v(k - |u|) \text{ if } k > |u|.$$

The *prefix relation* is denoted \sqsubseteq : a finite word u is a *prefix* of a finite word v (respectively, an infinite word v), denoted $u \sqsubseteq v$, if and only if there exists a finite word w (respectively, an infinite word w), such that $v = u.w$. The *set of ω -words* over the alphabet X is denoted by X^ω . An ω -language over an alphabet X is a subset of X^ω . The complement (in X^ω) of an ω -language $V \subseteq X^\omega$ is $X^\omega - V$, denoted V^- .

For $V \subseteq X^*$, the ω -power of V is :

$$V^\omega = \{\sigma = u_1 \dots u_n \dots \in X^\omega \mid \forall i \geq 1 \ u_i \in V\}.$$

We now define pushdown machines and the class of ω -context-free languages.

Definition 1. A *pushdown machine (PDM)* is a 6-tuple $M = (K, X, \Gamma, \delta, q_0, Z_0)$, where K is a finite set of states, X is a finite input alphabet, Γ is a finite pushdown alphabet, $q_0 \in K$ is the initial state, $Z_0 \in \Gamma$ is the start symbol, and δ is a mapping from $K \times (X \cup \{\lambda\}) \times \Gamma$ to finite subsets of $K \times \Gamma^*$.

If $\gamma \in \Gamma^+$ describes the pushdown store content, the leftmost symbol will be assumed to be on "top" of the store. A configuration of a PDM is a pair (q, γ) where $q \in K$ and $\gamma \in \Gamma^*$.

For $a \in X \cup \{\lambda\}$, $\beta, \gamma \in \Gamma^*$ and $Z \in \Gamma$, if (p, β) is in $\delta(q, a, Z)$, then we write $a : (q, Z\gamma) \mapsto_M (p, \beta\gamma)$.

\mapsto_M^* is the transitive and reflexive closure of \mapsto_M . (The subscript M will be omitted whenever the meaning remains clear).

Let $\sigma = a_1 a_2 \dots a_n \dots$ be an ω -word over X . An infinite sequence of configurations $r = (q_i, \gamma_i)_{i \geq 1}$ is called a complete run of M on σ , starting in configuration (p, γ) , iff:

1. $(q_1, \gamma_1) = (p, \gamma)$
2. for each $i \geq 1$, there exists $b_i \in X \cup \{\lambda\}$ satisfying $b_i : (q_i, \gamma_i) \mapsto_M (q_{i+1}, \gamma_{i+1})$ such that $a_1 a_2 \dots a_n \dots = b_1 b_2 \dots b_n \dots$

For every such run, $In(r)$ is the set of all states entered infinitely often during run r . A complete run r of M on σ , starting in configuration (q_0, Z_0) , will be simply called “a run of M on σ ”.

Definition 2. A Büchi pushdown automaton is a 7-tuple $M = (K, X, \Gamma, \delta, q_0, Z_0, F)$ where $M' = (K, X, \Gamma, \delta, q_0, Z_0)$ is a PDM and $F \subseteq K$ is the set of final states. The ω -language accepted by M is

$$L(M) = \{\sigma \in X^\omega \mid \text{there exists a complete run } r \text{ of } M \text{ on } \sigma \text{ such that } In(r) \cap F \neq \emptyset\}$$

Definition 3. A Muller pushdown automaton is a 7-tuple $M = (K, X, \Gamma, \delta, q_0, Z_0, \mathcal{F})$ where $M' = (K, X, \Gamma, \delta, q_0, Z_0)$ is a PDM and $\mathcal{F} \subseteq 2^K$ is the collection of designated state sets. The ω -language accepted by M is

$$L(M) = \{\sigma \in X^\omega \mid \text{there exists a complete run } r \text{ of } M \text{ on } \sigma \text{ such that } In(r) \in \mathcal{F}\}$$

Remark 4. We consider here two acceptance conditions for ω -words, the Büchi and the Muller acceptance conditions, respectively denoted 2-acceptance and 3-acceptance in [52] and in [20] and (inf, \sqcap) and $(inf, =)$ in [78]. We refer the reader to [19, 20, 78, 29] for consideration of weaker acceptance conditions, and to [46, 67] for the definitions of other usual ones like Rabin, Street, or parity acceptance conditions. Notice however that it seems that the latter ones have not been much considered in the study of context-free ω -languages but they are often involved in constructions concerning finite automata reading infinite words.

Notation. In the sequel we shall often abbreviate “Muller pushdown automaton” by MPDA and “Büchi pushdown automaton” by BPDA.

Cohen and Gold and independently Linna established a characterization theorem for ω -languages accepted by Büchi or Muller pushdown automata. We shall need the notion of “ ω -Kleene closure” which we now firstly define:

Definition 5. For any family \mathcal{L} of finitary languages, the ω -Kleene closure of \mathcal{L} is :

$$\omega-KC(\mathcal{L}) = \{\cup_{i=1}^n U_i \cdot V_i^\omega \mid \forall i \in [1, n] \ U_i, V_i \in \mathcal{L}\}$$

Theorem 6 (Linna [56], Cohen and Gold [19]). Let CFL be the class of context-free (finitary) languages. Then for any ω -language L the following three conditions are equivalent:

1. $L \in \omega\text{-KC}(CFL)$.
2. There exists a *BPDA* that accepts L .
3. There exists a *MPDA* that accepts L .

In [19] are also studied ω -languages generated by ω -context-free grammars and it is shown that each of the conditions 1), 2), and 3) of the above Theorem is also equivalent to: 4) L is generated by a context-free grammar G by leftmost derivations. These grammars are also studied by Nivat in [63, 64]. Then we can let the following definition:

Definition 7. An ω -language is a context-free ω -language iff it satisfies one of the conditions of the above Theorem. The class of context-free ω -languages will be denoted by CFL_ω .

If we omit the pushdown store in the above Theorem we obtain the characterization of languages accepted by classical Muller automata (MA) or Büchi automata (BA) :

Theorem 8. For any ω -language L , the following conditions are equivalent:

1. L belongs to $\omega\text{-KC}(REG)$,
where REG is the class of finitary regular languages.
2. There exists a MA that accepts L .
3. There exists a BA that accepts L .

An ω -language L satisfying one of the conditions of the above Theorem is called a regular ω -language. The class of regular ω -languages will be denoted by REG_ω .

It follows from Mc Naughton's Theorem that the expressive power of deterministic MA (DMA) is equal to the expressive power of non deterministic MA, i.e. that every regular ω -language is accepted by a deterministic Muller automaton, [62, 67]. Notice that Choueka gave a simplified proof of Mc Naughton's Theorem in [16]. Another variant was given by Rabin in [68]. Unlike the case of finite automata, deterministic *MPDA* do not define the same class of ω -languages as non deterministic *MPDA*. Let us now define deterministic pushdown machines.

Definition 9. A PDM $M = (K, X, \Gamma, \delta, q_0, Z_0)$ is said to be deterministic iff for each $q \in K, Z \in \Gamma$, and $a \in X$:

1. $\delta(q, a, Z)$ contains at most one element,
2. $\delta(q, \lambda, Z)$ contains at most one element, and
3. if $\delta(q, \lambda, Z)$ is non empty, then $\delta(q, a, Z)$ is empty for all $a \in X$.

It turned out that the class of ω -languages accepted by deterministic *BPDA* is strictly included into the class of ω -languages accepted by deterministic *MPDA*. This latest class is the class $DCFL_\omega$ of deterministic context-free ω -languages. We denote $DCFL$ the class of deterministic context-free (finitary) languages.

Proposition 10 ([20]).

1. $DCFL_\omega$ is closed under complementation, but is neither closed under union, nor under intersection.
2. $DCFL_\omega \subsetneq \omega\text{-KC}(DCFL) \subsetneq CFL_\omega$ (these inclusions are strict).

3 Topology

3.1 Borel hierarchy and analytic sets

We assume the reader to be familiar with basic notions of topology which may be found in [60, 55, 50, 78, 67]. There is a natural metric on the set X^ω of infinite words over a finite alphabet X containing at least two letters which is called the *prefix metric* and defined as follows. For $u, v \in X^\omega$ and $u \neq v$ let $\delta(u, v) = 2^{-l_{\text{pref}}(u, v)}$ where $l_{\text{pref}}(u, v)$ is the first integer n such that $u(n+1)$ is different from $v(n+1)$. This metric induces on X^ω the usual Cantor topology for which *open subsets* of X^ω are in the form $W.X^\omega$, where $W \subseteq X^*$. A set $L \subseteq X^\omega$ is a *closed set* iff its complement $X^\omega - L$ is an open set. Define now the *Borel Hierarchy* of subsets of X^ω :

Definition 11. For a non-null countable ordinal α , the classes Σ_α^0 and Π_α^0 of the Borel Hierarchy on the topological space X^ω are defined as follows:

Σ_1^0 is the class of open subsets of X^ω , Π_1^0 is the class of closed subsets of X^ω , and for any countable ordinal $\alpha \geq 2$:

Σ_α^0 is the class of countable unions of subsets of X^ω in $\bigcup_{\gamma < \alpha} \Pi_\gamma^0$.

Π_α^0 is the class of countable intersections of subsets of X^ω in $\bigcup_{\gamma < \alpha} \Sigma_\gamma^0$.

Recall some basic results about these classes :

Proposition 12.

- (a) $\Sigma_\alpha^0 \cup \Pi_\alpha^0 \subsetneq \Sigma_{\alpha+1}^0 \cap \Pi_{\alpha+1}^0$, for each countable ordinal $\alpha \geq 1$.
- (b) $\bigcup_{\gamma < \alpha} \Sigma_\gamma^0 = \bigcup_{\gamma < \alpha} \Pi_\gamma^0 \subsetneq \Sigma_\alpha^0 \cap \Pi_\alpha^0$, for each countable limit ordinal α .
- (c) A set $W \subseteq X^\omega$ is in the class Σ_α^0 iff its complement is in the class Π_α^0 .
- (d) $\Sigma_\alpha^0 - \Pi_\alpha^0 \neq \emptyset$ and $\Pi_\alpha^0 - \Sigma_\alpha^0 \neq \emptyset$ hold for every countable ordinal $\alpha \geq 1$.

For a countable ordinal α , a subset of X^ω is a Borel set of rank α iff it is in $\Sigma_\alpha^0 \cup \Pi_\alpha^0$ but not in $\bigcup_{\gamma < \alpha} (\Sigma_\gamma^0 \cup \Pi_\gamma^0)$.

There are also some subsets of X^ω which are not Borel. Indeed there exists another hierarchy beyond the Borel hierarchy, which is called the projective hierarchy and which is obtained from the Borel hierarchy by successive applications of operations of projection and complementation. The first level of the projective hierarchy is formed by the class of *analytic sets* and the class of *co-analytic sets* which are complements of analytic sets. In particular the class of Borel subsets of X^ω is strictly included into the class Σ_1^1 of *analytic sets* which are obtained by projection of Borel sets.

Definition 13. A subset A of X^ω is in the class Σ_1^1 of **analytic sets** iff there exist a finite alphabet Y and a Borel subset B of $(X \times Y)^\omega$ such that $x \in A \leftrightarrow \exists y \in Y^\omega$ such that $(x, y) \in B$, where (x, y) is the infinite word over the alphabet $X \times Y$ such that $(x, y)(i) = (x(i), y(i))$ for each integer $i \geq 1$.

Remark 14. In the above definition we could take B in the class Π_2^0 . Moreover analytic subsets of X^ω are the projections of Π_1^0 -subsets of $X^\omega \times \omega^\omega$, where ω^ω is the Baire space, [60].

