The theory of rational relations on transfinite strings

Christian Choffrut
LIAFA, Université Paris 7
2, pl. Jussieu 75251 Paris Cedex 05
France
cc@liafa.jussieu.fr

Serge Grigorieff
LIAFA, Université Paris 7
2, pl. Jussieu 75251 Paris Cedex 05
France
seg@liafa.jussieu.fr

August 28, 2002

1 Introduction

In order to explain the purpose of this paper, we recall briefly how the theory of "rationality" developed in the last fourty years.

The theory of sets of finite strings recognized by finite automata, also known as regular or rational sets was developed in the fifties. It rapidly extended in two directions. Indeed, by solving the decidability problem of the second order monadic theory of one successor, Büchi was led naturally to introduce the notion of finite automata working on infinite strings. He further extended this result to the monadic theory of all denumerable ordinals, and by doing so he again modified the original notion of finite automata to suit his new purpose, [9]. At this point, the equivalence between the notions of recognizability (by "finite automata"), rationality (by "rational expressions") and definability (by "monadic second order logics") was achieved as far as strings of denumerable lengths were concerned.

In the late sixties, Elgot and Mezei wrote an historical paper on rational relations [15] which was a successful attempt to construct the theory of relations between free monoids that could be recognized by so-called n-tape automata. Though hard to read, it contained the basic results of the theory. In the mid eighties Gire and Nivat showed that the theory of rational relations on finite strings carries over to infinite strings, [16]. More recently, Wilke made a breakthrough by giving an algebraic characterization of the rational infinite strings via the notion of "right binoids" (now known as "Wilke algebras"), [25]. This construct happens to be the natural extension of finite monoids for infinite strings. Elaborating on this notion, Bedon [2] showed that rational transfinite strings of countable length recognized by finite automata can also be recognized by finite algebras (the ω_1-algebras).

In the present paper we draw the theory further by showing that almost everything carries over to transfinite strings. The proof of most elementary properties are mere paraphrase of those for finite strings, so we state them here for self-containment but we leave the simple verifications to the reader. For a few results completely new proof techniques are required, like the equivalent of the "second factorization theorem" (in Eilenberg's terminology [14, p. 248]), cf. Thm. 32. Finally, some properties no longer hold beyond the length ω^ω like the uniformization problem, cf. Thm. 38 and §5.6.

Observe finally that our properties do not cover all those of [16] since the concatenation of strings is only partially defined. In some sense our general framework seems to be more natural. For example, when dealing with infinite strings the property that the family of relations recognized by automata is closed under concatenation needs some intricate case analysis ([16, pp. 107–110]) while a purely formal proof works for transfinite strings. Also, the notion of direct product of rational subsets (the "recognizable relations") is completely clarified now with the ω_1-algebras of Wilke.

2 Preliminaries

Here we present the basic facts on ordinals, transfinite strings and finite automata for transfinite strings, extending thus the elementary definitions usual for finite strings.

2.1 Ordinals

In this paper we shall deal with ordinals less than the first non-denumerable ordinal ω_1. We refer the interested reader to the standard textbooks, e.g., [24] and [23], for a thorough exposition of the material on ordinals. We recall an ordinal is *prime* if it cannot be expressed as the sum of two smaller ordinals; these ordinals are exactly the powers of ω. The (unique) Cantor's normal form of an ordinal $0 < \alpha < \omega_1$ is the sum

$$\alpha = \omega^{\lambda_n} a_n + \omega^{\lambda_{n-1}} a_{n-1} + \ldots + \omega^{\lambda_1} a_1 + \omega^{\lambda_0} a_0 \tag{1}$$

where $0 \leq n < \omega$, $\lambda_n > \ldots > \lambda_1 > \lambda_0$, $0 \leq a_i < \omega$ for all $i \leq n$ and $0 < a_n$. The ordinal λ_n is the *degree* and the ordinal λ_0 is the *type* of α. The degree and the type of 0 are equal to 0.

A few elementary definitions on sequences indexed by ordinals will be used in the sequel. We review them here.

Definition **1.** Let α be a limit ordinal. An increasing sequence $(\beta_\eta)_{\eta<\lambda}$ is *cofinal* to α if β_η is less than α for all $\eta < \lambda$ and if for all $\beta < \alpha$ there exists $0 \leq \eta < \lambda$ such that $\beta < \beta_\eta < \alpha$.

We shall also make use of the following well-known fact: if α is a countable ordinal then every sequence $(\beta_\eta)_{\eta<\lambda}$ which is *cofinal* to α contains an ω-subsequence $(\beta_{\eta_i})_{i<\omega}$ which is also *cofinal* to α.

Definition **2.** A (strictly) increasing sequence $(\beta_\eta)_{\eta<\lambda}$ is *continuous* if for all limit ordinals $\eta < \lambda$, $\sup_{\epsilon<\eta} \beta_\epsilon = \beta_\eta$ or equivalently, β_η is a limit ordinal and the sequence $(\beta_\epsilon)_{\epsilon<\eta}$ is cofinal to β_η.

By posing $\alpha = \lim_{\eta<\lambda} \beta_\eta$ if λ is a limit ordinal or $\alpha = \beta_{\lambda-1}$ if it is a successor ordinal, the previous property is equivalent to saying that the set of all ordinals less than α is precisely the union of all the semi-open intervals $[\beta_\eta, \beta_{\eta+1}[$ for all $0 \leq \eta + 1 < \lambda$. E.g., with $\lambda = \omega + 1$, the sequence $\beta_i = i$ for all $i < \omega$ and $\beta_\omega = \omega + 1$ is not continuous, nor is the sequence $\beta_i = i$ for all $i < \omega$ and $\beta_\omega = \omega \times 2$.

2.2 Transfinite strings

Given a finite alphabet A, a *string* is a mapping u from some $\alpha < \omega_1$ into A. Equivalently, u is a sequence of elements of A indexed by an ordinal α. We denote u_β the element indexed by $\beta < \alpha$ in this sequence. The ordinal α is the *length* of u, denoted by $|u|$. The collection of all strings is denoted by $A^{<\omega_1}$ and we will call them improperly *transfinite* though they might also be of finite length. The *empty string*, denoted by 1, is the string of length 0 and is the unit of $A^{<\omega_1}$ as a monoid. By extension, the *degree* of a string u is the degree of its length. For $a \in A$, $|u|_a$ denotes the *length in the letter* a of the string u, i.e., the ordinal of the subsequence indexed by the set of positions $\beta < \alpha$ for which $u_\beta = a$. The set of strings is partially ordered by the "prefix relation" : $u \leq v$ if there exists w such that $v = uw$.

2.3 Continuous mappings

From now on and unless otherwise stated, the term "increasing" when applied to strings, is to be understood relative to the prefix ordering.

The string v is the limit of the increasing sequence $(u_\eta)_{\eta<\lambda}$ of prefixes of v if for all prefixes w of v there exists an index η such that w is a prefix of u_η. E.g., $\lim_{i\to\omega}(ab)^i = \lim_{i\to\omega}(ab)^i a = (ab)^\omega$. Let $h : A^{<\omega_1} \to B^{<\omega_1}$ be increasing with respect to the prefix ordering. We say that h is *continuous* if it commutes with limits of strings, i.e., if for all limit ordinals λ and all string $u = \lim_{\eta\to\lambda} u_\eta \in A^{<\omega_1}$, equality $h(u) = \lim_{\eta\to\lambda} h(u_\eta)$ holds.

Any mapping h from A into $B^{<\omega_1}$ can be extended in a unique way as an increasing (relative to the prefix ordering) and continuous (in the above sense) function of $A^{<\omega_1}$ into $B^{<\omega_1}$. Thus, all morphisms will be determined by the images of the letters. This can be proved by transfinite induction by setting $h(ua) = h(u)h(a)$ for all $u \in A^{<\omega_1}$ and $a \in A$ and $h(u) = \sup\{h(v) \mid v \text{ is a prefix of } u\}$ if the length of u is a limit ordinal. This latter string is indeed well determined, it is the string of length $\alpha = \sup\{|h(v)| \mid v \text{ is a prefix of } u\}$ whose prefix of length $\beta < \alpha$ is the prefix of length β of every string $h(v)$ where v a prefix of u and $\beta \leq |h(v)| < \alpha$.

Similarly, every mapping h of A into the direct product $B_1^{<\omega_1} \times B_2^{<\omega_1}$ can be extended to a mapping of $A^{<\omega_1}$ into $B_1^{<\omega_1} \times B_2^{<\omega_1}$ by posing $h(u) = ((\pi_1 h)(u), (\pi_2 h)(u))$.

3 Finite automata on transfinite strings

There are two different ways of defining a finite automaton on transfinite strings. Both are due to Büchi [9]. The first one was extensively studied by Choueka and considers strings of length less than ω^{n+1} for a given n, see [13]. The second one deals with strings of arbitrary countable lengths and was investigated (and actually extended to uncountable ordinals) by Wojciechowski [27]. Both formalisms are equivalent when restricted to strings of length less than ω^{n+1} for some $n < \omega$, (cf. Theorem 22), so we shall use which one is more amenable depending on the question under investigation.

3.1 Büchi automata on transfinite strings

Here we recall precisely the above mentioned definitions of finite automata for transfinite strings starting with what is nowadays known as Büchi's automaton. The main new point is about the definition of limit transitions and relies on the notion of cofinal sequences, as defined in paragraph 2.1.

Definition **3.** Let Q be a finite set, let α be a limit ordinal and let $(q_\beta)_{\beta<\alpha}$ be a sequence of elements of Q. An element $q \in Q$ is *persistent* in the sequence if $\{\beta < \alpha \mid q = q_\beta\}$ is cofinal to α. In other words, for some increasing sequence $(\beta_i)_{i<\omega}$ cofinal to α we have $q = q_{\beta_i}$ for all $i < \omega$.

In the more familiar context of Büchi automata on ω-words, the states that are "infinitely repeated" in some infinite path are the persistent ones.

Definition **4.** A *Büchi automaton* is a quintuple $\mathcal{A} = (Q, A, Q_-, Q_+, E)$ where Q is the (finite) set of *states*, A is the finite *input alphabet*, $Q_- \subseteq Q$ is the set of *initial states*, $Q_+ \subseteq Q$ is the set of *final states* and the set E of *transitions* is a subset of $(Q \times A \times Q) \cup (2^Q \times Q)$.

If u is a string of length α, a *run* with *label* u is a sequence $(q_\beta)_{\beta\leq\alpha}$ of elements in Q which satisfies the following inductive conditions (recall u_β is the letter at position β in u, for all $\beta < \alpha$):

1. If $\beta < \alpha$ then $q_{\beta+1}$ satisfies $(q_\beta, u_\beta, q_{\beta+1}) \in E$.
2. If $\beta \leq \alpha$ is a limit ordinal then $(P, q_\beta) \in E$ where $P \subseteq Q$ is the set of persistent elements in the run $(q_\gamma)_{\gamma<\beta}$.

Observe that a run with label u has length $|u| + 1$.
The run is *successful* if $q_0 \in Q_-$ and $q_\alpha \in Q_+$.
A subset of $A^{<\omega_1}$ is Büchi recognizable if it is the set of labels of successful runs of some Büchi automaton. We denote Büchi$(A^{<\omega_1})$ the family of subsets of $A^{<\omega_1}$ which are recognizable by a Büchi automaton.