We now define completeness with regard to reduction by continuous functions. For a countable ordinal $\alpha \geq 1$, a set $F \subseteq X^\omega$ is said to be a Σ_α^0 (respectively, Π_α^0 , Σ_1^1)-complete set iff for any set $E \subseteq Y^\omega$ (with Y a finite alphabet): $E \in \Sigma_\alpha^0$ (respectively, $E \in \Pi_\alpha^0$, $E \in \Sigma_1^1$) iff there exists a continuous function $f : Y^\omega \rightarrow X^\omega$ such that $E = f^{-1}(F)$. Σ_n^0 (respectively Π_n^0)-complete sets, with n an integer ≥ 1 , are thoroughly characterized in [76].

In particular $\mathcal{R} = (0^*.1)^\omega$ is a well known example of Π_2^0 -complete subset of $\{0, 1\}^\omega$. It is the set of ω -words over $\{0, 1\}$ having infinitely many occurrences of the letter 1. Its complement $\{0, 1\}^\omega - (0^*.1)^\omega$ is a Σ_2^0 -complete subset of $\{0, 1\}^\omega$.

We recall now the definition of the arithmetical hierarchy of ω -languages which form the effective analogue to the hierarchy of Borel sets of finite rank.

Let X be a finite alphabet. An ω -language $L \subseteq X^\omega$ belongs to the class Σ_n if and only if there exists a recursive relation $R_L \subseteq (\mathbb{N})^{n-1} \times X^*$ such that

$$L = \{\sigma \in X^\omega \mid Q_1 a_1 Q_2 a_2 \dots Q_n a_n \quad (a_1, \dots, a_{n-1}, \sigma[a_n + 1]) \in R_L\}$$

where Q_1 is the existential quantifier \exists , and every other Q_i , for $2 \leq i \leq n$, is one of the quantifiers \forall or \exists (not necessarily in an alternating order). An ω -language $L \subseteq X^\omega$ belongs to the class Π_n if and only if its complement $X^\omega - L$ belongs to the class Σ_n . The inclusion relations that hold between the classes Σ_n and Π_n are the same as for the corresponding classes of the Borel hierarchy. The classes Σ_n and Π_n are included in the respective classes Σ_n^0 and Π_n^0 of the Borel hierarchy, and cardinality arguments suffice to show that these inclusions are strict.

As in the case of the Borel hierarchy, projections of arithmetical sets (of the second Π -class) lead beyond the arithmetical hierarchy, to the analytical hierarchy of ω -languages. The first class of this hierarchy is the class Σ_1^1 of *effective analytic sets* which are obtained by projection of arithmetical sets. An ω -language $L \subseteq X^\omega$ belongs to the class Σ_1^1 if and only if there exists a recursive relation $R_L \subseteq \mathbb{N} \times \{0, 1\}^* \times X^*$ such that:

$$L = \{\sigma \in X^\omega \mid \exists \tau (\tau \in \{0, 1\}^\omega \wedge \forall n \exists m ((n, \tau[m], \sigma[m]) \in R_L))\}$$

Then an ω -language $L \subseteq X^\omega$ is in the class Σ_1^1 iff it is the projection of an ω -language over the alphabet $X \times \{0, 1\}$ which is in the class Π_2 . The class Π_1^1 of *effective co-analytic sets* is simply the class of complements of effective analytic sets. We denote as usual $\Delta_1^1 = \Sigma_1^1 \cap \Pi_1^1$.

Recall that an ω -language $L \subseteq X^\omega$ is in the class Σ_1^1 iff it is accepted by a non deterministic Turing machine (reading ω -words) with a Büchi or Muller acceptance condition [78].

The Borel ranks of Δ_1^1 sets are the (recursive) ordinals $\gamma < \omega_1^{\text{CK}}$, where ω_1^{CK} is the first non-recursive ordinal, usually called the Church-Kleene ordinal. Moreover, for every non null ordinal $\alpha < \omega_1^{\text{CK}}$, there exist some Σ_α^0 -complete and some Π_α^0 -complete sets in the class Δ_1^1 .

On the other hand, Kechris, Marker and Sami proved in [51] that the supremum of the set of Borel ranks of (effective) Σ_1^1 -sets is the ordinal γ_2^1 . This ordinal is proved to be strictly greater than the ordinal δ_2^1 which is the first non Δ_2^1 ordinal. In particular, the ordinal γ_2^1 is strictly greater than the ordinal ω_1^{CK} . Remark that the exact value of the ordinal γ_2^1 may depend on axioms of set theory, see [51, 41] for more details. Notice also that it seems still unknown whether *every* non null ordinal $\gamma < \gamma_2^1$ is the Borel rank of a Σ_1^1 -set.

3.2 Wadge hierarchy

We now introduce the Wadge hierarchy, which is a great refinement of the Borel hierarchy defined via reductions by continuous functions, [23, 83].

Definition 15 (Wadge [83]). *Let X, Y be two finite alphabets. For $L \subseteq X^\omega$ and $L' \subseteq Y^\omega$, L is said to be Wadge reducible to L' ($L \leq_W L'$) iff there exists a continuous function $f : X^\omega \rightarrow Y^\omega$, such that $L = f^{-1}(L')$. L and L' are Wadge equivalent iff $L \leq_W L'$ and $L' \leq_W L$. This will be denoted by $L \equiv_W L'$. And we shall say that $L <_W L'$ iff $L \leq_W L'$ but not $L' \leq_W L$. A set $L \subseteq X^\omega$ is said to be self dual iff $L \equiv_W L^-$, and otherwise it is said to be non self dual.*

The relation \leq_W is reflexive and transitive, and \equiv_W is an equivalence relation.

The *equivalence classes* of \equiv_W are called *Wadge degrees*.

The Wadge hierarchy WH is the class of Borel subsets of a set X^ω , where X is a finite set, equipped with \leq_W and with \equiv_W .

For $L \subseteq X^\omega$ and $L' \subseteq Y^\omega$, if $L \leq_W L'$ and $L = f^{-1}(L')$ where f is a continuous function from X^ω into Y^ω , then f is called a continuous reduction of L to L' . Intuitively it means that L is less complicated than L' because to check whether $x \in L$ it suffices to check whether $f(x) \in L'$ where f is a continuous function. Hence the Wadge degree of an ω -language is a measure of its topological complexity.

Notice that in the above definition, we consider that a subset $L \subseteq X^\omega$ is given together with the alphabet X . This is important as it is shown by the following simple example.

Let $L_1 = \{0, 1\}^\omega \subseteq \{0, 1\}^\omega$ and $L_2 = \{0, 1\}^\omega \subseteq \{0, 1, 2\}^\omega$. So the languages L_1 and L_2 are equal but considered over the different alphabets $X_1 = \{0, 1\}$ and $X_2 = \{0, 1, 2\}$. It turns out that $L_1 <_W L_2$. In fact L_1 is open *and* closed in X_1^ω while L_2 is closed but non open in X_2^ω .

We can now define the *Wadge class* of a set L :

Definition 16. *Let L be a subset of X^ω . The Wadge class of L is :*

$$[L] = \{L' \mid L' \subseteq Y^\omega \text{ for a finite alphabet } Y \text{ and } L' \leq_W L\}.$$

Recall that each Borel class Σ_α^0 and Π_α^0 is a Wadge class.

A set $L \subseteq X^\omega$ is a Σ_α^0 (respectively Π_α^0)-complete set iff for any set $L' \subseteq Y^\omega$, L' is in Σ_α^0 (respectively Π_α^0) iff $L' \leq_W L$. It follows from the study of the Wadge hierarchy

that a set $L \subseteq X^\omega$ is a Σ_α^0 (respectively, Π_α^0)-complete set iff it is in Σ_α^0 but not in Π_α^0 (respectively, in Π_α^0 but not in Σ_α^0).

There is a close relationship between Wadge reducibility and games which we now introduce.

Definition 17. Let $L \subseteq X^\omega$ and $L' \subseteq Y^\omega$. The Wadge game $W(L, L')$ is a game with perfect information between two players, player 1 who is in charge of L and player 2 who is in charge of L' .

Player 1 first writes a letter $a_1 \in X$, then player 2 writes a letter $b_1 \in Y$, then player 1 writes a letter $a_2 \in X$, and so on.

The two players alternatively write letters a_n of X for player 1 and b_n of Y for player 2.

After ω steps, player 1 has written an ω -word $a \in X^\omega$ and player 2 has written an ω -word $b \in Y^\omega$. Player 2 is allowed to skip, even infinitely often, provided he really writes an ω -word in ω steps.

Player 2 wins the play iff $[a \in L \leftrightarrow b \in L']$, i.e. iff:

$$[(a \in L \text{ and } b \in L') \text{ or } (a \notin L \text{ and } b \notin L' \text{ and } b \text{ is infinite})].$$

Recall that a strategy for player 1 is a function $\sigma : (Y \cup \{s\})^* \rightarrow X$. And a strategy for player 2 is a function $f : X^+ \rightarrow Y \cup \{s\}$.

σ is a winning strategy for player 1 iff he always wins a play when he uses the strategy σ , i.e. when the n^{th} letter he writes is given by $a_n = \sigma(b_1 \dots b_{n-1})$, where b_i is the letter written by player 2 at step i and $b_i = s$ if player 2 skips at step i .

A winning strategy for player 2 is defined in a similar manner.

Martin's Theorem states that every Gale-Stewart Game $G(B)$, with B a Borel set, is determined, i.e. that one of the two players has a winning strategy in the game $G(B)$, see [50]. This implies the following determinacy result :

Theorem 18 (Wadge). Let $L \subseteq X^\omega$ and $L' \subseteq Y^\omega$ be two Borel sets, where X and Y are finite alphabets. Then the Wadge game $W(L, L')$ is determined : one of the two players has a winning strategy. And $L \leq_W L'$ iff player 2 has a winning strategy in the game $W(L, L')$.

Theorem 19 (Wadge). Up to the complement and \equiv_W , the class of Borel subsets of X^ω , for a finite alphabet X , is a well ordered hierarchy. There is an ordinal $|WH|$, called the length of the hierarchy, and a map d_W^0 from WH onto $|WH| - \{0\}$, such that for all $L, L' \subseteq X^\omega$:

$$\begin{aligned} d_W^0 L < d_W^0 L' &\leftrightarrow L <_W L' \text{ and} \\ d_W^0 L = d_W^0 L' &\leftrightarrow [L \equiv_W L' \text{ or } L \equiv_W L'^-]. \end{aligned}$$

The Wadge hierarchy of Borel sets of finite rank has length ${}^1\varepsilon_0$ where ${}^1\varepsilon_0$ is the limit of the ordinals α_n defined by $\alpha_1 = \omega_1$ and $\alpha_{n+1} = \omega_1^{\alpha_n}$ for n a non negative integer, ω_1 being the first non countable ordinal. Then ${}^1\varepsilon_0$ is the first fixed point of the ordinal exponentiation of base ω_1 . The length of the Wadge hierarchy of Borel sets in $\Delta_\omega^0 = \Sigma_\omega^0 \cap \Pi_\omega^0$ is the ω_1^{th} fixed point of the ordinal exponentiation of base ω_1 , which is a

much larger ordinal. The length of the whole Wadge hierarchy of Borel sets is a huge ordinal, with regard to the ω_1^{th} fixed point of the ordinal exponentiation of base ω_1 . It is described in [83, 23] by the use of the Veblen functions.

4 Topological complexity of context-free ω -languages

We recall first results about the topological complexity of regular ω -languages. Topological properties of regular ω -languages were first studied by L. H. Landweber in [52] where he characterized regular ω -languages in a given Borel class. It turned out that a regular ω -language is a Π_2^0 -set iff it is accepted by a deterministic Büchi automaton. On the other hand Mc Naughton's Theorem implies that regular ω -languages, accepted by deterministic Muller automata, are boolean combinations of regular ω -languages accepted by deterministic Büchi automata. Thus they are boolean combinations of Π_2^0 -sets hence Δ_3^0 -sets. Moreover Landweber proved that one can effectively determine the exact level of a given regular ω -language with regard to the Borel hierarchy.

A great improvement of these results was obtained by Wagner who determined in an effective way, using the notions of chains and superchains, the Wadge hierarchy of the class REG_ω , [84]. This hierarchy has length ω^ω and is now called the Wagner hierarchy, [69, 71, 72, 70, 78]. Wilke and Yoo proved in [86] that one can compute in polynomial time the Wadge degree of a regular ω -language. Later Carton and Perrin gave a presentation of the Wagner hierarchy using algebraic notions of ω -semigroups, [14, 13, 67]. This work was completed by Duparc and Riss in [27].

Context-free ω -languages beyond the class Δ_3^0 have been constructed for the first time in [32]. The construction used an operation of exponentiation of sets of finite or infinite words introduced by Duparc in his study of the Wadge hierarchy [23]. We are going now to recall these constructions although some stronger results on the topological complexity of context-free ω -languages were obtained later in [38, 41] by other methods. However the methods of [32] using Duparc's operation of exponentiation are also interesting and it gave other results on ambiguity and on ω -powers of context-free languages we can not (yet ?) get by other methods, see Sections 6 and 7 below.

Wadge gave a description of the Wadge hierarchy of Borel sets in [83]. Duparc recently got a new proof of Wadge's results and gave in [22, 23] a normal form of Borel sets in the class Δ_ω^0 , i.e. an inductive construction of a Borel set of every given degree smaller than the ω_1^{th} fixed point of the ordinal exponentiation of base ω_1 . The construction relies on set theoretic operations which are the counterpart of arithmetical operations over ordinals needed to compute the Wadge degrees.

Actually Duparc studied the Wadge hierarchy via the study of the conciliating hierarchy. Conciliating sets are sets of finite *or* infinite words over an alphabet X , i.e. subsets of $X^* \cup X^\omega = X^{\leq \omega}$. It turned out that the conciliating hierarchy is isomorphic to the Wadge hierarchy of non-self-dual Borel sets, via the correspondence $A \rightarrow A^d$ we recall now:

For a word $x \in (X \cup \{d\})^{\leq \omega}$ we denote by $x(/d)$ the sequence obtained from x by removing every occurrence of the letter d . Then for $A \subseteq X^{\leq \omega}$ and d a letter not in X , A^d is the ω -language over $X \cup \{d\}$ which is defined by :

$$A^d = \{x \in (X \cup \{d\})^\omega \mid x(/d) \in A\}.$$

We are going now to introduce the operation of exponentiation of conciliating sets.

Definition 20 (Duparc [23]). Let X be a finite alphabet, $\leftarrow \notin X$, and let x be a finite or infinite word over the alphabet $Y = X \cup \{\leftarrow\}$.

Then x^{\leftarrow} is inductively defined by:

$$\lambda^{\leftarrow} = \lambda,$$

and for a finite word $u \in (X \cup \{\leftarrow\})^*$:

$$(u.a)^{\leftarrow} = u^{\leftarrow}.a, \text{ if } a \in X,$$

$$(u.\leftarrow)^{\leftarrow} = u^{\leftarrow}(1).u^{\leftarrow}(2) \dots u^{\leftarrow}(|u^{\leftarrow}| - 1) \text{ if } |u^{\leftarrow}| > 0,$$

$$(u.\leftarrow)^{\leftarrow} = \lambda \text{ if } |u^{\leftarrow}| = 0,$$

and for u infinite:

$$(u)^{\leftarrow} = \lim_{n \in \omega} (u[n])^{\leftarrow}, \text{ where, given } \beta_n \text{ and } v \text{ in } X^*,$$

$$v \sqsubseteq \lim_{n \in \omega} \beta_n \leftrightarrow \exists n \forall p \geq n \quad \beta_p[[v]] = v.$$

(The finite or infinite word $\lim_{n \in \omega} \beta_n$ is determined by the set of its (finite) prefixes).

Remark 21. For $x \in Y^{\leq \omega}$, x^{\leftarrow} denotes the string x , once every \leftarrow occurring in x has been “evaluated” to the back space operation, proceeding from left to right inside x . In other words $x^{\leftarrow} = x$ from which every interval of the form “ $a \leftarrow$ ” ($a \in X$) is removed.

For example if $u = (a \leftarrow)^n$, for n an integer ≥ 1 , or $u = (a \leftarrow)^\omega$, or $u = (a \leftarrow \leftarrow)^\omega$, then $(u)^{\leftarrow} = \lambda$. If $u = (ab \leftarrow)^\omega$ then $(u)^{\leftarrow} = a^\omega$ and if $u = bb(\leftarrow a)^\omega$ then $(u)^{\leftarrow} = b$.

Let us notice that in Definition 20 the limit is not defined in the usual way:

for example if $u = bb(\leftarrow a)^\omega$ the finite word $u[n]^{\leftarrow}$ is alternatively equal to b or to ba : more precisely $u[2n+1]^{\leftarrow} = b$ and $u[2n+2]^{\leftarrow} = ba$ for every integer $n \geq 1$ (it holds also that $u[1]^{\leftarrow} = b$ and $u[2]^{\leftarrow} = bb$). Thus Definition 20 implies that $\lim_{n \in \omega} (u[n])^{\leftarrow} = b$ so $u^{\leftarrow} = b$.

We can now define the operation $A \rightarrow A^\sim$ of exponentiation of conciliating sets:

Definition 22 (Duparc [23]). For $A \subseteq X^{\leq \omega}$ and $\leftarrow \notin X$, let

$$A^\sim =_{df} \{x \in (X \cup \{\leftarrow\})^{\leq \omega} \mid x^{\leftarrow} \in A\}.$$

The operation \sim is monotone with regard to the Wadge ordering and produces some sets of higher complexity.

Theorem 23 (Duparc [23]). Let $A \subseteq X^{\leq \omega}$ and $n \geq 1$. if $A^d \subseteq (X \cup \{d\})^\omega$ is a Σ_n^0 -complete (respectively, Π_n^0 -complete) set, then $(A^\sim)^d$ is a Σ_{n+1}^0 -complete (respectively, Π_{n+1}^0 -complete) set.

It was proved in [32] that the class of context-free infinitary languages (which are unions of a context-free finitary language and of a context-free ω -language) is closed under the

operation $A \rightarrow A^\sim$. On the other hand $A \rightarrow A^d$ is an operation from the class of context-free infinitary languages into the class of context-free ω -languages. This implies that, for each integer $n \geq 1$, there exist some context-free ω -languages which are Σ_n^0 -complete and some others which are Π_n^0 -complete.

Theorem 24 ([32]). *For each non negative integer $n \geq 1$, there exist Σ_n^0 -complete context-free ω -languages A_n and Π_n^0 -complete context-free ω -languages B_n .*

Proof. For $n = 1$ consider the Σ_1^0 -complete regular ω -language

$$A_1 = \{\alpha \in \{0, 1\}^\omega \mid \exists i \ \alpha(i) = 1\}$$

and the Π_1^0 -complete regular ω -language

$$B_1 = \{\alpha \in \{0, 1\}^\omega \mid \forall i \ \alpha(i) = 0\}.$$

These languages are context-free ω -languages because $REG_\omega \subseteq CFL_\omega$.

Now consider the Σ_2^0 -complete regular ω -language

$$A_2 = \{\alpha \in \{0, 1\}^\omega \mid \exists^{<\omega} i \ \alpha(i) = 1\}$$

and the Π_2^0 -complete regular ω -language

$$B_2 = \{\alpha \in \{0, 1\}^\omega \mid \exists^\omega i \ \alpha(i) = 0\},$$

where $\exists^{<\omega} i$ means: "there exist only finitely many i such that ...", and

$\exists^\omega i$ means: "there exist infinitely many i such that ...".

A_2 and B_2 are context-free ω -languages because they are regular ω -languages.

To obtain context-free ω -languages of greater Borel ranks, consider now O_1 (respectively, C_1) subsets of $\{0, 1\}^{\leq\omega}$ such that $(O_1)^d$ (respectively, $(C_1)^d$) are Σ_1^0 -complete (respectively Π_1^0 -complete).

For example $O_1 = \{x \in \{0, 1\}^{\leq\omega} \mid \exists i \ x(i) = 1\}$ and $C_1 = \{\lambda\}$.

We can apply $n \geq 1$ times the operation of exponentiation of sets.

More precisely, we define, for a set $A \subseteq X^{\leq\omega}$:

$$A^{\sim.0} = A$$

$$A^{\sim.1} = A^\sim \text{ and}$$

$$A^{\sim.(n+1)} = (A^{\sim.n})^\sim.$$

Now apply n times (for an integer $n \geq 1$) the operation \sim (with different new letters $\leftarrow_1, \leftarrow_2, \leftarrow_3, \dots, \leftarrow_n$) to O_1 and C_1 .

By Theorem 23, it holds that for an integer $n \geq 1$:

$$(O_1^{\sim.n})^d \text{ is a } \Sigma_{n+1}^0\text{-complete subset of } \{0, 1, \leftarrow_1, \dots, \leftarrow_n, d\}^\omega.$$

$$(C_1^{\sim.n})^d \text{ is a } \Pi_{n+1}^0\text{-complete subset of } \{0, 1, \leftarrow_1, \dots, \leftarrow_n, d\}^\omega.$$

And it is easy to see that O_1 and C_1 are in the form $E \cup F$ where E is a finitary context-free language and F is a context-free ω -language. Then the ω -languages $(O_1^{\sim.n})^d$ and $(C_1^{\sim.n})^d$ are context-free. Hence the class CFL_ω exhausts the finite ranks of the Borel hierarchy: we obtain the context-free ω -languages $A_n = (O_1^{\sim.(n-1)})^d$ and $B_n = (C_1^{\sim.(n-1)})^d$, for $n \geq 3$. \square

This gave a partial answer to questions of Thomas and Lescow [55] about the hierarchy of context-free ω -languages.

A natural question now arose: Do the decidability results of [52] extend to context-free ω -languages? Unfortunately the answer is no. Cohen and Gold proved that one cannot decide whether a given context-free ω -language is in the class Π_1^0 , Σ_1^0 , or Π_2^0 , [19]. This result was first extended to all classes Σ_n^0 and Π_n^0 , for n an integer ≥ 1 , using the undecidability of the Post Correspondence Problem, [32].

Later, the coding of an infinite number of erasers \leftarrow_n , $n \geq 1$, and an iteration of the operation of exponentiation were used to prove that there exist some context-free ω -languages which are Borel of infinite rank, [36].

Using the correspondences between the operation of exponentiation of sets and the ordinal exponentiation of base ω_1 , and between the Wadge's operation of sum of sets, [83, 23], and the ordinal sum, it was proved in [33] that the length of the Wadge hierarchy of the class CFL_ω is at least ε_0 , the first fixed point of the ordinal exponentiation of base ω . Next were constructed some Δ_ω^0 context-free ω -languages in ε_ω Wadge degrees, where ε_ω is the ω^{th} fixed point of the ordinal exponentiation of base ω , and also some Σ_ω^0 -complete context-free ω -languages, [31, 39]. Notice that the Wadge hierarchy of *non-deterministic* context-free ω -languages is not effective, [33].