A word of caution. The way we introduced the relation E is a slight modification of Büchi's original treatment of limit cases (for which the set of persistent states itself is considered as the limit state). However, this does not change the family of recognized subsets of $A^{<\omega_1}$ as the reader may easily verify.

REMARK 5. The family of Büchi recognizable languages is easily seen to be closed by union and intersection. Closure by complementation is a difficult result (Büchi, 1965, [8], cf. also [9]).

The next result insures that given an automaton, the set of strings that label some path visiting a specified subset of states is rational. It will be used later in Theorem 32.

LEMMA **6.** *Let $\mathcal{A}$ be a Büchi automaton, Q its set of states and $V \subseteq Q$ a fixed subset. For all $p, q \in Q$, the set $X_{p,V,q}$ of strings that label some run from p to q which visits exactly the states in V is Büchi recognizable.*

Proof. If $\mathcal{A} = (Q, A, Q_-, Q_+, E)$ then $X_{p,V,q}$ is recognized by the automaton $\mathcal{B} = (Q \times 2^Q, A, \{(p, \emptyset)\}, \{(q, V)\}, F)$ where the transition set F satisfies the following conditions

1. for all $q_1, q_2 \in Q$, $W \subseteq Q$, $a \in A$ we have
$$((q_1, W), a, (q_2, W \cup \{q_2\})) \in F \text{ if and only if } (q_1, a, q_2) \in E$$
2. for all $\{q_1, \ldots, q_k\} \subseteq Q$, $W \subseteq Q$, $r \in Q$ we have
$$(S, (r, W \cup \{r\})) \in F \text{ if and only if } (\{q_1, \ldots, q_k\}, r) \in E$$
where $S = \{(q_1, W), \ldots, (q_k, W)\}$.

□

3.2 Transfinite pumping lemma

The classical pumping lemma admits a transfinite version. First, we introduce some notation and mention a technical tool.

NOTATION 7. Let $\rho = (q_\alpha)_{\alpha \le |u|}$ be a run of a Büchi automaton $\mathcal{A}$. We denote $\texttt{Visit}(\rho \restriction [\beta, \gamma[)$ and $\texttt{Persist}(\rho \restriction [\beta, \gamma[)$ the respective sets of visited and persistent states of the subrun $(q_\alpha)_{\beta \le \alpha < \gamma}$.

PROPOSITION **8.** *Let $\rho = (q_\alpha)_{\alpha \le \mu}$ be a run of a Büchi automaton $\mathcal{A} = (Q, A, Q_-, Q_+, E)$ such that there are ζ, λ satisfying*

1. *λ is limit and $\zeta + \lambda \le \mu$ and $q_\zeta = q_{\zeta+\lambda}$*
2. $\texttt{Visit}(\rho \restriction [\zeta, \zeta + \lambda[) = \texttt{Persist}(\rho \restriction [\zeta, \zeta + \lambda[)$

Let $\sigma = (q_\alpha)_{0 \le \alpha < \zeta}$, $\tau = (q_\alpha)_{\zeta \le \alpha < \zeta+\lambda}$, $\nu = (q_\alpha)_{\zeta+\lambda \le \alpha < \mu}$, (so that $\rho = \sigma\tau\upsilon$ and $\tau = \rho \restriction [\zeta, \zeta + \lambda[$). Then $\sigma\tau^\xi\nu$ is a run for all $\xi < \omega_1$.

Proof. We prove by induction on $\xi \ge 1$ that
(*) τ^ξ is a run
(**) $\texttt{Visit}(\tau^\xi) = \texttt{Persist}(\tau^\xi) = \texttt{Persist}(\tau)$.
Observe that if properties (*),(**) are true for ξ_1, ξ_2 then they are also true for $\tau_1^\xi \tau_2^\xi$. Now, if $\xi = \xi_1 + \xi_2 + \ldots$ and properties (*),(**) are true for the τ_i^ξ's ($i \in \omega$) then they are also true for τ^ξ. Lastly, case $\xi = 1$ is condition 2 of the Proposition. □

LEMMA **9 (Transfinite pumping lemma).**
1) If a Büchi automaton $\mathcal{A}$ with n states accepts some word u with length $\ge \omega^n$ then there is a factorization $u = xyz$ such that
- $|y| = \omega^i$ for some $i \ge 1$,
- $xy^\xi z$ is accepted by $\mathcal{A}$ for all $\xi < \omega_1$.
In particular, $\mathcal{A}$ accepts words of arbitrary length $< \omega_1$.
2) If for every n a Büchi automaton $\mathcal{A}$ accepts some word with length $\ge \omega^n$ then $\mathcal{A}$ accepts words of arbitrary length $< \omega_1$.
In particular if $\mathcal{A}$ accepts some word with length $\ge \omega^\omega$ then $\mathcal{A}$ accepts words of arbitrary length $< \omega_1$.

Proof. Point 2 is a straightforward application of Point 1. We keep notations of Def. 4. Observe that for every run σ on a word with limit length, there exists some step η after which all visited states are persistent ones. Applying this fact to an accepting run ρ on input u, we inductively define $\eta_0 < \omega^n$, $\eta_1 < \omega^{n-1} \ldots \eta_{n-1} < \omega$ as follows:

- η_0 is the least $\eta < \omega^n$ such that all states visited during steps in $[\eta, \omega^n[$ are persistent ones.
- η_1 is the least $\eta < \omega^{n-1}$ such that all states visited during steps in $[\eta_0 + \eta, \eta_0 + \omega^{n-1}[$ are persistent ones.
- ...
- For $i < n$, η_i is the least $\eta < \omega^{n-i}$ such that all states visited during steps in $[\eta_0 + \eta_1 + \ldots + \eta_{n-i-1} + \eta, \eta_0 + \eta_1 + \ldots + \eta_{n-i-1} + \omega^{n-i}[$ are persistent ones.

Let $u = hbt$ (for "head'", "body" and "tail") where $|h| = \eta_0 + \eta_1 + \ldots + \eta_{n-1}$ and $|b| = \omega^n$. By construction, for all $1 \leq j \leq n$ we have

$$\texttt{Visit}(\rho \restriction [|h|, |h| + \omega^j[) = \texttt{Persist}(\rho \restriction [|h|, |h| + \omega^j[).$$

hence also for all $0 \leq k < j \leq n$

$$(*) \qquad \texttt{Visit}(\rho \restriction [|h| + \omega^k, |h| + \omega^j[) = \texttt{Persist}(\rho \restriction [|h| + \omega^k, |h| + \omega^j[)$$

Consider now the sequence of states

$$q_{|h|}, q_{|h|+\omega}, \ldots, q_{|h|+\omega^n}$$

Since there are n states, there must exist $0 \leq k < j \leq n$ such that $q_{|h|+\omega^k} = q_{|h|+\omega^j}$. Equality (*) allows to conclude using Prop. 8. □

REMARK 10. 1) The above transfinite pumping lemma is optimal. The following n states automaton $\mathcal{A} = (Q, A, Q_-, Q_+, E)$ recognizes the set of words of length $< \omega^n$ on alphabet A :

$Q = \{0, 1, \ldots, n-1\}$, $Q_- = \{0\}$, $Q_+ = Q$

$E = \{(i, a, 0,) \mid a \in A, 1 \leq i \leq n-1\}$
$\cup\{(\{0, \ldots, i\}, i+1) \mid 0 \leq i \leq n-2\}$

2) The transfinite pumping lemma can be improved as follows: instead of considering the number of states of $\mathcal{A}$, one can consider the (smaller) number of states which appear at limit transitions, i.e. states q such that $(X, q) \in E$ for some $X \subseteq Q$.

DEFINITION **11.** Büchi$(A^{<\alpha}) = \{X \cap A^{<\alpha} \mid X \in$ Büchi$(A^{<\omega_1})\}$.

As a corollary of Lemma 9 and Remarks 5,10, we get

Corollary 12.
If $\omega^\omega \leq \alpha < \omega_1$ *then* $\text{Büchi}(A^{<\omega_1}) \cap P(A^{<\alpha}) = \bigcup_{n<\omega} \text{Büchi}(A^{<\omega^{n+1}})$.
In other words, $\text{Büchi}(A^{<\alpha}) = \text{Büchi}(A^{<\omega_1}) \cap P(A^{<\alpha})$ *if and only if* $\alpha < \omega^\omega$.

3.3 Choueka automata on transfinite strings

Corollary 12 shows that if $\alpha \geq \omega^\omega$ then $\text{Büchi}(A^{<\alpha})$ has no intrinsic characterization in terms of Büchi automata.
As for $\alpha < \omega^\omega$, there is a variant of Büchi automata which captures exactly the family $\text{Büchi}(A^{<\alpha})$: the so-called Choueka automata which deal with strings having length less than ω^{n+1} for some $n < \omega$. They require a special notion of limit, known as "Choueka-continuity" in the literature, see [13]. In case $\alpha < \omega^{n+1}$ is a limit ordinal with type k (such that $0 < k \leq n$) then $\alpha = \beta + \omega^k$ for some β and there is a canonical sequence cofinal to α, namely $(\beta + \omega^{k-1}m)_{m<\omega}$ which is the one considered at limit steps for Choueka automata.

Given a set Q, we pose $[Q]^0 = Q$, $[Q]^k = 2^{[Q]^{k-1}} \setminus \emptyset$ if $k > 0$ and $[Q]_0^n = \bigcup_{0\leq k\leq n} [Q]^k$.

The restriction to successor ordinals in the next definition is motivated by the two following facts:
i) runs on inputs with length α have length $\alpha + 1$,
ii) successor ordinals are exactly those ordinals which contain limits of sequences of smaller ordinals.

Definition **13.** A transfinite sequence s of length $\alpha + 1 < \omega^\omega$ is *Choueka-continuous over the set* Q if for all $\beta \leq \alpha$ of the form $\beta = \gamma + \omega^{i+1}$ (i.e. i is the type of β) we have $s_\beta = \{e \in [Q]^i \mid \exists^\infty k,\ \ e = s_{\gamma+\omega^i.k}\}$. In particular, s maps $\alpha + 1$ into $[Q]_0^{t-1}$ (and not into Q) where t is the type of α.

Definition **14.** A *Choueka automaton* is a quintuple $\mathcal{A} = (Q, A, E, Q_-, Q_+)$ where Q is the (finite) set of *states*, A is the *input alphabet*, $Q_- \subseteq Q$ is the set of *initial states*, $Q_+ \subseteq [Q]_0^n$ is the set of *final states* and $E \subseteq [Q]_0^n \times A \times Q$ is the set of *transitions.*

The notion of *run* with *label* $u = (u_\beta)_{\beta<\alpha}$ as interpreted for Büchi's automata, extends naturally to Choueka's automata. Indeed, it is a Choueka-continuous sequence $(q_\beta)_{\beta\leq\alpha} \in [Q]_0^n$ satisfying $(q_\beta, u_\beta, q_{\beta+1}) \in E$ for all $\beta < \alpha$. Observe that it is clear from Definitions 13 and 14 that $q_\beta \in [Q]^k$ if β is of type k.