The question of the existence of non-Borel context-free ω -languages was solved by Finkel and Ressayre. Using a coding of infinite binary trees labeled in a finite alphabet X , it was proved that there exist some non-Borel, and even Σ_1^1 -complete, context-free ω -languages, and that one cannot decide whether a given context-free ω -language is a Borel set, [35]. Amazingly there is a simple finitary language V accepted by a 1-counter automaton such that V^ω is Σ_1^1 -complete; we shall recall it in Section 7 below on ω -powers.

But a complete and very surprising result was obtained in [38, 41], which extended previous results. A simulation of multicounter automata by 1-counter automata was used in [38, 41]. We firstly recall now the definition of these automata, in order to sketch the constructions involved in these simulations.

Definition 25. *Let k be an integer ≥ 1 . A k -counter machine (k -CM) is a 4-tuple $\mathcal{M}=(K, X, \Delta, q_0)$, where K is a finite set of states, X is a finite input alphabet, $q_0 \in K$ is the initial state, and $\Delta \subseteq K \times (X \cup \{\lambda\}) \times \{0, 1\}^k \times K \times \{0, 1, -1\}^k$ is the transition relation. The k -counter machine \mathcal{M} is said to be real time iff: $\Delta \subseteq K \times X \times \{0, 1\}^k \times K \times \{0, 1, -1\}^k$, i.e. iff there are not any λ -transitions.*

If the machine \mathcal{M} is in state q and $c_i \in \mathbf{N}$ is the content of the i^{th} counter \mathcal{C}_i then the configuration (or global state) of \mathcal{M} is the $(k+1)$ -tuple (q, c_1, \dots, c_k) .

For $a \in X \cup \{\lambda\}$, $q, q' \in K$ and $(c_1, \dots, c_k) \in \mathbf{N}^k$ such that $c_j = 0$ for $j \in E \subseteq \{1, \dots, k\}$ and $c_j > 0$ for $j \notin E$, if $(q, a, i_1, \dots, i_k, q', j_1, \dots, j_k) \in \Delta$ where $i_j = 0$ for $j \in E$ and $i_j = 1$ for $j \notin E$, then we write:

$$a : (q, c_1, \dots, c_k) \mapsto_{\mathcal{M}} (q', c_1 + j_1, \dots, c_k + j_k)$$

Thus we see that the transition relation must satisfy:

if $(q, a, i_1, \dots, i_k, q', j_1, \dots, j_k) \in \Delta$ and $i_m = 0$ for some $m \in \{1, \dots, k\}$, then $j_m = 0$ or $j_m = 1$ (but j_m cannot be equal to -1).

Let $\sigma = a_1 a_2 \dots a_n \dots$ be an ω -word over X . An ω -sequence of configurations $r = (q_i, c_1^i, \dots, c_k^i)_{i \geq 1}$ is called a run of \mathcal{M} on σ , starting in configuration (p, c_1, \dots, c_k) , iff:

- (1) $(q_1, c_1^1, \dots, c_k^1) = (p, c_1, \dots, c_k)$
- (2) for each $i \geq 1$, there exists $b_i \in X \cup \{\lambda\}$ such that $b_i : (q_i, c_1^i, \dots, c_k^i) \mapsto_{\mathcal{M}} (q_{i+1}, c_1^{i+1}, \dots, c_k^{i+1})$ such that either $a_1 a_2 \dots a_n \dots = b_1 b_2 \dots b_n \dots$ or $b_1 b_2 \dots b_n \dots$ is a finite prefix of $a_1 a_2 \dots a_n \dots$.

The run r is said to be complete when $a_1 a_2 \dots a_n \dots = b_1 b_2 \dots b_n \dots$.

For every such run, $\text{In}(r)$ is the set of all states entered infinitely often during run r .

A complete run r of \mathcal{M} on σ , starting in configuration $(q_0, 0, \dots, 0)$, will be simply called “a run of \mathcal{M} on σ ”.

Definition 26. A Büchi k -counter automaton is a 5-tuple $\mathcal{M} = (K, X, \Delta, q_0, F)$, where $\mathcal{M}' = (K, X, \Delta, q_0)$ is a k -counter machine and $F \subseteq K$ is the set of accepting states. The ω -language accepted by \mathcal{M} is

$$L(\mathcal{M}) = \{\sigma \in X^\omega \mid \text{there exists a run } r \text{ of } \mathcal{M} \text{ on } \sigma \text{ such that } \text{In}(r) \cap F \neq \emptyset\}$$

The notion of Muller k -counter automaton is defined in a similar way. One can see that an ω -language is accepted by a (real time) Büchi k -counter automaton iff it is accepted by a (real time) Muller k -counter automaton [29]. Notice that this result is no longer true in the deterministic case.

We denote $\mathbf{BC}(k)$ (respectively, $\mathbf{r-BC}(k)$) the class of Büchi k -counter automata (respectively, of real time Büchi k -counter automata).

We denote $\mathbf{BCL}(k)_\omega$ (respectively, $\mathbf{r-BCL}(k)_\omega$) the class of ω -languages accepted by Büchi k -counter automata (respectively, by real time Büchi k -counter automata).

Remark that 1-counter automata introduced above are equivalent to pushdown automata whose stack alphabet is in the form $\{Z_0, A\}$ where Z_0 is the bottom symbol which always remains at the bottom of the stack and appears only there and A is another stack symbol. The pushdown stack may be seen like a counter whose content is the integer N if the stack content is the word $A^N.Z_0$.

In the model introduced here the counter value cannot be increased by more than 1 during a single transition. However this does not change the class of ω -languages accepted by such automata. So the class $\mathbf{BCL}(1)_\omega$ is equal to the class $\mathbf{1-ICL}_\omega$, introduced in [33], and it is a strict subclass of the class \mathbf{CFL}_ω of context-free ω -languages accepted by Büchi pushdown automata.

We state now the surprising result proved in [41], using multicounter-automata.

Theorem 27 ([41]). *The Wadge hierarchy of the class $\mathbf{r-BCL}(1)_\omega$, hence also of the class \mathbf{CFL}_ω , or of every class \mathcal{C} such that $\mathbf{r-BCL}(1)_\omega \subseteq \mathcal{C} \subseteq \Sigma_1^1$, is the Wadge*

hierarchy of the class Σ_1^1 of ω -languages accepted by Turing machines with a Büchi acceptance condition.

We now sketch the proof of this result. It is well known that every Turing machine can be simulated by a (non real time) 2-counter automaton, see [49]. Thus the Wadge hierarchy of the class $\mathbf{BCL}(2)_\omega$ is also the Wadge hierarchy of the class of ω -languages accepted by Büchi Turing machines.

One can then find, from an ω -language $L \subseteq X^\omega$ in $\mathbf{BCL}(2)_\omega$, another ω -language $\theta_S(L)$ which will be of the same topological complexity but accepted by a *real-time* 8-counter Büchi automaton. The idea is to add firstly a storage type called a queue to a 2-counter Büchi automaton in order to read ω -words in real-time. Then the queue can be simulated by two pushdown stacks or by four counters. This simulation is not done in real-time but a crucial fact is that one can bound the number of transitions needed to simulate the queue. This allows to pad the strings in L with enough extra letters so that the new words will be read in real-time by a 8-counter Büchi automaton. The padding is obtained via the function θ_S which we define now.

Let X be an alphabet having at least two letters, E be a new letter not in X , S be an integer ≥ 1 , and $\theta_S : X^\omega \rightarrow (X \cup \{E\})^\omega$ be the function defined, for all $x \in X^\omega$, by:

$$\theta_S(x) = x(1).E^S.x(2).E^{S^2}.x(3).E^{S^3}.x(4) \dots x(n).E^{S^n}.x(n+1).E^{S^{n+1}} \dots$$

It turns out that if $L \subseteq X^\omega$ is in $\mathbf{BCL}(2)_\omega$ then there exists an integer $S \geq 1$ such that $\theta_S(L)$ is in the class $\mathbf{r-BCL}(8)_\omega$, and, except for some special few cases, $\theta_S(L) \equiv_W L$.

The next step is to simulate a *real-time* 8-counter Büchi automaton, using only a *real-time* 1-counter Büchi automaton.

Consider the product of the eight first prime numbers:

$$K = 2 \times 3 \times 5 \times 7 \times 11 \times 13 \times 17 \times 19 = 9699690$$

Then an ω -word $x \in X^\omega$ can be coded by the ω -word

$$h(x) = A.0^K.x(1).B.0^{K^2}.A.0^{K^2}.x(2).B.0^{K^3}.A.0^{K^3}.x(3).B \dots B.0^{K^n}.A.0^{K^n}.x(n).B \dots$$

where A , B and 0 are new letters not in X . The mapping $h : X^\omega \rightarrow (X \cup \{A, B, 0\})^\omega$ is continuous. It is easy to see that the ω -language $h(X^\omega)^-$ is an open subset of $(X \cup \{A, B, 0\})^\omega$ and that it is in the class $\mathbf{r-BCL}(1)_\omega$.

If $L(\mathcal{A}) \subseteq X^\omega$ is accepted by a real time 8-counter Büchi automaton \mathcal{A} , then one can construct effectively from \mathcal{A} a 1-counter Büchi automaton \mathcal{B} , reading words over the alphabet $X \cup \{A, B, 0\}$, such that $L(\mathcal{A}) = h^{-1}(L(\mathcal{B}))$, i.e.

$$\forall x \in X^\omega \quad h(x) \in L(\mathcal{B}) \iff x \in L(\mathcal{A})$$

In fact, the simulation, during the reading of $h(x)$ by the 1-counter Büchi automaton \mathcal{B} , of the behaviour of the real time 8-counter Büchi automaton \mathcal{A} reading x , can be achieved, using the coding of the content (c_1, c_2, \dots, c_8) of eight counters by the product $2^{c_1} \times 3^{c_2} \times \dots \times (17)^{c_7} \times (19)^{c_8}$, and the **special shape** of ω -words in $h(X^\omega)$ which allows the propagation of the value of the counters of \mathcal{A} . A crucial fact here is that $h(X^\omega)^-$ is in the class $\mathbf{r-BCL}(1)_\omega$. Thus the ω -language

$$h(L(\mathcal{A})) \cup h(X^\omega)^- = L(\mathcal{B}) \cup h(X^\omega)^-$$

is in the class $\mathbf{BCL}(1)_\omega$ and it has the same topological complexity as the ω -language $L(\mathcal{A})$, (except the special few cases where $d_W(L(\mathcal{A})) \leq \omega$).

One can see, from the construction of \mathcal{B} , that at most $(K-1)$ consecutive λ -transitions can occur during the reading of an ω -word x by \mathcal{B} . It is then easy to see that the ω -language $\phi(h(L(\mathcal{A})) \cup h(X^\omega)^-)$ is an ω -language in the class $\mathbf{r-BCL}(1)_\omega$ which has the same topological complexity as the ω -language $L(\mathcal{A})$, where ϕ is the mapping from $(X \cup \{A, B, 0\})^\omega$ into $(X \cup \{A, B, F, 0\})^\omega$, with F a new letter, which is defined by:

$$\phi(x) = F^{K-1}.x(1).F^{K-1}.x(2).F^{K-1}.x(3) \dots F^{K-1}.x(n).F^{K-1}.x(n+1).F^{K-1} \dots$$

Altogether these constructions are used in [41] to prove Theorem 27. As the Wadge hierarchy is a refinement of the Borel hierarchy and, for any countable ordinal α , Σ_α^0 -complete sets (respectively, Π_α^0 -complete sets) form a single Wadge degree, this implies also the following result.

Theorem 28. *Let \mathcal{C} be a class of ω -languages such that:*

$$\mathbf{r-BCL}(1)_\omega \subseteq \mathcal{C} \subseteq \Sigma_1^1.$$

- (a) *The Borel hierarchy of the class \mathcal{C} is equal to the Borel hierarchy of the class Σ_1^1 .*
- (b) *$\gamma_2^1 = \text{Sup} \{ \alpha \mid \exists L \in \mathcal{C} \text{ such that } L \text{ is a Borel set of rank } \alpha \}$.*
- (c) *For every non null ordinal $\alpha < \omega_1^{\text{CK}}$, there exists some Σ_α^0 -complete and some Π_α^0 -complete ω -languages in the class \mathcal{C} .*

Notice that similar methods have next be used to get another surprising result: the Wadge hierarchy, hence also the Borel hierarchy, of infinitary rational relations accepted by 2-tape Büchi automata is equal to the Wadge hierarchy of the class $\mathbf{r-BCL}(1)_\omega$ or of the class Σ_1^1 , [42, 43].

5 Topological complexity of deterministic context-free ω -languages

We have seen in the previous section that all *non-deterministic* finite machines accept ω -languages of the same topological complexity, as soon as they can simulate a real time 1-counter automaton.

This result is still true in the *deterministic* case if we consider only the Borel hierarchy. Recall that regular ω -languages accepted by Büchi automata are Π_2^0 -sets and

ω -languages accepted by Muller automata are boolean combinations of Π_2^0 -sets hence Δ_3^0 -sets. Engelfriet and Hoogeboom proved that this result holds also for all ω -languages accepted by *deterministic* \mathbf{X} -automata, i.e. automata equipped with a storage type \mathbf{X} , including the cases of k -counter automata, pushdown automata, Petri nets, Turing machines. In particular, ω -languages accepted by deterministic Büchi Turing machines are Π_2^0 -sets and ω -languages accepted by deterministic Muller Turing machines are Δ_3^0 -sets.

It turned out that this is no longer true if we consider the much finer Wadge hierarchy to measure the complexity of ω -languages. The Wadge hierarchy is suitable to distinguish the accepting power of deterministic finite machines reading infinite words. Recall that the Wadge hierarchy of regular ω -languages, now called the Wagner hierarchy, has been effectively determined by Wagner; it has length ω^ω [84, 69, 70].

Its extension to *deterministic* context-free ω -languages has been determined by Duparc, its length is $\omega^{(\omega^2)}$ [26, 24]. To determine the Wadge hierarchy of the class $DCFL_\omega$, Duparc first defined operations on DMPDA which correspond to ordinal operations of sum, multiplication by ω , and multiplication by ω_1 , over Wadge degrees. In this way are constructed some DMPDA accepting ω -languages of every Wadge degree in the form :

$$d_W^0(A) = \omega_1^{n_j} \cdot \delta_j + \omega_1^{n_{j-1}} \cdot \delta_{j-1} + \dots + \omega_1^{n_1} \cdot \delta_1$$

where $j > 0$ is an integer, $n_j > n_{j-1} > \dots > n_1$ are integers ≥ 0 , and $\delta_j, \delta_{j-1}, \dots, \delta_1$ are non null ordinals $< \omega^\omega$.

On the other hand it is known that the Wadge degree α of a boolean combination of Π_2^0 -sets is smaller than the ordinal ω_1^ω thus it has a Cantor normal form :

$$\alpha = \omega_1^{n_j} \cdot \delta_j + \omega_1^{n_{j-1}} \cdot \delta_{j-1} + \dots + \omega_1^{n_1} \cdot \delta_1$$

where $j > 0$ is an integer, $n_j > n_{j-1} > \dots > n_1$ are integers ≥ 0 , and $\delta_j, \delta_{j-1}, \dots, \delta_1$ are non null ordinals $< \omega_1$, i.e. non null countable ordinals. In a second step it is proved in [24], using infinite multi-player games, that if such an ordinal α is the Wadge degree of a deterministic context-free ω -language, then all the ordinals $\delta_j, \delta_{j-1}, \dots, \delta_1$ appearing in its Cantor normal form are smaller than the ordinal $< \omega^\omega$. Thus the Wadge hierarchy of the class $DCFL_\omega$ is completely determined.

Theorem 29 (Duparc [24]). *The Wadge degrees of deterministic context-free ω -languages are exactly the ordinals in the form :*

$$\alpha = \omega_1^{n_j} \cdot \delta_j + \omega_1^{n_{j-1}} \cdot \delta_{j-1} + \dots + \omega_1^{n_1} \cdot \delta_1$$

where $j > 0$ is an integer, $n_j > n_{j-1} > \dots > n_1$ are integers ≥ 0 , and $\delta_j, \delta_{j-1}, \dots, \delta_1$ are non null ordinals $< \omega^\omega$.

The length of the Wadge hierarchy of the class $DCFL_\omega$ is the ordinal $(\omega^\omega)^\omega = \omega^{(\omega^2)}$.

Notice that the Wadge hierarchy of $DCFL_\omega$ is not determined in an effective way in [24]. The question of the decidability of problems like: “given two DMPDA \mathcal{A} and \mathcal{B} ,

does $L(\mathcal{A}) \leq_W L(\mathcal{B})$ hold ?” or “given a DMPDA \mathcal{A} can we compute $d_W^0(L(\mathcal{A}))$?” naturally arises.

Cohen and Gold proved that one can decide whether an effectively given ω -language in $DCFL_\omega$ is an open or a closed set [19]. Linna characterized the ω -languages accepted by DBPDA as the Π_2^0 -sets in $DCFL_\omega$ and proved in [58] that one can decide whether an effectively given ω -language accepted by a DMPDA is a Π_2^0 -set or a Σ_2^0 -set.

Using a recent result of Walukiewicz on infinite games played on pushdown graphs, [85], these decidability results were extended in [32] where it was proved that one can decide whether a *deterministic* context-free ω -language accepted by a given DMPDA is in a given Borel class Σ_1^0 , Π_1^0 , Σ_2^0 , or Π_2^0 or even in the wadge class $[L]$ given by any regular ω -language L .

An effective extension of the Wagner hierarchy to ω -languages accepted by Muller *deterministic* real time blind (i. e. without zero-test) 1-counter automata has been determined in [30]. Recall that blind 1-counter automata form a subclass of 1-counter automata hence also of pushdown automata. A blind 1-counter Muller automaton is just a Muller pushdown automaton $M = (K, X, \Gamma, \delta, q_0, Z_0, \mathcal{F})$ such that $\Gamma = \{Z_0, I\}$ where Z_0 is the bottom symbol and always remains at the bottom of the store. Moreover every transition which is enabled at zero level is also enabled at non zero level, i.e. if $\delta(q, a, Z_0) = (p, I^n Z_0)$, for some $p, q \in K$, $a \in X$ and $n \geq 0$, then $\delta(q, a, I) = (p, I^{n+1})$. But the converse may not be true, i.e. some transition may be enabled at non zero level but not at zero level. Notice that blind 1-counter automata are sometimes called partially blind 1-counter automata as in [47].

The Wadge hierarchy of blind counter ω -languages, accepted by deterministic Muller real time blind 1-counter automata (MBCA), is studied in [30] in a similar way as Wagner studied the Wadge hierarchy of regular ω -languages in [84]. Chains and superchains for MBCA are defined as Wagner did for Muller automata. The essential difference between the two hierarchies relies on the existence of superchains of transfinite length $\alpha < \omega^2$ for MBCA when in the case of Muller automata the superchains have only finite lengths. The hierarchy of ω -languages accepted by MBCA is effective and leads to effective winning strategies in Wadge games between two players in charge of ω -languages accepted by MBCA. Concerning the length of the Wadge hierarchy of MBCA the following result is proved :

Theorem 30 (Finkel [30]).

- (a) *The length of the Wadge hierarchy of blind counter ω -languages in Δ_2^0 is ω^2 .*
- (b) *The length of the Wadge hierarchy of blind counter ω -languages is the ordinal ω^ω (hence it is equal to the length of the Wagner hierarchy).*

Notice that the length of the Wadge hierarchy of blind counter ω -languages is equal to the length of the Wagner hierarchy although it is actually a strict extension of the Wagner hierarchy, as shown already in item (a) of the above theorem. The Wadge degrees of blind counter ω -languages are the ordinals in the form :

$$\alpha = \omega_1^{n_j} \cdot \delta_j + \omega_1^{n_{j-1}} \cdot \delta_{j-1} + \dots + \omega_1^{n_1} \cdot \delta_1$$

where $j > 0$ is an integer, $n_j > n_{j-1} > \dots > n_1$ are integers ≥ 0 , and $\delta_j, \delta_{j-1}, \dots, \delta_1$ are non null ordinals $< \omega^2$. Recall that in the case of Muller automata, the ordinals $\delta_j, \delta_{j-1}, \dots, \delta_1$ are non-negative integers, i.e. non null ordinals $< \omega$.

Notice that Selivanov has recently determined the Wadge hierarchy of ω -languages accepted by *deterministic* Turing machines; its length is $(\omega_1^{\text{CK}})^\omega$ [72, 71]. The ω -languages accepted by deterministic Muller Turing machines or equivalently which are boolean combinations of arithmetical Π_2^0 -sets have Wadge degrees in the form :

$$\alpha = \omega_1^{n_j} \cdot \delta_j + \omega_1^{n_{j-1}} \cdot \delta_{j-1} + \dots + \omega_1^{n_1} \cdot \delta_1$$

where $j > 0$ is an integer, $n_j > n_{j-1} > \dots > n_1$ are integers ≥ 0 , and $\delta_j, \delta_{j-1}, \dots, \delta_1$ are non null ordinals $< \omega_1^{\text{CK}}$.

6 Topology and ambiguity in context-free ω -languages

The notions of ambiguity and of degrees of ambiguity are well known and important in the study of context-free languages. These notions have been extended to context-free ω -languages accepted by Büchi or Muller pushdown automata in [34]. Notice that it is proved in [34] that these notions are independent of the Büchi or Muller acceptance condition. So in the sequel we shall only consider the Büchi acceptance condition.

We now firstly introduce a slight modification in the definition of a run of a Büchi pushdown automaton, which will be used in this section.

Definition 31. Let $\mathcal{A} = (K, X, \Gamma, \delta, q_0, Z_0, F)$ be a Büchi pushdown automaton. Let $\sigma = a_1 a_2 \dots a_n \dots$ be an ω -word over X . A run of \mathcal{A} on σ is an infinite sequence $r = (q_i, \gamma_i, \varepsilon_i)_{i \geq 1}$ where $(q_i, \gamma_i)_{i \geq 1}$ is an infinite sequence of configurations of \mathcal{A} and, for all $i \geq 1$, $\varepsilon_i \in \{0, 1\}$ and:

1. $(q_1, \gamma_1) = (q_0, Z_0)$
2. for each $i \geq 1$, there exists $b_i \in X \cup \{\lambda\}$ satisfying
 - $b_i : (q_i, \gamma_i) \mapsto_{\mathcal{A}} (q_{i+1}, \gamma_{i+1})$
 - and $(\varepsilon_i = 0 \text{ iff } b_i = \lambda)$
 - and such that $a_1 a_2 \dots a_n \dots = b_1 b_2 \dots b_n \dots$

As before the ω -language accepted by \mathcal{A} is

$$L(\mathcal{A}) = \{\sigma \in X^\omega \mid \text{there exists a run } r \text{ of } \mathcal{A} \text{ on } \sigma \text{ such that } I_n(r) \cap F \neq \emptyset\}$$

Notice that the numbers $\varepsilon_i \in \{0, 1\}$ are introduced in the above definition in order to distinguish runs of a BPDA which go through the same infinite sequence of configurations but for which λ -transitions do not occur at the same steps of the computations.