A run is *successful* if its first state q_0 is an initial state and its last state q_α is a final state. A subset of $A^{<\omega^{n+1}}$ is *Choueka recognizable* if it is the set of labels of successful runs of some Choueka automaton.

REMARK 15. The family of Choueka recognizable languages is easily seen to be closed by union and intersection. Closure by complementation is also true (Choueka, [13]).

3.4 Rational sets of transfinite strings

A fundamental result is that Kleene's theorem can be extended to transfinite strings for both notions of automata. To that purpose, one has to consider two new operations on sets of strings.

DEFINITION **16.** 1) The ω-*power* of a set X, denoted X^ω, is the set of strings obtained by concatenating ω-sequences of strings in X.
2) The ω_1-*iteration* of a set X is the set $X^{<\omega_1} = \bigcup_{\alpha<\omega_1} X^\alpha$, where X^α is the set of strings obtained by concatenating α-sequences of strings in X (in particular, $X^0 = \{1\}$).
3) The n-*trace*-ω-*power* of a set X is $n\text{-trace}(X^\omega) = X^\omega \cap A^{<\omega^{n+1}}$.

REMARK 17. 1) $X^{<\omega_1}$ is the closure of $X \cup \{1\}$ under ω-power, i.e. it is the smallest set Y which contains $X \cup \{1\}$ and is closed under ω-power: $X \cup \{1\} \subseteq Y = Y^\omega$. It is also the closure of $X \cup \{1\}$ under Kleene-star and ω-power, i.e. the smallest set Y such that $X \cup \{1\} \subseteq Y = Y^* = Y^\omega$. In fact, since $(X \cup \{1\})^\omega \supseteq X^*$, such a Y necessarily satisfies $Y = Y^2$ and an easy induction over ordinals less than ω_1 (using equalities $Y = Y^2$ and $Y = Y^\omega$ for the respective successor and limit cases) shows that $Y = Y^\alpha$ for all $\alpha < \omega_1$, hence $Y = Y^{<\omega_1}$.
2) Observe that there is no need for a similar closure operation relative to the n-trace-ω-power. In fact, an easy induction over i shows that the i-th iterate of n-trace-ω-power of $X \cup \{1\}$ is $(\bigcup_{\alpha\leq\omega^i} X^\alpha) \cap A^{<\omega^{n+1}}$. All this iterates coincide for $i \geq n+1$, therefore the $(n+1)$-th iterate of n-trace-ω-power of $X \cup \{1\}$ is closed under n-trace-ω-power.

DEFINITION **18.** 1) [Wojciechowski, 1985 [27]] $\mathrm{Rat}(A^{<\omega_1})$ denotes the least family of subsets of $A^{<\omega_1}$ that contains the empty set, the singleton sets consisting of a letter and is closed under set union, concatenation, Kleene-star, ω-power and ω_1-iteration.
2) [Choueka, 1978 [13]] $\mathrm{Rat}(A^{<\omega^{n+1}})$ denotes the least family of subsets of $A^{<\omega^{n+1}}$ that contains the empty set, the singleton sets consisting of a letter and is closed under set union, concatenation, Kleene-star, and the n-trace-ω-power.

The following can be proven with the same structural induction technique as that for rational subsets of free monoids. Recall that a *rational substitution* of C into $A^{<\omega_1}$ is a mapping σ that assigns a rational subset of $A^{<\omega_1}$ to each $c \in C$. Given an ordinal α, one can extend σ to $A^{<\alpha}$ by setting $\sigma((u_\beta)_{\beta<\alpha}) = \prod_{\beta<\alpha} \sigma(u_\beta)$.

PROPOSITION **19.** *If* $\sigma : C \to \mathrm{Rat}(A^{<\omega_1})$ *is a rational substitution and* $X \in \mathrm{Rat}(C^{<\omega_1})$ *then* $\sigma(X) = \bigcup\{\sigma(x) \mid x \in X\}$ *is in* $\mathrm{Rat}(A^{<\omega_1})$.
An analog property holds with ω^{n+1} *in place of* ω_1.

3.5 Kleene type results

As for finite strings, there is an equivalence for subsets of transfinite strings between recognizability via some type of finite automaton and expressability via some type of operations. We state these results for transfinite lengths less than ω^n for some $n < \omega$ and for arbitrary lengths less than ω_1.

THEOREM **20** (Choueka, 1974,[13], cf. also Bedon, 1996, [1]).
$\mathrm{Rat}(A^{<\omega^{n+1}})$ *is exactly the family of Choueka recognizable subsets of* $A^{<\omega^{n+1}}$.

THEOREM **21** (Wojciechowski, 1985,[27]).
$\mathrm{Rat}(A^{<\omega_1})$ *is exactly the family of Büchi recognizable subsets of* $A^{<\omega_1}$.

Recall (Remark 10) that $A^{<\omega^{n+1}}$ (considered as a subset of $A^{<\omega_1}$) is Büchi recognizable. Now, as a corollary of the above theorems, we see that Choueka recognizability is a special case of Büchi's recognizability. This is proven in [1] but it can more easily be seen by arguing on the lengths of the strings.

Theorem **22** ([13],[1]). *Let $X \subseteq A^{<\omega^{n+1}}$. The following conditions are equivalent:*

1) *X is Büchi recognizable*

2) *X is Choueka recognizable*

3) *$X = Y \cap A^{<\omega^{n+1}}$ for some Büchi recognizable set $Y \subseteq A^{<\omega_1}$*

4 Rational relations on transfinite strings

4.1 Two-tape Büchi and Choueka automata

The idea of Büchi and Choueka automata described above extends to two-tape automata.

Definition **23.** A *two-tape Büchi automaton* on transfinite strings is a construct $\mathcal{A} = (Q, A, B, Q_-, Q_+, E)$ where Q, Q_-, Q_+, A are as in Definition 4, B is a finite alphabet and where $E = E_A \cup E_B \cup E_\ell$ with

$$E_A \subseteq Q \times (A \times \{1\}) \times Q \tag{2}$$

$$E_B \subseteq Q \times (\{1\} \times B) \times Q \tag{3}$$

$$E_\ell \subseteq 2^Q \times Q \tag{4}$$

The elements of (2), (3) and (4), are respectively called the A, B and limit transitions. The notions of run, label, successful run are straightforward extensions of the corresponding notions in finite automaton, the only difference being that labels are pairs of strings rather than single strings. The relation in $A^{<\omega_1} \times B^{<\omega_1}$ defined by the two-tape Büchi automaton is the set of pairs (u, v) that are the labels of some successful run and is said to be *Büchi recognizable.*

Definition **24.** A *two-tape Choueka automaton* on transfinite strings is a construct $\mathcal{A} = (Q, A, B, Q_-, Q_+, E)$ where Q, Q_-, Q_+, A are as in Definition 14, B is a finite alphabet and where $E \subseteq [Q]_0^n \times ((A \times \{1\}) \cup (\{1\} \times B)) \times Q$.

The relation in $A^{<\omega^{n+1}} \times B^{<\omega^{n+1}}$ defined by the two-tape Choueka automaton is said to be *Choueka recognizable.*

4.2 Rational relations

The operations needed for defining rational relations are those introduced for subsets extended in the usual way to pairs of strings. For instance Definition 16 obviously extends from strings to pairs of strings which leads to the following extension of Definition 18.

Definition **25.** 1) $\mathrm{Rat}(A^{<\omega_1} \times B^{<\omega_1})$ denotes the least family of subsets of $A^{<\omega_1} \times B^{<\omega_1}$ that contains the empty set, the singleton sets $\{(a,1)\}$, $\{(1,b)\}$ for $a \in A$, $b \in B$ and is closed under set union, concatenation, Kleene-star, ω-power and ω_1-iteration.
2) $\mathrm{Rat}(A^{<\omega^{n+1}} \times B^{<\omega^{n+1}})$ denotes the least family of subsets of $A^{<\omega_1} \times B^{<\omega_1}$ that contains the empty set, the singleton sets $\{(a,1)\}$, $\{(1,b)\}$ for $a \in A$, $b \in B$ and is closed under set union, concatenation, Kleene-star, and the n-trace of ω-power.

The closure under rational substitution seen in Proposition 19 extends easily to relations.

Proposition **26.** *1) If* $\sigma : C \to \mathrm{Rat}(A^{<\omega_1})$ *and* $\tau : C \to \mathrm{Rat}(B^{<\omega_1})$ *are rational substitutions and* $X \in \mathrm{Rat}(C^{<\omega_1})$ *then* $(\sigma,\tau)(X) = \bigcup_{x\in X} \sigma(x) \times \tau(x)$ *is in* $\mathrm{Rat}(A^{<\omega_1} \times B^{<\omega_1})$.
2) If $\sigma : C \to \mathrm{Rat}(A^{<\omega^{n+1}})$ *and* $\tau : C \to \mathrm{Rat}(B^{<\omega^{n+1}})$ *are rational substitutions and* $X \in \mathrm{Rat}(C^{<\omega^{p+1}})$ *then* $(\sigma,\tau)(X) = \bigcup_{x\in X} \sigma(x) \times \tau(x)$ *is in* $\mathrm{Rat}(A^{<\omega^{np+1}} \times B^{<\omega^{np+1}})$.

4.3 The first factorization theorem

We recall that a morphism $\phi : A^* \to B^*$ is *alphabetic* whenever $\phi(A) \subseteq B \cup \{1\}$ holds and that it is *strictly alphabetic* whenever $\phi(A) \subseteq B$ holds. These notions make sense when applied to morphisms from $A^{<\omega_1}$ into $B^{<\omega_1}$.

Proposition 26 admits a reciprocal which is given some normalized forms in Proposition 27 below and Theorems 32,33. The first form we consider is the transfinite extension of a result on finite strings first observed by Nivat, 1968, [22], which is called the first factorization theorem in Eilenberg [14] p. 240. Its proof is a paraphrase of the proof for finite strings, cf., e.g., [6, Thm. III.4.1.] and consists in considering pairs $(a,1),(1,b)$ as letters of a new alphabet C and introducing the projections from C onto $A \cup \{1\}$ and $B \cup \{1\}$.

Proposition **27.** *The following conditions are equivalent*

1) $R \in \mathrm{Rat}(A^{<\omega_1} \times B^{<\omega_1})$

2) *there exist a finite alphabet C, a rational subset K of $C^{<\omega_1}$ and two alphabetic morphisms $\pi_A : C \to A \cup \{1\}$, $\pi_B : C \to B \cup \{1\}$ such that $R = \{(\pi_A(x), \pi_B(x)) \mid x \in K\}$. Moreover, one can suppose*

$$(*) \qquad \forall c \in C(\pi_A(c) \text{ or } \pi_B(c) \text{ is not empty})$$

An analog equivalence holds with ω^{n+1} in place of ω_1.

As a corollary, we get the relational version of Theorem 22.