As usual the cardinal of ω is denoted \aleph_0 and the cardinal of the continuum is denoted 2^{\aleph_0} . The latter is also the cardinal of the set of real numbers or of the set X^ω for every finite alphabet X having at least two letters.

We are now ready to define degrees of ambiguity for BPDA and for context-free ω -languages.

Definition 32. Let \mathcal{A} be a BPDA reading infinite words over the alphabet X . For $x \in X^\omega$ let $\alpha_{\mathcal{A}}(x)$ be the cardinal of the set of accepting runs of \mathcal{A} on x .

Lemma 33 ([34]). Let \mathcal{A} be a BPDA reading infinite words over the alphabet X . Then for all $x \in X^\omega$ it holds that $\alpha_{\mathcal{A}}(x) \in \mathbb{N} \cup \{\aleph_0, 2^{\aleph_0}\}$.

Definition 34. Let \mathcal{A} be a BPDA reading infinite words over the alphabet X .

- (a) If $\sup\{\alpha_{\mathcal{A}}(x) \mid x \in X^\omega\} \in \mathbb{N} \cup \{2^{\aleph_0}\}$, then $\alpha_{\mathcal{A}} = \sup\{\alpha_{\mathcal{A}}(x) \mid x \in X^\omega\}$.
- (b) If $\sup\{\alpha_{\mathcal{A}}(x) \mid x \in X^\omega\} = \aleph_0$ and there is no word $x \in X^\omega$ such that $\alpha_{\mathcal{A}}(x) = \aleph_0$, then $\alpha_{\mathcal{A}} = \aleph_0^-$.
(\aleph_0^- does not represent a cardinal but is a new symbol that is introduced here to conveniently speak of this situation).
- (c) If $\sup\{\alpha_{\mathcal{A}}(x) \mid x \in X^\omega\} = \aleph_0$ and there exists (at least) one word $x \in X^\omega$ such that $\alpha_{\mathcal{A}}(x) = \aleph_0$, then $\alpha_{\mathcal{A}} = \aleph_0$.

Notice that for a BPDA \mathcal{A} , $\alpha_{\mathcal{A}} = 0$ iff \mathcal{A} does not accept any ω -word.

We shall consider below that $\mathbb{N} \cup \{\aleph_0^-, \aleph_0, 2^{\aleph_0}\}$ is linearly ordered by the relation $<$, which is defined by: $\forall k \in \mathbb{N}, k < k+1 < \aleph_0^- < \aleph_0 < 2^{\aleph_0}$.

Definition 35. For $k \in \mathbb{N} \cup \{\aleph_0^-, \aleph_0, 2^{\aleph_0}\}$ let

$$CFL_\omega(\alpha \leq k) = \{L(\mathcal{A}) \mid \mathcal{A} \text{ is a BPDA with } \alpha_{\mathcal{A}} \leq k\}$$

$$CFL_\omega(\alpha < k) = \{L(\mathcal{A}) \mid \mathcal{A} \text{ is a BPDA with } \alpha_{\mathcal{A}} < k\}$$

$NA-CFL_\omega = CFL_\omega(\alpha \leq 1)$ is the class of non ambiguous context-free ω -languages.

For every integer k such that $k \geq 2$, or $k \in \{\aleph_0^-, \aleph_0, 2^{\aleph_0}\}$,

$$A(k) - CFL_\omega = CFL_\omega(\alpha \leq k) - CFL_\omega(\alpha < k)$$

If $L \in A(k) - CFL_\omega$ with $k \in \mathbb{N}, k \geq 2$, or $k \in \{\aleph_0^-, \aleph_0, 2^{\aleph_0}\}$, then L is said to be inherently ambiguous of degree k .

Notice that one can define in a similar way the degree of ambiguity of a finitary context-free language. If M is a pushdown automaton accepting finite words by final states (or by final states and topmost stack letter) then $\alpha_M \in \mathbb{N}$ or $\alpha_M = \aleph_0^-$ or $\alpha_M = \aleph_0$. However every context-free language is accepted by a pushdown automaton M with $\alpha_M \leq \aleph_0^-$, [3]. We denote the class of non ambiguous context-free languages by $NA-CFL$ and the class of inherently ambiguous context-free languages by $A-CFL$. Then one can state the following result.

Theorem 36 ([34]).

$$NA-CFL_\omega \subsetneq \omega-KC(NA-CFL)$$

$$A-CFL_\omega \not\subseteq \omega-KC(A-CFL)$$

We now come to the study of links between topology and ambiguity in context-free ω -languages [34, 45].

Using a Theorem of Lusin and Novikov, and another theorem of descriptive set theory, see [50, page 123], Simonnet proved the following strong result which shows that non-Borel context-free ω -languages have a maximum degree of ambiguity.

Theorem 37 (Simonnet [45]). *Let $L(\mathcal{A})$ be a context-free ω -language accepted by a BPDA \mathcal{A} such that $L(\mathcal{A})$ is an analytic but non Borel set. The set of ω -words, which have 2^{\aleph_0} accepting runs by \mathcal{A} , has cardinality 2^{\aleph_0} .*

On the other hand, it turned out that, informally speaking, the operation $A \rightarrow A^\sim$ conserves globally the degrees of ambiguity of infinitary context-free languages (which are unions of a finitary context-free language and of a context-free ω -language). Then, starting from known examples of finitary context-free languages of a given degree of ambiguity, are constructed in [34] some context-free ω -languages of any finite Borel rank and which are non-ambiguous or of any finite degree of ambiguity or of degree \aleph_0^- .

Theorem 38.

1. For each non negative integer $n \geq 1$, there exist Σ_n^0 -complete non ambiguous context-free ω -languages A_n and Π_n^0 -complete non ambiguous context-free ω -languages B_n .
2. Let k be an integer ≥ 2 or $k = \aleph_0^-$. Then for each integer $n \geq 1$, there exist Σ_n^0 -complete context-free ω -languages $E_n(k)$ and Π_n^0 -complete context-free ω -languages $F_n(k)$ which are in $A(k) - CFL_\omega$, i.e. which are inherently ambiguous of degree k .

Notice that the ω -languages A_n and B_n are simply those which were constructed in the proof of Theorem 24. On the other hand it is easy to see that the BPDA accepting the context-free ω -language which is Borel of infinite rank, constructed in [36] using an iteration of the operation $A \rightarrow A^\sim$, has an infinite degree of ambiguity. And 1-counter Büchi automata accepting context-free ω -languages of any Borel rank of an effective analytic set, constructed via simulation of multicounter automata, may also have a great degree of ambiguity. So this left open some questions we shall detail in the last section.

We indicate now a new result which follows easily from the proof of Theorem 27 sketched in Section 4 above, see [41]. Consider an ω -language L accepted by a **deterministic** Muller Turing machine or equivalently by a **deterministic** 2-counter Muller automaton. We get first an ω -language $\theta_S(L) \subseteq X^\omega$ which has the same topological complexity (except for finite Wadge degrees), and which is accepted by a **deterministic** real time 8-counter Muller automaton \mathcal{A} .

Then one can construct from \mathcal{A} a 1-counter Muller automaton \mathcal{B} , reading words over the alphabet $X \cup \{A, B, 0\}$, such that $h(L(\mathcal{A})) \cup h(X^\omega)^- = L(\mathcal{B}) \cup h(X^\omega)^-$, where $h : X^\omega \rightarrow (X \cup \{A, B, 0\})^\omega$ is the mapping defined in Section 4. Notice that the 1-counter Muller automaton \mathcal{B} which is constructed is now also **deterministic**.

On the other hand it is easy to see, from the decomposition given in [41, Proof of Lemma

5.3], that the ω -language $h(X^\omega)^-$ is accepted by a 1-counter Büchi automaton which has degree of ambiguity 2 and the ω -language $L(\mathcal{B})$ is in $NA - CFL_\omega = CFL_\omega(\alpha \leq 1)$ because it is accepted by a **deterministic** 1-counter Muller automaton. Then we can easily infer, using [34, Theorem 5.16 (c)] that the ω -language $h(L(\mathcal{A})) \cup h(X^\omega)^- = L(\mathcal{B}) \cup h(X^\omega)^-$ is in $CFL_\omega(\alpha \leq 3)$. And this ω -language has the same complexity as $L(\mathcal{A})$. Thus we can state the following result.

Theorem 39. *For each ω -language L accepted by a **deterministic Muller Turing machine** there is an ω -language $L' \in CFL_\omega(\alpha \leq 3)$, accepted by a 1-counter Muller automaton \mathcal{D} with $\alpha_{\mathcal{D}} \leq 3$, such that $L \equiv_W L'$.*

7 ω -Powers of context-free languages

The ω -powers of finitary languages are ω -languages in the form V^ω , where V is a finitary language over a finite alphabet X . They appear very naturally in the characterization of the class REG_ω of regular ω -languages (respectively, of the class CFL_ω of context-free ω -languages) as the ω -Kleene closure of the family REG of regular finitary languages (respectively, of the family CF of context-free finitary languages). The question of the topological complexity of ω -powers naturally arises and was raised by Niwinski [66], Simonnet [75], and Staiger [79].

An ω -power of a finitary language is always an analytic set because it is either the continuous image of a compact set $\{0, 1, \dots, n\}^\omega$ for $n \geq 0$ or of the Baire space ω^ω . The first example of finitary language L such that L^ω is analytic but not Borel, and even Σ_1^1 -complete, was obtained in [35]. Amazingly the language L was very simple and even accepted by a 1-counter automaton. It was obtained via a coding of infinite labelled binary trees.

We now give a simple construction of this language L using the notion of substitution which we now recall. A *substitution* is defined by a mapping $f : X \rightarrow \mathcal{P}(\Gamma^*)$, where $X = \{a_1, \dots, a_n\}$ and Γ are two finite alphabets, $f : a_i \rightarrow L_i$ where for all integers $i \in [1; n]$, $f(a_i) = L_i$ is a finitary language over the alphabet Γ .

Now this mapping is extended in the usual manner to finite words: $f(a_{i_1} \dots a_{i_n}) = L_{i_1} \dots L_{i_n}$, and to finitary languages $L \subseteq X^*$: $f(L) = \cup_{x \in L} f(x)$. If for each integer $i \in [1; n]$ the language L_i does not contain the empty word, then the mapping f may be extended to ω -words: $f(x(1) \dots x(n) \dots) = \{u_1 \dots u_n \dots \mid \forall i \geq 1 \quad u_i \in f(x(i))\}$ and to ω -languages $L \subseteq X^\omega$ by setting $f(L) = \cup_{x \in L} f(x)$.

Let now $X = \{0, 1\}$ and d be a new letter not in X and

$$D = \{u.d.v \mid u, v \in X^* \text{ and } (|v| = 2|u|) \text{ or } (|v| = 2|u| + 1)\}$$

$D \subseteq (X \cup \{d\})^*$ is a context-free language accepted by a 1-counter automaton. Let $g : X \rightarrow \mathcal{P}((X \cup \{d\})^*)$ be the substitution defined by $g(a) = a.D$. As $W = 0^*1$ is regular, $L = g(W)$ is a context-free language and it is accepted by a 1-counter automaton. Moreover one can prove that $(g(W))^\omega$ is Σ_1^1 -complete, hence a non Borel

set. This is done by reducing to this ω -language a well-known example of Σ_1^1 -complete set : the set of infinite binary trees labelled in the alphabet $\{0, 1\}$ which have an infinite branch in the Π_2^0 -complete set $(0^*.1)^\omega$, see [35] for more details.