Theorem **28.** *Let $R \subseteq A^{<\omega^{n+1}} \times B^{<\omega^{n+1}}$. The following conditions are equivalent:*

1) *R is Büchi recognizable*

2) *R is Choueka recognizable*

3) *$R = S \cap (A^{<\omega^{n+1}} \times B^{<\omega^{n+1}})$ for some Büchi recognizable relation S.*

Proof. **1)** $\Rightarrow$ **2)**. Let K, π_A, π_B be as in **2)** of Proposition 27 with condition (*). Clearly, $|\pi_A(x)| + |\pi_B(x)| \geq |x|$ for all $x \in C^{<\omega_1}$. Since $R \subseteq A^{<\omega^{n+1}} \times B^{<\omega^{n+1}}$ we see that $K \subseteq C^{<\omega^{n+1}}$. We conclude by Theorem 22 and Proposition 27 applied to ω^{n+1}.
2) $\Rightarrow$ **3)**. If R is Choueka recognizable then by Proposition 27 with condition (*) applied to ω^{n+1} and Theorem 22 there exists $K = K' \cap C^{<\omega^{n+1}}$ where K' is Büchi recognizable such that $R = \{(\pi_A(x), \pi_B(x)) \mid x \in K\}$. Then $R = S \cap (A^{<\omega^{n+1}} \times B^{<\omega^{n+1}})$ where $S = \{(\pi_A(x), \pi_B(x)) \mid x \in K'\}$.
3) $\Rightarrow$ **1)**. Let $S = \{(\pi_A(x), \pi_B(x)) \mid x \in K\}$ for some $K \in \mathrm{Rat}(C^{<\omega_1})$ with condition (*). Then $R = \{(\pi_A(x), \pi_B(x)) \mid x \in K \cap C^{<\omega^{n+1}}\}$. □

Another corollary of the first factorization theorem is the closure under composition property. The proof is also a paraphrase of the same property for finite strings, [6, Thm III.4.4.]. We reproduce it here for the sake of completeness.

Proposition **29 (Closure under composition).**
If $R \in \mathrm{Rat}(A^{<\omega_1} \times B^{<\omega_1})$ *and* $S \in \mathrm{Rat}(B^{<\omega_1} \times C^{<\omega_1})$ *then* $R \circ S \in \mathrm{Rat}(A^{<\omega_1} \times C^{<\omega_1})$. *The same holds with* $A^{<\omega^{n+1}}, B^{<\omega^{n+1}}, C^{<\omega^{n+1}}$ *for all* $0 \leq n < \omega$.

Proof. The proof goes exactly as in the finite length case, see [6]. Denote by π_A^{AB} the projection of $A \cup B$ onto A and by π_{AB}^{ABC} the projection of $A \cup B \cup C$ onto $A \cup B$. Let $R = \{(\pi_A^{AB}(x), \pi_B^{AB}(x)) \mid x \in K\}$, $S = \{(\pi_B^{BC}(x), \pi_C^{BC}(x)) \mid x \in L\}$, where $K \subseteq (A \cup B)^{<\omega_1}$ and $L \subseteq (B \cup C)^{<\omega_1}$ are rational languages. Now, the composition $R \circ S$ can be written $R \circ S = \{(\pi_A^{ABC}(x), \pi_C^{ABC}(x)) \mid x \in M\}$, where $M = (\pi_{AB}^{ABC})^{-1}(K) \cap (\pi_{BC}^{ABC})^{-1}(L)$. The inclusion from right to left is straightforward. As for the left to right inclusion, observe that if A, B, C are disjoint, $u \in (A \cup B)^{<\omega_1}$ and $v \in (B \cup C)^{<\omega_1}$ are such that $\pi_B^{AB}(u) = \pi_B^{BC}(v)$ then there exists $w \in (A \cup B \cup C)^{<\omega_1}$ such that $\pi_{AB}^{ABC}(w) = u$ and $\pi_{BC}^{ABC}(w) = v$. Finally, observe that M is rational due to the commutation of π^{-1} with set union, concatenation, Kleene-star, ω-power and ω_1-iteration and second to the closure under intersection of the family of rational sets of transfinite strings. □

4.4 Rational Büchi transducers

Here we modify the notion of finite transducer in such a way as transforming it into a Büchi automaton on A where each transition is equipped with an output in $\mathrm{Rat}(B^{<\omega_1})$. In other words, as for finite and infinite strings, there is an alternative definition where the third component of the transitions are rational subsets of $B^{<\omega_1}$.

Definition **30.** A *Büchi transducer* on transfinite strings is a construct $\mathcal{T} = (Q, A, B, Q_-, E, F)$ where Q, A, B, Q_- are as in Definition 4, E is a finite subset of $(Q \times A \times \mathrm{Rat}(B^{<\omega_1}) \times Q) \cup (2^Q \times Q)$ and F is a mapping from Q to $\mathrm{Rat}(B^{<\omega_1})$.
A *run* of $\mathcal{T}$ on a transfinite input string $(a_\eta)_{\eta<\lambda} \in A^{<\omega_1}$ is a pair of sequences $((q_\eta)_{\eta\leq\lambda}, (X_\eta)_{\eta\leq\lambda})$ where

1) $q_0 \in Q_-$ and for all $\eta < \lambda$ we have $(q_\eta, a_\eta, X_\eta, q_{\eta+1}) \in E$

2) if $\eta \leq \lambda$ is limit then $(P, q_\eta) \in E$ where P is the set of persistent states in $(q_\zeta)_{\zeta<\eta}$.

3) $X_\lambda = F(q_\lambda)$

The *output* of the run is the concatenation product $X = (\prod_{\eta \leq \lambda} X_\eta)$

Transducer $\mathcal{T}$ *defines the relation* $R \subseteq A^{<\omega_1} \times B^{<\omega_1}$ which associates with every transfinite input string u the union set of all outputs of runs on u.

4.5 Eilenberg's second factorization theorem: the case $< \omega_1$

The following technical result happens to be the crux for the transfinite extension of Eilenberg's second factorization theorem.

LEMMA **31.** *Let Q be a finite set, let α and λ be limit ordinals and let $(q_\beta)_{\beta<\alpha}$ be a sequence of elements of Q. Let $(\beta_\eta)_{\eta<\lambda}$ be an increasing continuous (see 2.1) sequence of ordinals which is cofinal to α. Consider the sequence $(Q_\eta)_{\eta<\lambda}$ of subsets of Q where $Q_\eta = \{q_\gamma \mid \beta_\eta \leq \gamma < \beta_{\eta+1}\}$. Then the set of persistent elements in $(q_\beta)_{\beta<\alpha}$ is the union of the sets which are persistent elements in the sequence $(Q_\eta)_{\eta<\lambda}$.*

Proof. Denote by $\{Q^{(1)}, \ldots Q^{(k)}\}$ the collection of persistent elements in the sequence $(Q_\eta)_{\eta<\lambda}$ and set $P = Q^{(1)} \cup \ldots \cup Q^{(k)}$.

1) Let q be some persistent element in the sequence $(q_\beta)_{\beta<\alpha}$, say $q = q_{\gamma_i}$ for some sequence $(\gamma_i)_{i<\omega}$ cofinal to α. Since $(\beta_\eta)_{\eta<\lambda}$ is cofinal to α, for every $i < \omega$ there exists η such that $\gamma_i < \beta_\eta$. Since $(\beta_\eta)_{\eta<\lambda}$ is continuous, the least such η is necessarily a successor ordinal, we denote it by $\eta_i + 1$. Thus, $\gamma_i \in [\beta_{\eta_i}, \beta_{\eta_i+1}[$. Observe that the sequence $(\eta_i)_{i<\omega}$ is necessarily cofinal to λ. In fact, if $\eta < \lambda$ then $\eta \leq \eta_i$ for all $i < \omega$ such that $\beta_\eta < \gamma_i$. Since $q = q_{\gamma_i}$ we have $q \in Q_{\eta_i}$ for all i. Since Q is finite, there exists an infinite set $I \subseteq \omega$ such that all Q_{η_i} , $i \in I$, are equal. Let Q' be their common value. The sequence $(\eta_i)_{i\in I}$ being cofinal to λ (as is $(\eta_i)_{i\in\omega}$) we see that Q' is a persistent element in the sequence $(Q_\eta)_{\eta<\lambda}$. Thus, with the above notations, Q' is among $\{Q^{(1)}, \ldots Q^{(k)}\}$. Since $q \in Q'$ we conclude that $q \in P$.

2) Conversely, let $q \in Q^{(j)}$, with $1 \leq j \leq k$. Then $Q^{(j)} = Q_{\eta_i}$ for some increasing sequence $(\eta_i)_{i<\omega}$ cofinal to λ. Choose an arbitrary element $\epsilon_i \in [\beta_{\eta_i}, \beta_{\eta_i+1}[$ with $q_{\epsilon_i} = q$. Then the sequence $(\epsilon_i)_{i<\omega}$ is cofinal to α and thus q is persistent in $(q_\beta)_{\beta<\alpha}$. □

The equivalence of properties 1 and 4 in the next theorem is a second normalized form for a reciprocal of Proposition 26 introducing *strictly alphabetic morphisms*, i.e. morphisms $\phi : C^{<\omega_1} \to D^{<\omega_1}$ satisfying $\phi(C) \subseteq D$. It is the transfinite version of Eilenberg's second factorization theorem ([14], 1974, p. 248).

THEOREM **32.** *Given a relation $R \subseteq A^{<\omega_1} \times B^{<\omega_1}$, the following properties are equivalent*

1) *R is rational*

2) *R is defined by some 2-tape Büchi automaton*

3) *R is defined by some Büchi transducer*

4) *there exist finite alphabets C, D, a rational subset $K \subseteq C^{<\omega_1}D$, a strictly alphabetic morphism $\phi : C \to A$ and a rational substitution $\sigma : (C \cup D) \to Rat(B^{<\omega_1})$ such that $R = \bigcup_{d\in D, xd\in K}\{\phi(x)\} \times \sigma(xd)$*

Proof. With no loss of generality we can suppose that A and B are disjoint.

1) $\Leftrightarrow$ **2)**. As in the finite string case, there is a one-to-one correspondence between 2-tape Büchi automata on alphabets A, B and one tape Büchi automata working on alphabet $(A \times \{1\}) \cup (\{1\} \times B)$. The relation associated to the 2-tape Büchi automaton being defined as in condition 2 in Proposition 27. Thus, **1)** $\Leftrightarrow$ **2)** is a mere reformulation of that last proposition.

3) $\Rightarrow$ **4)**. Given the Büchi transducer $\mathcal{T} = (Q, A, B, Q_-, E, F)$, let C, D be disjoint finite sets of new symbols such that C is in one-to-one correspondence with the set of quadruples (q, a, X, p) of E, and D is in one-to-one correspondence with the set of pairs $(q, X) \in Q \times \mathrm{Rat}(B^{<\omega_1})$ such that $F(q) = X$. Denote by $[q, a, X, p]$ or $[q, X]$ whichever, the element of C, D in this correspondence. Define a strictly alphabetic morphism $\phi : C \to A$ and a rational substitution $\sigma : (C \cup D) \to \mathrm{Rat}(B^{<\omega_1})$ as follows: $\phi([q, a, X, p]) = a$, $\sigma([q, a, X, p]) = \sigma([q, X]) = X$. Transform $\mathcal{T}$ into an one-tape Büchi automaton $\mathcal{A} = (Q \cup \{q_+\}, C \cup D, Q_-, \{q_+\}, E')$ by defining $(q, [q, a, X, p], p) \in E'$ if and only if $(q, a, X, p) \in E$ and $(q, [q, X], q_+) \in E'$ if and only if $F(q) = X$. Clearly, $\mathcal{A}$ recognizes a subset K of $C^{<\omega_1}D$ such that $R = \{(\phi(x), \sigma(xd)) \mid d \in D, xd \in K\}$

4) $\Rightarrow$ **1)**. This is a direct consequence of Proposition 26.

2) $\Rightarrow$ 3). This is the last implication that remains to be proved. The rest of this paragraph is devoted to its proof. Using condition 2 in Proposition 27, let $R = \{(\pi_A(x), \pi_B(x)) \mid x \in K\}$ where $K \in Rat((A \cup B)^{<\omega_1})$ is recognized by the (1-tape) automaton $\mathcal{A} = (Q, A \cup B, Q_-, Q_+, E)$. We shall associate to $\mathcal{A}$ a Büchi transducer $\mathcal{T}$ which, in a first approach, follows the same construction as for finite strings.

Preliminary analysis
In a run of $\mathcal{A}$, the set of transitions can be divided into those that are labeled by letters in A and those that are labelled by letters in B. Thus, by tracking the A-transitions, we may view a run of $\mathcal{A}$ as starting with a certain number (possibly zero) of B-transitions, followed by one transition in A, followed by a sequence (possible empty) of B-transitions, followed by one A-transitions, etc .... This sequence may be followed by a last sequence of B-transitions.
This grouping process is indeed possible in the transfinite case.

Claim: Every transfinite string w over $A \cup B$ of length λ in the subalphabet A (i.e. $|w|_A = \lambda$) can be factored in a unique way as

$$w = (\prod_{\eta<\lambda} v_\eta a_\eta) v_\lambda \tag{5}$$

with $a_\eta \in A$ and $v_\eta \in B^{<\omega_1}$.
Indeed, let $\pi_A(w) = (a_\eta)_{\eta<\lambda}$. For $\eta < \lambda$ consider the shortest prefix w_η of w containing all occurrences a_ζ with $\zeta < \eta$. Then $w = w_\eta z$ holds for some transfinite string z and v_η is exactly the longest prefix of z not containing an occurrence of A.

Consider an $\mathcal{A}$-run $(q_\beta)_{\beta\leq\alpha}$ on input $w = (\prod_{\eta<\lambda} v_\eta a_\eta)v_\lambda \in (A\cup B)^{<\omega_1}$. where $\alpha = |w|$ and $\lambda = |w_A|$. Let $\beta_\eta = |\prod_{\zeta<\eta} v_\zeta a_\zeta| = \sum_{\zeta<\eta}(|v_\zeta| + 1)$, so that q_{β_η} is the state reached after processing the η first occurrences of A. A direct naive extension to the transfinite of the intuition used for finite strings would define $\mathcal{T}$ so that

1. the $\mathcal{A}$-run $q_{\beta_\eta} \xrightarrow{v_\eta} q_{\beta_\eta+|v_\eta|} \xrightarrow{a_\eta} q_{\beta_{\eta+1}}$ is replaced by a single $\mathcal{T}$-transition $q_{\beta_\eta} \xrightarrow{a_\eta, X_\eta} q_{\beta_{\eta+1}}$ where the output X_η is the rational set

consisting of all strings $v' \in B^{<\omega_1}$ such that there is an $\mathcal{A}$-run $q_{\beta_\eta} \xrightarrow{v'} q_{\beta_\eta+|v_\eta|}$

2. the last part $q_{\beta_\lambda} \xrightarrow{v_\lambda} q_\alpha$ of the $\mathcal{A}$-run is replaced by a last $\mathcal{T}$-output $X_\lambda = F(q_{\eta_\lambda})$ equal to the the rational set consisting of all strings $v' \in B^{<\omega_1}$ such that there is an $\mathcal{A}$-run $q_{\beta_\lambda} \xrightarrow{v'} q_\alpha$

However, we also have to consider limit transitions which ask for a memorization of the set of persistent states in the run. This leads to the following modification in the above assertion 1:

3. the $\mathcal{T}$-output X_η is reduced to the set of all strings $v' \in B^{<\omega_1}$ such that there is an $\mathcal{A}$-run $q_{\beta_\eta} \xrightarrow{v'} q_{\beta_\eta+|v_\eta|}$ and the two runs from q_{β_η} to $q_{\beta_\eta+|v_\eta|}$ with labels v_η and v' visit exactly the same set of states, namely $Q_\eta = \{q_\beta \mid \beta_\eta \leq \beta < \beta_{\eta+1}\}$. That X_η is indeed rational is insured by Lemma 6.

   Also, $\mathcal{T}$-states will have two components: one for the current $\mathcal{A}$-state and the other to memorize the set Q_η. Since that last set is known at the end of the run on input v_η, it will be stored in the second component of the state $(q_{\beta_{\eta+1}}, V_{\eta+1})$ reached by $\mathcal{T}$ after processing letter a_η. For limit η we shall put $V_\eta = \emptyset$.

Construction of the Büchi transducer
This leads to the following construction of $\mathcal{T} = (Q \times 2^Q, A, B, Q_- \times \{\emptyset\}, E', F)$.
Let $\mathcal{A}'$ be the automaton obtained from $\mathcal{A}$ by deleting all A-transitions. As in Lemma 6, for all $q, r \in Q$ and $V \subseteq Q$ we denote by $X_{q,V,r}$ the set of strings v in $B^{<\omega_1}$ such that there is an $\mathcal{A}'$-run from state q to state r which visits exactly the states in V.
Initial or successor transitions
We set $((q, W), a, X, (p, V)) \in E'$ if and only if $X = \bigcup_{r \text{ such that } (r,a,p)\in E} X_{q,V,r}$
.
Limit transitions
For all $S = \{(q_1, V_1), \ldots (q_k, V_k)\}$ we set $(S, (q, \emptyset)) \in E'$ if $(P, q) \in E$ where $P = V_1 \cup \ldots \cup V_k$
Final transitions We set $F((q, V)) = X$ if and only if X is the

family of labels of runs of $\mathcal{A}'$ going from state q to some state in Q_+.

Let us verify that R is exactly the relation associated to transducer $\mathcal{T}$. Suppose $(u,v) \in R$. Then there exists $w \in (A \cup B)^{<\omega_1}$ such that $(u,v) = (\pi_A(w), \pi_B(w))$ and there exists an accepting $\mathcal{A}$-run $(q_\beta)_{\beta \leq \alpha}$ on input w with $\alpha = |w|$. We keep the notations of the preliminary analysis. and show that

1. $(q_{\beta_\eta})_{\eta \leq \lambda}$ is a run of $\mathcal{T}$ on input $u = (a_\eta)_{\eta<\lambda}$ with outputs $(X_\eta)_{\eta \leq \lambda}$
2. $v \in \prod_{\eta \leq \lambda} X_\eta$.

The preliminary analysis and the very definition of $\mathcal{T}$ show that $v_\eta \in X_\eta$ for $\eta < \lambda$. Also, since the $\mathcal{A}$-run on w is accepting, $q_\alpha \in Q_+$ so that $X_\lambda \neq \emptyset$ and $v_\lambda \in X_\lambda$. This proves assertion 2 above.

As for assertion 1, the case of initial and successor steps is clear from the preliminary analysis and the very definition of $\mathcal{T}$.
Concerning limit transitions, first observe that μ is a limit ordinal if and only if $\eta_\mu = |\prod_{\gamma<\mu} v_\gamma a_\gamma|$ is a limit ordinal. In particular, the sequence $(\beta_\eta)_{\eta<\lambda}$ is continuous. Recall that $V_{\eta+1} = Q_\eta$ for all η and $V_\eta = \emptyset$ for limit η. Suppose $\theta \leq \lambda$ is limit. The family of persistent elements in the sequence $(Q_\eta)_{\eta<\theta}$ is exactly that of persistent elements in the sequence $(V_\eta)_{\eta<\theta}$ augmented with the empty set (which appears at limit steps) in case θ has type ≥ 2, i.e. is a limit of limit ordinals. Thus, these families have the same union. Lemma 31 insures that this union set is exactly the set of persistent elements of the sequence $(q_\beta)_{\beta<\beta_\theta}$. The definitions of limit transitions for $\mathcal{A}$ and $\mathcal{T}$ now show that $(q_{\beta_\theta}, \emptyset)$ is a valid limit state of $\mathcal{T}$ for the run on u. This proves assertion 1 above. Thus, (u,v) is in the relation associated to $\mathcal{T}$.

Conversely, suppose (u,v) is in the relation associated to $\mathcal{T}$ and $u = (a_\eta)_{\eta<\lambda}$ and $v = \prod_{\eta\leq\lambda} v_\eta$ where v_η is in the output X_η of $\mathcal{T}$ relative to the transition on letter a_η if $\eta < \lambda$ or in the very last output given by the function F if $\eta = \lambda$. Similar arguments allow to construct an accepting $\mathcal{A}$-run on input $w = (\prod_{\eta<\lambda} v_\eta a_\eta) v_\lambda$, which shows that $(u,v) \in R$. □

4.6 Eilenberg's second factorization theorem: the case $< \omega^{n+1}$

Büchi two-tape automata and Büchi transducers have obvious Choueka counterparts in which limit transitions are treated in the Choueka way. In the notion of Choueka transducer there is no need to restrict the output rational sets to $Rat(B^{<\omega^{n+1}})$, such outputs can be taken in $Rat(B^{<\omega_1})$.

Using Theorem 22 and Proposition 27, we now prove that Theorem 32 implies its Choueka version as a corollary.

THEOREM **33.**
1) Given a relation $R \subseteq A^{<\omega^{n+1}} \times B^{<\omega_1}$, the following properties are equivalent

1) *R is rational*

2) *R is defined by some Choueka transducer*

3) *there exist finite alphabets C, D, a rational subset $K \subseteq C^{<\omega^{n+1}}D$, a strictly alphabetic morphism $\phi : C \to A$ and a rational substitution $\sigma : (C \cup D) \to Rat(B^{<\omega_1})$ such that $R = \bigcup_{d \in D, xd \in K} \{(\phi(x)\} \times \sigma(xd)$*

2) In case $R \subseteq A^{<\omega^{n+1}} \times B^{<\omega^{n+1}}$, the above properties are also equivalent to

4) *R is defined by some 2-tape Choueka automaton*

Moreover, in that case, in condition **3)** *we can suppose σ to have range in* $\mathrm{Rat}(B^{<\omega^{n+1}})$.

Proof. 1) Implications **2)** $\Rightarrow$ **3)** $\Rightarrow$ **1)** go as the corresponding implications **2)** $\Rightarrow$ **4)** $\Rightarrow$ **1)** in Theorem 32.
For implication **3)** $\Rightarrow$ **2)**, on input $u \in A^{<\omega^{n+1}}$, the wanted Choueka transducer non deterministically guesses some string $xd \in C^{<\omega^{n+1}}D$, outputs $\sigma(xd)$ and checks whether $xd \in K$ and $u = \phi(x)$ (which can be done step by step since ϕ is strictly alphabetical).
For implication **1)** $\Rightarrow$ **3)** suppose **1)**. Using the analog implication in Theorem 32, let C, D, $\phi : C \to A$, $\sigma : C \cup D \to \mathrm{Rat}(B^{<\omega_1})$ and $K \in \mathrm{Rat}(C^{<\omega_1}D)$ be such that $R = \bigcup_{d \in D, xd \in K} \{\phi(x)\} \times \sigma(xd)$. Up to a restriction of alphabet C, we can suppose that $\sigma(xd) \neq \emptyset$ for all $xd \in K$. Since ϕ is strictly alphabetic and $\phi(xd) \in A^{<\omega^{n+1}}$ whenever $xd \in K$, we see that $K \subseteq C^{<\omega^{n+1}}$. Now, using Theorem 22, we get

the desired conclusion.
2) Obvious. □

4.7 Recognizable relations

We recall that an ω_1-Wilke algebra is a semigroup S equipped with an additional unary operation $x \to x^\omega$, subject to the two axioms

1. for all x, y, $x(yx)^\omega = (xy)^\omega$ holds
2. for all x, for all integers $n < \omega$, $(x^n)^\omega = x^\omega$ holds.