Remark 40. *The ω -language $(g(W))^\omega$ is context-free. By Theorem 37 every BPDA accepting $(g(W))^\omega$ has the maximum ambiguity and $(g(W))^\omega \in A(2^{\aleph_0}) - CFL_\omega$. On the other hand we can prove that $g(W)$ is a non ambiguous context-free language. This is used in [45] to prove that neither unambiguity nor ambiguity of context-free languages are preserved under the operation $V \rightarrow V^\omega$.*

Concerning Borel ω -powers, it has been proved in [32] that for each integer $n \geq 1$, there exist some ω -powers of context-free languages which are Π_n^0 -complete Borel sets. These results were obtained by the use of a new operation $V \rightarrow V^\approx$ over ω -languages, which is a slight modification of the operation $V \rightarrow V^\sim$. The new operation $V \rightarrow V^\approx$ preserves ω -powers and context-freeness. More precisely if $V = W^\omega$ for some context-free language W , then $V^\approx = T^\omega$ for some context-free language T which is obtained from W by application of a given context-free substitution. And it follows easily from [23] that if $V \subseteq X^\omega$ is a Π_n^0 -complete set, for some integer $n \geq 2$, then V^\approx is a Π_{n+1}^0 -complete set. Then, starting from the Π_2^0 -complete set $(0^*.1)^\omega$, we get some Π_n^0 -complete ω -powers of context-free languages for each integer $n \geq 3$.

An iteration of the operation $V \rightarrow V^\approx$ was used in [37] to prove that there exists a finitary language V such that V^ω is a Borel set of infinite rank. The language V was a simple recursive language but it was not context-free. Later, with a modification of the construction, using a coding of an infinity of erasers previously defined in [36], Finkel and Duparc got a context-free language V such that V^ω is a Borel set above the class Δ_ω^0 , [25].

The question of the Borel hierarchy of ω -powers of finitary languages has been solved very recently by Finkel and Lecomte in [44], where a very surprising result is proved, showing that actually ω -powers exhibit a great topological complexity. For every non-null countable ordinal α there exist some Σ_α^0 -complete ω -powers and also some Π_α^0 -complete ω -powers. But the ω -powers constructed in [44] are not ω -powers of context-free languages, except for the case of a Σ_2^0 -complete set. Notice also that an example of a regular language L such that L^ω is Σ_1^0 -complete was given by Simonnet in [75], see also [54].

8 Perspectives and open questions

We give below a list of some open questions which arise naturally. The problems listed here seem important for a better comprehension of context-free ω -languages but the list is not exhaustive.

8.1 Effective results

In the *non-deterministic* case, the Borel and Wadge hierarchies of context-free ω -languages are not effective, [32, 35, 33]. This is not surprising since most decision problems

on context-free languages are undecidable. On the other hand we can expect some decidability results in the case of *deterministic* context-free ω -languages. We have already cited some of them : we can decide whether a deterministic context-free ω -language is in a given Borel class or even in the Wadge class $[L]$ of a given regular ω -language L . The most challenging question in this area would be to find an effective procedure to determine the Wadge degree of an ω -language in the class $DCFL_\omega$.

Recall that the Wadge hierarchy of the class $DCFL_\omega$ is determined in a non-effective way in [24]. On the other hand the Wadge hierarchy of the class of blind counter ω -languages is determined in an effective way, using notions of chains and superchains, in [30]. There is a gap between the two hierarchies because (blind) 1-counter automata are much less expressive than pushdown automata. One could try to extend the methods of [30] to the study of *deterministic* pushdown automata.

Another question concerns the complexity of decidable problems. A first question would be the following one. Could we extend the results of Wilke and Yoo to the class of blind counter ω -languages, i.e. is the Wadge degree of a blind counter ω -language computable in polynomial time ? Otherwise what is the complexity of this problem ? Of course the question may be further asked for classes of ω -languages which are located between the classes of blind counter ω -languages and of deterministic context-free ω -languages.

Another interesting question would be to determine the Wadge hierarchy of ω -languages accepted by deterministic higher order pushdown automata (even firstly in a non effective way), [28, 11].

8.2 Topology and ambiguity

Simonnet's Theorem 37 states that non-Borel context-free ω -languages have a maximum degree of ambiguity, i.e. are in the class $A(2^{\aleph_0}) - CFL_\omega$. On the other hand, there exist some non-ambiguous context-free ω -languages of every finite Borel rank. The question naturally arises whether there exist some non-ambiguous context-free ω -languages which are Wadge equivalent to any given **Borel** context-free ω -language (or equivalently to any **Borel** Σ_1^1 -set, by Theorem 28). This may be connected to a result of Arnold who proved in [2] that every Borel subset of X^ω , for a finite alphabet X , is accepted by a *non-ambiguous* finitely branching transition system with Büchi acceptance condition. By Theorem 38, if k is an integer ≥ 2 or $k = \aleph_0^-$, then for each integer $n \geq 1$, there exist Σ_n^0 -complete context-free ω -languages $E_n(k)$ and Π_n^0 -complete context-free ω -languages $F_n(k)$ which are in $A(k) - CFL_\omega$, i.e. which are inherently ambiguous of degree k . More generally the question arises : determine the Borel ranks and the Wadge degrees of context-free ω -languages in classes $CFL_\omega(\alpha \leq k)$ or $A(k) - CFL_\omega$ where $k \in \mathbb{N} \cup \{\aleph_0^-, \aleph_0, 2^{\aleph_0}\}$ ($k \geq 2$ in the case of $A(k) - CFL_\omega$). A first result in this direction is Theorem 39 stated in Section 6.

8.3 ω -Powers

The results of [32, 35, 37, 44] show that ω -powers of finitary languages have actually a great topological complexity. Concerning ω -powers of context-free languages we do not know yet what are all their infinite Borel ranks. However the results of [41] suggest that ω -powers of context-free languages or even of languages accepted by 1-counter

automata exhibit also a great topological complexity.

Indeed Theorem 28 states that there are ω -languages accepted by Büchi 1-counter automata of every Borel rank (and even of every Wadge degree) of an effective analytic set. On the other hand each ω -language accepted by a Büchi 1-counter automaton can be written as a finite union $L = \bigcup_{1 \leq i \leq n} U_i \cdot V_i^\omega$, where for each integer i , U_i and V_i are finitary languages accepted by 1-counter automata. Then we can conjecture that there exist some ω -powers of languages accepted by 1-counter automata which have Borel ranks up to the ordinal γ_2^1 , although these languages are located at the very low level in the complexity hierarchy of finitary languages.

Recall that a finitary language L is a code (respectively, an ω -code) if every word of L^+ (respectively, every ω -word of L^ω) has a unique decomposition in words of L , [6]. It is proved in [45] that if V is a context-free language such that V^ω is a non Borel set then there are 2^{\aleph_0} ω -words of V^ω which have 2^{\aleph_0} decompositions in words of V ; in particular, V is really not an ω -code although it is proved in [45] that V may be a code (see the example $V = g(W)$ given in Section 7). The following question about **Borel** ω -powers now arises : are there some context-free codes (respectively, ω -codes) V such that V^ω is Σ_α^0 -complete or Π_α^0 -complete for a given countable ordinal $\alpha < \gamma_2^1$?

References

1. A. Andretta and R. Camerlo. The use of complexity hierarchies in descriptive set theory and automata theory. *Task Quarterly*, 9(3):337–356, 2005.
2. A. Arnold. Topological characterizations of infinite behaviours of transition systems. In *Proceedings of the International Conference on Automata, Languages and Programming, 10th Colloquium, Barcelona, Spain, July 18-22, 1983*, volume 154 of *Lecture Notes in Computer Science*, pages 28–38. Springer, 1983.
3. J.-M. Autebert, J. Berstel, and L. Boasson. Context free languages and pushdown automata. In *Handbook of formal languages, Vol. 1*. Springer-Verlag, 1996.
4. Danièle Beauquier. Some results about finite and infinite behaviours of a pushdown automaton. In *Proceedings of the International Conference Automata, Languages and Programming, 11th Colloquium, Antwerp, Belgium, July 16-20, 1984*, volume 172 of *Lecture Notes in Computer Science*, pages 187–195. Springer, 1984.
5. J. Berstel. *Transductions and context free languages*. Teubner Studienbücher Informatik, 1979.
6. J. Berstel and D. Perrin. *Theory of codes*. Academic Press, 1985.
7. L. Boasson. Context-free sets of infinite words. In *Proceedings of the International Conference Theoretical Computer Science, 4th GI-Conference, Aachen, Germany, March 26-28, 1979*, volume 67 of *Lecture Notes in Computer Science*, pages 1–9. Springer, 1979.
8. L. Boasson and M. Nivat. Adherences of languages. *Journal of Computer and System Science*, 20(3):285–309, 1980.
9. T. Cachat. Symbolic strategy synthesis for games on pushdown graphs. In *Proceedings of the International Conference ICALP 2002*, volume 2380 of *Lecture Notes in Computer Science*, pages 704–715. Springer, 2002.
10. T. Cachat, J. Duparc, and W. Thomas. Solving pushdown games with a Σ_3 -winning condition. In *Proceedings of the 11th Annual Conference of the European Association for Computer Science Logic, CSL 2002*, volume 2471 of *Lecture Notes in Computer Science*, pages 322–336. Springer, 2003.

11. Thierry Cachat and Igor Walukiewicz. The complexity of games on higher order pushdown automata. available at <http://fr.arxiv.org/abs/0705.0262>, May 2007.
12. B. Cagnard and P. Simonnet. Baire and automata. *Discrete Mathematics and Theoretical Computer Science*, 9(2):255–296, 2007.
13. O. Carton and D. Perrin. Chains and superchains for ω -rational sets, automata and semigroups. *International Journal of Algebra and Computation*, 7(7):673–695, 1997.
14. O. Carton and D. Perrin. The Wagner hierarchy of ω -rational sets. *International Journal of Algebra and Computation*, 9(5):597–620, 1999.
15. N. Chomsky. Three models for the description of language. *IRE Transactions on Information Theory*, 2(3):113–124, 1956.
16. Y. Choueka. Theories of automata on omega-tapes: A simplified approach. *Journal of Computer and System Science*, 8(2):117–141, 1974.
17. Y. Choueka. Finite automata, definable sets, and regular expressions over ω^n -tapes. *Journal of Computer and System Science*, 17(1):81–97, 1978.
18. Y. Choueka and D. Peleg. A note of omega-regular languages. *Bulletin of the EATCS*, 21:21–23, 1983.
19. R.S. Cohen and A.Y. Gold. Theory of ω -languages, parts one and two. *Journal of Computer and System Science*, 15:169–208, 1977.
20. R.S. Cohen and A.Y. Gold. ω -computations on deterministic pushdown machines. *Journal of Computer and System Science*, 16:275–300, 1978.
21. R.S. Cohen and A.Y. Gold. ω -computations on Turing machines. *Theoretical Computer Science*, 6:1–23, 1978.
22. J. Duparc. *La forme Normale des Boréliens de rang finis*. PhD thesis, Université Paris VII, 1995.
23. J. Duparc. Wadge hierarchy and Veblen hierarchy: Part 1: Borel sets of finite rank. *Journal of Symbolic Logic*, 66(1):56–86, 2001.
24. J. Duparc. A hierarchy of deterministic context free ω -languages. *Theoretical Computer Science*, 290(3):1253–1300, 2003.
25. J. Duparc and O. Finkel. An ω -power of a context free language which is Borel above Δ_ω^0 . In *Proceedings of the International Conference Foundations of the Formal Sciences V : Infinite Games, November 26th to 29th, 2004, Bonn, Germany*, volume 11 of *Studies in Logic, College Publications at King's College*, pages 109–122. London, 2007.
26. J. Duparc, O. Finkel, and J.-P. Ressayre. Computer science and the fine structure of Borel sets. *Theoretical Computer Science*, 257(1–2):85–105, 2001.
27. J. Duparc and M. Riss. The missing link for ω -rational sets, automata, and semigroups. *to appear in International Journal of Algebra and Computation*, 2005.
28. J. Engelfriet. Iterated pushdown automata and complexity classes. In *Proceedings of the Fifteenth Annual ACM Symposium on Theory of Computing, 25-27 April 1983, Boston, Massachusetts, USA*, pages 365–373. ACM, 1983.
29. J Engelfriet and H. J. Hoogeboom. X-automata on ω -words. *Theoretical Computer Science*, 110(1):1–51, 1993.
30. O. Finkel. An effective extension of the Wagner hierarchy to blind counter automata. In *Proceedings of Computer Science Logic, 15th International Workshop, CSL 2001*, volume 2142 of *Lecture Notes in Computer Science*, pages 369–383. Springer, 2001.
31. O. Finkel. On the Wadge hierarchy of omega context free languages. In *Proceedings of the International Workshop on Logic and Complexity in Computer Science, held in Honor of Anatol Slissenko for his 60th Birthday*, pages 69–79, Créteil, France, 2001.
32. O. Finkel. Topological properties of omega context free languages. *Theoretical Computer Science*, 262(1–2):669–697, 2001.
33. O. Finkel. Wadge hierarchy of omega context free languages. *Theoretical Computer Science*, 269(1–2):283–315, 2001.