A subset $K \subseteq A^{<\omega_1}$ is *recognizable* if there exists a morphism ϕ (relative to the structure of ω_1-Wilke algebras) from $A^{<\omega_1}$ onto a finite ω_1-Wilke algebra S and a subset $T \subseteq S$ such that $K = \phi^{-1}(T)$.
The notion of recognizability extends to relations: $R \subseteq A^{<\omega_1} \times B^{<\omega_1}$ is *recognizable* if there exists a morphism ϕ in the category of ω_1-Wilke algebras from $A^{<\omega_1} \times B^{<\omega_1}$ onto a finite ω_1-Wilke algebra S and a subset $T \subseteq S$ such that $R = \phi^{-1}(T)$.
It is clear again that the traditional properties of recognizable sets and relations on finite strings extend to transfinite strings.

PROPOSITION **34.** *1) Recognizable relations on $A^{<\omega_1} \times B^{<\omega_1}$ form a Boolean algebra.*
2) If $\phi : A \to B^*$ is a monoid morphism, and if $K \subseteq B^{<\omega_1}$ is recognizable, then $\phi^{-1}(K)$ is a recognizable subset of $A^{<\omega_1}$.
3) $K \subseteq A^{<\omega_1}$ is recognizable if and only if it is rational.
4) A rational relation is recognizable if and only if there exists an integer p and rational sets $X_1, \ldots, X_p \in Rat(A^{<\omega_1})$ and $Y_1, \ldots, Y_p \in Rat(B^{<\omega_1})$ such that $R = \bigcup_{1 \le i \le p} X_i \times Y_i$.

5 Uniformization

Uniformizing a rational relation in $\mathrm{Rat}(A^{<\omega_1} \times B^{<\omega_1})$ consists of finding a function whose graph is a rational relation and whose domain coincides with that of the given relation.
In the finite case, Eilenberg proved that this can be achieved as follows:

- For relations recognizable by synchonous automata, just choose for each $u \in A^*$ in the domain of the relation the minimal $v \in B^*$ that is associated with u (relative to the length-lexicographical ordering relative to some prescribed order on alphabet B.
- The passage from synchronous to general rational relations uses Eilenberg's second factorization theorem.

However this does not carry over to infinite strings, let alone to transfinite strings. The reason is that the lexicographic order is no more well-founded for infinite strings of any fixed length. As in [10], we will use a "greedy ordering" on the runs on a given input in order to "rationally" choose a second component associated with a given input.

5.1 The greedy ordering on Choueka-continuous sequences

Recall that runs on inputs of length α ($\alpha < \omega_1$) are sequences in $([Q]_0^t)^{\alpha+1}$ (where t is the type of α) which are Choueka-continuous over Q (cf. Definition 13) .

We fix the finite set Q and some finite total orderings $<^0, <^1, \ldots, <^n$ on Q and its successive power sets $[Q]^1, \ldots, [Q]^n$.

The purpose of this subsection is to define a total ordering on the set of all Choueka-continuous sequences of fixed length $\alpha + 1 < \omega^{n+1}$ such that every set of runs associated with a given input possesses a minimal element.

We now detail the inductive construction of an operation which associates to every ordinal $\alpha < \omega^{n+1}$ a total ordering (which we shall call "greedy") $\prec^{\alpha+1}_{\text{greedy}}$ on the set `Choueka`$(Q, \alpha + 1)$ of Choueka-continuous sequences over Q of length $\alpha + 1$. The definition is first given for ordinals of the form $\omega^i + 1$ and uses an induction.

INITIAL CASE $\alpha + 1 = 1$. `Choueka`$(Q, \alpha + 1)$ is just Q and we let $\prec^1_{\text{greedy}}$ be $<^0$.

INDUCTIVE STEP: FROM 1 TO $\omega + 1$ AND FROM $\omega^i + 1$ TO $\omega^{i+1} + 1$ ($i > 0$).
Let $i \geq 0$. To any $(\omega^{i+1} + 1)$-Choueka-continuous sequence ξ we associate a sequence

$$\phi(\xi) = (\mathcal{U},\ \beta_{-1},\ \xi \restriction [0, \beta_{-1}],\ \beta_0,\ \xi \restriction [\beta_{-1}, \beta_0],\ \beta_1,\ \xi \restriction [\beta_0, \beta_1],\ \ldots)$$

The first component is the last element of the sequence ξ, i.e. $\mathcal{U} = \xi(\omega^{i+1})$, and lies in the set $[Q]^{i+1}$. When ξ varies, such elements can be compared via $<^{i+1}$.

The second component β_{-1} is the greatest ordinal $\beta \in \{\omega^i.k \mid k < \omega\}$ such that $\xi(\beta) \notin \mathcal{U}$ (recall that $\mathcal{U}$ is the set of values infinitely often repeated in the sequence $(\xi(\omega^i), \xi(\omega^i.2), \xi(\omega^i.3), \ldots)$, so that such a greatest β does indeed exist).

Let $\mathcal{U} = \{U_0, \ldots, U_{m-1}\}$. For $0 \leq j < m$ and $0 \leq p < \omega$ the component β_{mp+j} is the smallest ordinal $\beta \in \{\omega^i.k \mid k < \omega\}$ such that $\beta_{mp+j-1} < \beta_{mp+j}$ and $\xi(\beta) = U_j$. Such a β does indeed exist since U_j is infinitely often repeated in the sequence $(\xi(\omega^i), \xi(\omega^i.2), \xi(\omega^i.3), \ldots)$.

For $-1 \leq k < \omega$, letting p be such that $\beta_{k+1} = \beta_k + \omega^i.t$, the $(1+2k)$-th component $\xi \restriction [\beta_k, \beta_{k+1}]$ is in $\texttt{Choueka}(Q, \omega^i.t+1)$ if $i > 0$ or in $\texttt{Choueka}(Q,t)$ (i.e. Q^t) if $i = 0$. When ξ varies among strings for which $\phi(\xi)$ has fixed components $\beta_{-1}, \ldots, \beta_{k+1}$, such $(1+2k)$-th components can be compared via the lexicographic t-power of $\prec^{\omega^i+1}_{\text{greedy}}$ if $i > 0$ or of $\prec^1_{\text{greedy}}$ if $i = 0$.

The greedy ordering $\prec^{\omega^{i+1}+1}_{\text{greedy}}$ is defined from $\prec^{\omega^i+1}_{\text{greedy}}$ (in case $i > 0$) or from $\prec^1_{\text{greedy}}$, i.e. $<$ (in case $i = 0$) as follows. To compare two different $(\omega^{i+1}+1)$-Choueka-continuous sequences η, ξ, we consider the sequences $\phi(\xi)$ and $\phi(\eta)$, look at the first component on which they differ and compare ξ, η according to these components. Though the $(1+2k)$-th component of $\phi(\xi)$ lies in a set depending on ξ, such a comparison really makes sense. In fact, if η, ξ cannot be compared via their first $2k$ components, then their $1+2k$-th components lie in the very same set $\texttt{Choueka}(Q, \gamma+1)$ (where γ is of the form $\omega^i.t$ for some $t \geq 0$).

Clearly, $\prec^{\omega^{i+1}+1}_{\text{greedy}}$ is a total ordering on $\texttt{Choueka}(Q, \omega^{i+1}+1)$.

The greedy ordering on $\texttt{Choueka}(Q, \alpha+1)$.
Let $\alpha = \omega^{i_1} + \omega^{i_2} + \ldots + \omega^{i_m}$ where $\omega > i_1 \geq i_2 \geq \ldots \geq i_m \geq 0$ and $0 \leq m < \omega$. For $0 < j \leq m$ set $\sigma_j = \omega^{i_1} + \omega^{i_2} + \ldots + \omega^{i_j}$.

To each $(\alpha+1)$-sequence ξ we associate the m-sequence:

$$\theta(\xi) = (\xi \restriction [0, \sigma_1],\ \xi \restriction [\sigma_1, \sigma_2],\ \ldots,\ \xi \restriction [\sigma_{m-1}, \sigma_m])$$

of Choueka-continuous sequences of lengths ω^{i_1} , ω^{i_2} ,..., ω^{i_m}.

We define the greedy ordering $\prec^{\alpha+1}_{\text{greedy}}$ on $\texttt{Choueka}(Q, \alpha+1)$ as follows. To compare two different $(\alpha+1)$-Choueka-continuous sequences

η, ξ, we consider the m-tuples $\theta(\eta)$ and $\theta(\xi)$, we look at the first component on which they differ, say it has rank j, and compare ξ, η according to the greedy ordering $\prec_{\text{greedy}}^{\omega^{i_j}+1}$ on this component.

5.2 Two properties of the greedy ordering

We first prove that for every $n < \omega$ the union of the greedy orderings $\prec_{\text{greedy}}^{\alpha+1}$ for $\alpha < \omega^{n+1}$ (which is a partial ordering comparing strings having the same length $< \omega^{n+1}$) is synchronous rational.

LEMMA **35.** *There exists a formula $\Phi(\eta, \xi)$ of second order monadic logic (in the language of order) which uniformly defines the relation*

$$\{(\eta, \xi) \mid \exists\alpha < \omega^{n+1}(\eta, \xi \in \texttt{Choueka}(Q, \alpha+1) \text{ and } \eta \prec_{greedy}^{\alpha+1} \xi\}$$

in any structure $(\theta, <)$ where $\theta \geq \omega^{n+1}$ (Convention: as usual, strings over an alphabet with m letters are interpreted as t-tuples of (bounded) subsets of θ with $t = \lceil \log(m) \rceil$).
In particular, this relation is rational and even synchronous rational, i.e. recognized by an automaton which reads its tapes in a synchronous way.

Proof. 1) First, we express types of ordinals (cf. 2.1) in second order monadic logic. Let $\texttt{Lim}(\alpha) \equiv \forall\beta < \alpha \ \ \exists\gamma \ (\beta < \gamma < \alpha)$ asserts that α is a limit ordinal. Then

$$\texttt{Type}_0(\alpha) \equiv \neg\texttt{Lim}(\alpha) \qquad \texttt{Type}_{\geq 1}(\alpha) \equiv \texttt{Lim}(\alpha)$$

and more generally for $n \geq 1$:

$$\texttt{Type}_{\geq n}(\alpha) \equiv \forall\beta < \alpha \ \ \exists\gamma \ (\texttt{Type}_{\geq n-1}(\beta) \wedge \beta < \gamma < \alpha)$$

We may express the exact type by

$$\texttt{Type}_n(\alpha) \equiv \texttt{Type}_{\geq n}(\alpha) \wedge \neg\texttt{Type}_{\geq n+1}(\alpha)$$

If $\beta < \gamma < \omega^{n+1}$ then by relativizing in the above formulas all quantifiers to the interval $[\beta, \gamma[$, we can express the type of this interval.
2) Thus, the relations

$$i \leq n \text{ and } \gamma = \beta + \omega^i$$
$$i \leq n \text{ and } \exists k < \omega \ (\gamma = \beta + \omega^i.k)$$
$$i \leq n \text{ and } u = (\xi(\omega^i) \ , \ \xi(\omega^i.2) \ , \ \xi(\omega^i.3) \ , \ldots)$$
$$i \leq n \text{ and } u = (\xi \restriction [0\omega^i] \ , \ \xi \restriction [\omega^i\omega^i.2] \ , \ \xi \restriction [\omega^i.k\omega^i.(k+1)] \ , \ \ldots)$$

are also expressible.

3) The prefix and lexicographic orderings on transfinite sequences on some finite ordered alphabet are easy to express by second order formulas.
If $\prec$ is an expressible ordering over strings in $\texttt{Choueka}(Q, \omega^i + 1)$ then so is the relation

$$\beta, \gamma \in \{\omega^i.k \mid k < \omega\} \text{ and } \xi \restriction [\beta, \gamma] \prec^{[\beta,\gamma]\text{-lex}} \eta \restriction [\beta, \gamma]$$

where $\prec^{[\beta,\gamma]\text{-lex}}$ is the lexicographic extension of $\prec$ to strings indexed in $[\beta, \gamma]$.
Consequently, for $i \leq n$, the greedy orderings $\prec^{\omega^i+1}_{\text{greedy}}$ are expressible.
4) In order to express the greedy ordering $\prec^{\alpha+1}_{\text{greedy}}$ there remains to deal with the decomposition sequences of ordinals $\alpha < \omega^{n+1}$, namely

$$\alpha = \omega^{i_1} + \omega^{i_2} + \ldots + \omega^{i_m}$$

where $\omega > i_1 \geq i_2 \geq \ldots \geq i_m \geq 0$ and $0 \leq m < \omega$.
For $\beta < \gamma$ let's say that $[\beta, \gamma]$ is a block if it has order type ω^i for some $i \leq n$ and the type of β is at least i. Clearly, blocks of α are the pieces of the decomposition sequence of α. Now, the relation

$$[\beta, \gamma] \text{ is a block of } \alpha \text{ and } u = \xi \restriction [\beta, \gamma]$$

is easy to express via types. □

Let $\alpha < \omega^{n+1}$. We consider the compact product topology on the product set $\prod_{\beta \leq \alpha} [Q]^{\tau(\beta)}$ (where $\tau(\beta)$ denotes the type of β, cf. §2.1). Since this set contains $\texttt{Choueka}(Q, \alpha+1)$, it induces a topology on this last set (Caution: the induced topology is not compact!).

Lemma **36.** *Every closed non empty subset of* $\texttt{Choueka}(Q, \alpha + 1)$ *has a smallest element for the* $\prec^{\alpha+1}_{greedy}$ *ordering.*

Proof. The way $\prec^{\alpha+1}_{\text{greedy}}$ is defined from the $\prec^{\omega^i+1}_{\text{greedy}}$ orderings makes it clear that we can reduce to the case where $\alpha = 0$ or $\alpha = \omega^i$. The first case is trivial. We argue by induction on i for the second case. following the construction detailed in paragraph 5.1 of $\prec^{\omega^{i+1}+1}_{\text{greedy}}$ from $\prec^{\omega^i+1}_{\text{greedy}}$ in case $i > 0$ or from $\prec^{1}_{\text{greedy}} = <$ in case $i = 0$.

Let F be a non empty closed subset of $\texttt{Choueka}(Q, \omega^{i+1} + 1)$. We want to minimize the sequence

$$\phi(\xi) = (\mathcal{U},\ \beta_{-1},\ \xi \restriction [0, \beta_{-1}],\ \beta_0,\ \xi \restriction [\beta_{-1}, \beta_0],\ \beta_1,\ \xi \restriction [\beta_0, \beta_1],\ \ldots)$$

where ξ varies in F.

Let $\mathcal{V}$ be the smallest value of the first component $\mathcal{U} = \xi(\omega^{i+1})$ of $\phi(\xi)$ when ξ varies in F. Restrict F to the subset F' of strings $\xi \in F$ with first component $\mathcal{V}$. Clearly, F' is still closed and non empty.

The following argument is a variation from our paper [10] p. 68 about uniformization. We inductively define a sequence of integers $(k_1\ ,\ k_2\ ,\ k_3, \ldots)$ and a sequence of strings $(\zeta_1\ ,\ \zeta_2\ ,\ \zeta_3\ , \ldots)$ as follows.

Let λ_{-1} be the smallest value of the second component β_{-1} of $\phi(\xi)$ when ξ varies in F'. We let k_{-1} be such that $\lambda_{-1} = \omega^i.k_{-1}$ and we define ζ_1 as the smallest value of $\xi \upharpoonright [0, \omega^i.k_{-1}]$ (with respect to the lexicographic k_{-1}-power of $\prec^{\omega^i+1}_{\text{greedy}}$) when ξ varies over sequences such in F' such that $\beta_{-1} = \lambda_{-1}$.

For $t \geq 0$ let λ_t be the smallest value of the component β_t of $\phi(\xi)$ for $\xi \in F'$ extending the concatenation of strings $\zeta_1, \ldots, \zeta_{t-1}$. We let k_t be such that $\lambda_t = \omega^i.k_t$ and we define ζ_t as the smallest value of $\xi \upharpoonright [\omega^i.k_{t-1}, \omega^i.k_t]$ (with respect to the lexicographic $(k_t - k_{t-1})$-power of $\prec^{\omega^i+1}_{\text{greedy}}$) for $\xi \in F'$ extending the concatenation of strings $\zeta_1, \ldots, \zeta_{t-1}$ and such that $\beta_t = \lambda_t$.

Let ζ be the concatenation of the sequence of strings $\zeta_1\ ,\ \zeta_2\ ,\ \zeta_3\ , \ldots$ and of the element $\mathcal{V}$. It is clear that ζ has length $\omega^{i+1} + 1$ and that $\zeta(\omega^i.k) \in \mathcal{V}$ whenever $k > k_{-1}$. Also, each element $U_j \in \mathcal{V}$ appears infinitely often in the sequence $(\zeta(\omega^i), \zeta(\omega^i.2, \ldots)$ since it appears in every string ζ_{mp+j} for $p \geq 0$. This gives the Choueka-continuity condition for level ω^{i+1}. As for levels $\omega, \ldots, \omega^i$, the Choueka-continuity conditions are inherited from the ζ_t's. Thus, $\zeta \in \texttt{Choueka}(Q, \omega^{i+1}+1)$. By very construction ζ is the limit of strings in F, hence is in F. Also, by definition of the $\prec^{\omega^i+1}_{\text{greedy}}$ ordering, it is the smallest string in F. $\square$

5.3 Uniformization of relations with domain bounded below ω^ω

We first state a simple lemma.

Lemma **37.** *The set of accepting runs of a Choueka automaton on an input with length $\alpha < \omega^\omega$ is a closed subset of* $\texttt{Choueka}(Q, \alpha + 1)$.

Proof. Suppose $\zeta \in \texttt{Choueka}(Q, \alpha + 1)$ is the pointwise limit of runs r_0, r_1,... on input u. Then for every $\alpha <\mid u \mid$ there exists t such that

$\zeta(\alpha) = r_t(\alpha)$ and $\zeta(\alpha+1) = r_t(\alpha+1)$. Since r_t is a run, the triple $(\zeta(\alpha), u_\alpha, \zeta(\alpha+1))$ is in the transition relation. Thus, ζ satisfies the initial and successor conditions for runs. Also, there exists t such that $\zeta(|\ u\ |) = r_t(|\ u\ |)$. Since r_t is a run, $\zeta(|\ u\ |)$ is a final state. Thus, ζ satisfies the final condition for runs.
Finally, since ζ is Choueka-continuous, it also satisfies the limit condition for runs. Thus, ζ is indeed an accepting run. □

THEOREM **38.** *Every rational transfinite relation* $R \subseteq A^{<\omega^{n+1}} \times B^{<\omega_1}$ *can be uniformized by a rational relation.*

Proof. Consider a Choueka transducer defining R. Without loss of generality, one can suppose that for each transition (q, X, r) the output $X \in \mathrm{Rat}(B^{<\omega_1})$ is uniquely determined by the state r. For each such non empty X choose a witness string $v_X \in X$ (there are only finitely many choices since the X's involved in transitions are finitely many). According to the previous lemma, given an input $u \in A^\alpha$ (where necessarily $\alpha < \omega^{n+1}$), the set of accepting runs on u is a closed subset of $\texttt{Choueka}(X, \alpha+1)$. Arguing in second order monadic logic, Lemmas 37, 36, 35, allow to definably, hence rationally, associate to each input u a uniquely determined run. Now, from the run we rationally go to the transfinite sequence of ouput rational sets and also to the final output set. From this sequence and set we rationally get the transfinite sequence of witness strings and the final string, the concatenation of which gives a unique v such that $(u, v) \in R$. □

5.4 Uniformization of relations with range bounded below ω^ω

THEOREM **39.** *Every rational transfinite relation* $R \subseteq A^{<\omega_1} \times B^{<\omega^{n+1}}$ *can be uniformized by a rational relation.*

Proof. □

5.5 Uniformization of relations with relationality degree bounded below ω^ω

A close examination of the proof of Thm. 39 shows that the important point is not that the range consists of bounded strings but that the

factorizations $(u,v) = \prod_{\alpha<\theta}(u_\alpha,1)(1,v_\alpha)$ of the input associated to the runs on input (u,v) are such that $\theta < \omega^{n+1}$. I.e. the number of alternations of the tape readings is bounded by ω^{n+1}.
Formalizing this notion, we get a symmetric extension of Theorems 38, 39.

5.6 From ω^ω on

The following two technical properties are elementary properties which can be found resp. in [24] and [11].

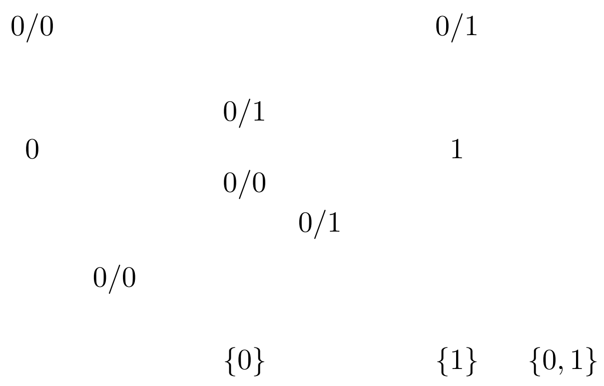
LEMMA **40.** *Let* $0 < k < \omega$. *If* $\alpha_i < \omega^k$ *for all* $i < \omega$ *then* $\sum_{i<\omega} \alpha_i \leq \omega^k$.