34. O. Finkel. Ambiguity in omega context free languages. *Theoretical Computer Science*, 301(1-3):217–270, 2003.
35. O. Finkel. Borel hierarchy and omega context free languages. *Theoretical Computer Science*, 290(3):1385–1405, 2003.
36. O. Finkel. On omega context free languages which are Borel sets of infinite rank. *Theoretical Computer Science*, 299(1-3):327–346, 2003.
37. O. Finkel. An omega-power of a finitary language which is a Borel set of infinite rank. *Fundamenta Informaticae*, 62(3–4):333–342, 2004.
38. O. Finkel. Borel ranks and Wadge degrees of context free ω -languages. In *Proceedings of New Computational Paradigms: First Conference on Computability in Europe, CiE 2005, Amsterdam, The Netherlands*, volume 3526 of *Lecture Notes in Computer Science*, pages 129–138. Springer, 2005.
39. O. Finkel. On the length of the Wadge hierarchy of ω -context free languages. *Journal of Automata, Languages and Combinatorics*, 10(4):439–464, 2005.
40. O. Finkel. On winning conditions of high Borel complexity in pushdown games. *Fundamenta Informaticae*, 66(3):277–298, 2005.
41. O. Finkel. Borel ranks and Wadge degrees of omega context free languages. *Mathematical Structures in Computer Science*, 16(5):813–840, 2006.
42. O. Finkel. On the accepting power of two-tape Büchi automata. In *Proceedings of the 23rd International Symposium on Theoretical Aspects of Computer Science, STACS 2006*, volume 3884 of *Lecture Notes in Computer Science*, pages 301–312. Springer, 2006.
43. O. Finkel. Wadge degrees of infinitary rational relations. *Special Issue on Intensional Programming and Semantics in honour of Bill Wadge on the occasion of his 60th cycle, Mathematics in Computer Science*, 2(1):85–102, 2008.
44. O. Finkel and D. Lecomte. There exist some ω -powers of any Borel rank. In *Proceedings of the 16th EACSL Annual International Conference on Computer Science and Logic, CSL 2007, Lausanne, Switzerland, September 11-15, 2007*, volume 4646 of *Lecture Notes in Computer Science*, pages 115–129. Springer, 2007. to appear.
45. O. Finkel and P. Simonnet. Topology and ambiguity in omega context free languages. *Bulletin of the Belgian Mathematical Society*, 10(5):707–722, 2003.
46. E. Grädel, W. Thomas, and W. Wilke, editors. *Automata, Logics, and Infinite Games: A Guide to Current Research [outcome of a Dagstuhl seminar, February 2001]*, volume 2500 of *Lecture Notes in Computer Science*. Springer, 2002.
47. S.A. Greibach. Remarks on blind and partially blind one way multicounter machines. *Theoretical Computer Science*, 7:311–324, 1978.
48. J. E. Hopcroft, R. Motwani, and J. D. Ullman. *Introduction to automata theory, languages, and computation*. Addison-Wesley Publishing Co., Reading, Mass., 2001. Addison-Wesley Series in Computer Science.
49. J. E. Hopcroft and J. D. Ullman. *Introduction to automata theory, languages, and computation*. Addison-Wesley Publishing Co., Reading, Mass., 1979. Addison-Wesley Series in Computer Science.
50. A. S. Kechris. *Classical descriptive set theory*. Springer-Verlag, New York, 1995.
51. A. S. Kechris, D. Marker, and R. L. Sami. Π_1^1 Borel sets. *Journal of Symbolic Logic*, 54(3):915–920, 1989.
52. L.H. Landweber. Decision problems for ω -automata. *Mathematical Systems Theory*, 3(4):376–384, 1969.
53. D. Lecomte. Sur les ensembles de phrases infinies constructibles a partir d’un dictionnaire sur un alphabet fini. In *Séminaire d’Initiation a l’Analyse, Vol. 1*. Université Paris 6, 2001–2002.
54. D. Lecomte. Omega-powers and descriptive set theory. *Journal of Symbolic Logic*, 70(4):1210–1232, 2005.

55. H. Lescow and W. Thomas. Logical specifications of infinite computations. In J. W. de Bakker, Willem P. de Roever, and Grzegorz Rozenberg, editors, *A Decade of Concurrency*, volume 803 of *Lecture Notes in Computer Science*, pages 583–621. Springer, 1994.
56. M. Linna. On ω -words and ω -computations. *Ann. Univ. Turku. Ser A I*, 168:53, 1975.
57. M. Linna. On omega-sets associated with context-free languages. *Information and Control*, 31(3):272–293, 1976.
58. M. Linna. A decidability result for deterministic ω -context-free languages. *Theoretical Computer Science*, 4:83–98, 1977.
59. D. Mihoubi. Characterization and closure properties of linear omega-languages. *Theoretical Computer Science*, 191(1-2):79–95, 1998.
60. Y. N. Moschovakis. *Descriptive set theory*. North-Holland Publishing Co., Amsterdam, 1980.
61. D. E. Muller. Infinite sequences and finite machines. In *Proceedings of the Fourth Annual Symposium on Switching Circuit Theory and Logical Design, 28-30 October 1963, Chicago, Illinois, USA*, pages 3–16. IEEE, 1963.
62. R. Mac Naughton. Testing and generating infinite sequences by a finite automaton. *Information and Control*, 9:521–530, 1966.
63. M. Nivat. Mots infinis engendrés par une grammaire algébrique. *RAIRO Informatique Théorique et Applications*, 11:311–327, 1977.
64. M. Nivat. Sur les ensembles de mots infinis engendrés par une grammaire algébrique. *RAIRO Informatique Théorique et Applications*, 12(3):259–278, 1978.
65. D. Niwinski. Fixed-point characterization of context-free ∞ -languages. *Information and Control*, 61(3):247–276, 1984.
66. D. Niwinski. A problem on ω -powers. In *1990 Workshop on Logics and Recognizable Sets*, University of Kiel, 1990.
67. D. Perrin and J.-E. Pin. *Infinite words, automata, semigroups, logic and games*, volume 141 of *Pure and Applied Mathematics*. Elsevier, 2004.
68. M. O. Rabin. Decidability of second-order theories and automata on infinite trees. *Transactions of the American Mathematical Society*, 141:1–35, 1969.
69. V.L. Selivanov. Fine hierarchy of regular ω -languages. In *Proceedings of the International Joint Conference on the Theory and Practice of Software Development TAPSOFT-95, in Aarhus, Denmark*, volume 915 of *Lecture Notes in Computer Science*, pages 277–287. Springer, 1995.
70. V.L. Selivanov. Fine hierarchy of regular ω -languages. *Theoretical Computer Science*, 191:37–59, 1998.
71. V.L. Selivanov. Wadge degrees of ω -languages of deterministic Turing machines. *RAIRO-Theoretical Informatics and Applications*, 37(1):67–83, 2003.
72. V.L. Selivanov. Wadge degrees of ω -languages of deterministic Turing machines. In *Proceedings of the International Conference STACS 2003, 20th Annual Symposium on Theoretical Aspects of Computer Science, Berlin, Germany*, volume 2607 of *Lecture Notes in Computer Science*, pages 97–108. Springer, 2003.
73. O. Serre. *Contribution à l'étude des jeux sur des graphes de processus à pile*. PhD thesis, Université Paris VII, 2004.
74. O. Serre. Games with winning conditions of high Borel complexity. In *Proceedings of the International Conference ICALP 2004*, volume 3142 of *Lecture Notes in Computer Science*, pages 1150–1162. Springer, 2004.
75. P. Simonnet. *Automates et théorie descriptive*. PhD thesis, Université Paris VII, 1992.
76. L. Staiger. Hierarchies of recursive ω -languages. *Elektronische Informationsverarbeitung und Kybernetik*, 22(5-6):219–241, 1986.

77. L. Staiger. Research in the theory of ω -languages. *Journal of Information Processing and Cybernetics*, 23(8-9):415–439, 1987. Mathematical aspects of informatics (Mägdesprung, 1986).
78. L. Staiger. ω -languages. In *Handbook of formal languages, Vol. 3*, pages 339–387. Springer, Berlin, 1997.
79. L. Staiger. On ω -power languages. In *New Trends in Formal Languages, Control, Cooperation, and Combinatorics*, volume 1218 of *Lecture Notes in Computer Science*, pages 377–393. Springer-Verlag, Berlin, 1997.
80. L. Staiger and K. Wagner. Automatentheoretische und automatenfreie Charakterisierungen topologischer Klassen regulärer Folgenmengen. *Elektron. Informationsverarbeitung. Kybernetik*, 10:379–392, 1974.
81. W. Thomas. Automata on infinite objects. In J. van Leeuwen, editor, *Handbook of Theoretical Computer Science*, volume B, Formal models and semantics, pages 135–191. Elsevier, 1990.
82. W. Thomas. Infinite games and verification (extended abstract of a tutorial). In *Proceedings of the International Conference on Computer Aided Verification CAV 2002*, volume 2404 of *Lecture Notes in Computer Science*, pages 58–64. Springer, 2002.
83. W. Wadge. *Reducibility and determinateness in the Baire space*. PhD thesis, University of California, Berkeley, 1983.
84. K. Wagner. On ω -regular sets. *Information and Control*, 43(2):123–177, 1979.
85. I. Walukiewicz. Pushdown processes: games and model checking. *Information and Computation*, 157:234–263, 2000.
86. T. Wilke and H. Yoo. Computing the Wadge degree, the Lifschitz degree, and the Rabin index of a regular language of infinite words in polynomial time. In P.D. Mosses, M. Nielsen, and M.I. Schwartzbach, editors, *TAPSOFT 95*, volume 915 of *Lect. Notes in Comp. Sci.*, pages 288–302. Springer Verlag, Berlin, Heidelberg, New York, 1995.