LEMMA **41.** *For any two transfinite strings* x, y, *equality* $y = xy$ *holds if and only if* x^ω *is a prefix of* y.

Given a limit ordinal $\lambda < \omega^\omega$ we denote by W_λ the set of strings $w \in \{0,1\}^\lambda$ such that for all non trivial factorizations $w = w_1 w_2$, the prefix w_1 contains finitely many occurrences of the symbol 1 and w_2 contains at least one occurrence of 1. This is equivalent to saying that the positions of the occurrences of 1 define an ω-sequence which is cofinal with λ. Consider the relation

$$\texttt{CofinalSeq} = \{(0^\lambda, w) \mid \lambda \text{ is limit } < \omega^\omega \text{ and } w \in W_\lambda\}$$

It is easy to see that $\texttt{CofinalSeq} = \mathit{Rel}(\mathcal{A}) \cap (\{0\}^{<\omega^\omega} \times \{0,1\}^{<\omega^\omega})$ where $\mathit{Rel}(\mathcal{A})$ is the relation recognized by the following *deterministic synchronous* 2-tape automaton (where the final state $\{1\}$ takes into account the case $\lambda = \mu + \omega$ and w has a suffix 1^ω).



Extending Def. 11 to relations (and observing that Thm. 21 goes through for relations), let's write $\mathrm{Rat}(A^{<\alpha} \times B^{<\beta})$ in place of $\{R \cap (A^{<\alpha} \times B^{<\beta}) \mid R$ is Büchi recognizable$\}$.

PROPOSITION **42.** *The relation* `CofinalSeq`*, which is in* $\mathrm{Rat}(A^{<\omega^\omega} \times B^{<\omega^\omega})$*, where* $A = B = \{0, 1\}$*, is not uniformizable.*

Proof. Let R be a rational relation on the aphabet $\{0, 1\}$ which, for every $i \in \omega$, accepts some $(0^{\omega^i}, u_i) \in$ `CofinalSeq`. The proof consists in showing that for some integer i and some string $v \neq u_i$ both $(0^{\omega^i}, u_i)$ and $(0^{\omega^i}, v)$ belong to R. We assume the relation is recognized by some Büchi automaton to which any run refers and we denote by π_1 and π_2 the projection of the labels of a run onto $A^{<\omega^\omega}$ and $B^{<\omega^\omega}$ respectively (the "first" and the "second" component). For all pairs $(x, y) \in A^{<\omega^\omega} \times B^{<\omega^\omega}$ let us denote by $M(x, y)$ the Boolean $Q \times Q$-matrix whose (q, r)-entry is 1 if and only if there is a run labelled by (x, y) which leads from state q to state r.

Some familiarity with run factorizations are necessary to understand the technique used in the proof. Given a run ρ in an 2-automaton and its label (ξ, η) it is not the case that each factorization of the label is the label of a prefix of the run, which is a big departure from the 1-automaton case. What can be guaranteed is the following which could be called a "run factorization driven by an output factorization". Let α be an ordinal and let $\prod_{r<\alpha} \eta_r$ be a factorization of η. Then there exists a (non necessarily unique) factorization of ρ into $\rho = \prod_{r<\alpha} \rho_r$ such that $\pi_2(\rho_r) = \eta_r$ for all $r < \alpha$.
Apply this observation to some $(0^{\omega^i}, u) \in$ `CofinalSeq`. Since the

length $|u|_1$ is ω, the output label can be factored as $u = \prod_{r<\omega} 0^{\xi_r}1$. Lemma 40 shows in particular (and this will be used below heavily) that for some $r < \omega$, $\xi_r \geq \omega^{i-1}$ holds. The idea of the proof is to substitute some factor ρ'_r for the r-th factor ρ_r in the run driven by the output factorization, in such a way as to keep the input label and modify the output label (actually what we do is slightly different but the basic idea is there). This method is akin to that used in the famous "pumping lemma" but it is more elaborate in the sense that the factor ρ that we plug in does not require considering a single run but rather an infinite collection of runs.

Let us now be more technical. For all $i < \omega$, fix a successful run ρ_i labelled by $(0^{\omega^i}, u_i)$. We consider the following predicate

$$\text{the image by } \pi_1 \text{ of some proper prefix of } \rho_i \text{ equals } 0^{\omega^i} \tag{6}$$

(Observe, as an example, no run in the above synchronous automaton recognizing `CofinalSeq` satifies the predicate).

Case 1: there exist infinitely many runs ρ_i satisfying predicate (6). For every such run we choose an arbitrary factorization $\rho_i = h_i b_i t_i$ ("head", "body" and "tail") such that $\pi_1(h_i) = 0^{\omega^i}$ and $\pi_2(b_i) = 0^{\omega^{i-1}}$ (the existence of such a b_i is guaranteed by Lemma 40). Since Q is finite there exist two integers $k < j$ such that ρ_k and ρ_j satisfy predicate (6) and $M(\texttt{label}(b_k)) = M(\texttt{label}(b_j))$ holds. The two runs $h_j b_j t_j$ and $h_j b_k t_j$ are successful and we claim that they have the same input labels and two different output labels. Indeed, since $|\pi_1(b_j)| = |\pi_1(b_k)| = |\pi_1(t_j)| = |\pi_1(t_k)| = 0$ holds, we clearly have $|\pi_1(h_j b_j t_j)| = |\pi_1(h_j b_k t_j)|$ and therefore $\pi_1(h_j b_j t_j) = \pi_1(h_j b_k t_j)$. Concerning the output labels, the condition $\pi_2(h_j b_j t_j) = \pi_2(h_j b_k t_j)$ would yield, after cancelling out the common prefix $\pi_2(h_j b_k)$, equality $0^{\omega^{j-1}}\pi_2(t_j) = 0^{\omega^{k-1}}\pi_2(t_j)$, hence $\pi_2(t_j) = 0^{\omega^{j-1}}\pi_2(t_j)$. This implies by Lemma 41, that $\pi_2(t_j)$ has a prefix equal to 0^{ω^j}, a contradiction to the fact that letter 1 occurs cofinally in u_j.

Case 2: for sufficiently large i, ρ_i does not satisfy predicate (6). For all sufficiently large i, choose a fatorization $\rho_i = h_i b_i t_i$ such that $\texttt{label}(b_i) = (0^{\xi_i}, 0^{\omega^{i-1}})$ for some $\xi_i < \omega^i$ (this condition is guaranteed by Lemma 40). There exist two integers k, j such that $i < k < j$ and

$M(\texttt{label}(b_k)) = M(\texttt{label}(b_j))$ holds. Then the two runs $\rho_j = h_j b_j t_j$ and $h_j b_k t_j$ are successful and we claim that they have the same input labels and two different output labels. Indeed, $|\pi_1(\rho_j)| = \omega^j$ and $|\pi_1(t_j)| \neq 0$ implies $|\pi_1(t_j)| = \omega^j$ and $|\pi_1(h_j b_j)| < \omega^j$, thus $|\pi_1(h_j b_k)| < \omega^j$ and finally $|\pi_1(h_j b_k t_j)| = \omega^j$ proving the equality of the two input labels (both equal to 0^{ω^j}). Concerning the output labels, the condition $\pi_2(h_j b_j t_j) = \pi_2(h_j b_k t_j)$ would yield, by cancelling out the common prefix $\pi_2(h_j)0^{\omega^k}$ equality $\pi_2(t_j) = 0^{\omega^{j-1}}\pi_2(t_j)$. By Lemma 41 this implies that $\pi_2(t_j)$ has a prefix equal to 0^{ω^j}, a contradiction to the fact that letter 1 occurs cofinally in u_j. $\square$

References

[1] N. Bedon. Finite automata and ordinals. *Theoret. Comput. Sci.*, 156(1–2):119–144, 1996.

[2] N. Bedon. Automata, semigroups and recognizability of words on ordinals. *Internat. J. Algebra Comput.*, 8(1):1–21, 1998.

[3] N. Bedon. *Langages reconnaissables de mots indexés par des ordinaux.* PhD thesis, Université de Marne la Vallée, 1998.

[4] N. Bedon and O. Carton. An Eilenberg theorem for words on countable ordinals. In *Proceedings of Latin'98 Theoretical Informatics*, number 1380 in LNCS, pages 53–64. Springer-Verlag, 1998.

[5] B. Bérard and C. Picaronny. Accepting Zeno words without making time stand still. In *Proceedings of the 22nd. Symp. Math. Found. Comp. Sci. (MFCS'97, Bratislava, Slovakia)*, number 1295 in LNCS, pages 149–158. Springer-Verlag, 1997.

[6] J. Berstel. *Transductions and context-free languages.* B. G. Teubner, 1979.

[7] Alexis Bès. Decision problems related to the elementary theory of oordinal multiplication. 1999. Submitted to Fundamenta Mathematicae.

[8] J. Büchi. Decision methods in the theory of ordinals. *Bull. AMS*, 71:767–770, 1965.

[9] J. Büchi and D. Siefkes. The monadic theory of ω_1. In *Decidable Theories II*, number 328 in Lecture Notes in Mathematics. Springer-Verlag, 1973.

[10] C. Choffrut and Serge Grigorieff. Uniformization of rational relations. In G. Paun J. Karhumäki, H. Maurer and G. Rozenberg, editors, *Jewels are Forever*, pages 59–71, 1999.

[11] C. Choffrut and S. Horváth. Transfinite equations in transfinite strings. *Internat. J. Algebra Comput.*, 10(5):625–649, 2000.

[12] C. Choffrut and J. Karhumäki. Combinatorics of words. In G. Rozenberg and A. Salomaa, editors, *Handbook of Formal Languages*, volume 1, pages 329–438. World Scientific, 1997.

[13] S. C. Choueka. Finite automata, definable sets and regular expressions over ω^n-tapes. *J. Comput. System Sci.*, 17:81–97, 1978.

[14] S. Eilenberg. *Automata, languages and machines*, volume A. Academic Press, 1974.

[15] C.C. Elgot and J.E. Mezei. On Relations Defined by Finite Automata. *IBM Journal*, 10:47–68, 1965.

[16] F. Gire and M. Nivat. Relations rationnelles infinitaires. *Calcolo*, XXI:91–125, 1984.

[17] M. R. Hansen, P. K. Pandya, and A. C. Zhou. Finite divergence. *Theoret. Comput. Sci.*, 138:113–139, 1995.

[18] H. Johnson. Rational equivalence relations. *Theoret. Comput. Sci.*, 47:39–60, 1986.

[19] D. Klaua. *Konstruktion ganzer, rationaler und reeller Ordinazahlen und die diskontinuierliche Struktur der transfiniten reellen Zahlenräume*. Akademie-Verlag, Berlin (DDR), 1961.

[20] D. Klaua. *Allgemeine Mengenlehre*. Akademie Verlag, 1969.

[21] Lothaire. *Combinatorics on Words*, volume 17 of *Encyclopedia of Mathematics and its Applications*. Addison-Wesley, gian-carlo rota edition, 1983.

[22] M. Nivat. Transductions des langages de Chomsky. *ANN. Inst. Fourier*, 108:339–456, 1968.

[23] J. G. Rosenstein. *Linear ordering.* Academic Press, New-York, 1982.

[24] W. Sierpinski. *Cardinal and Ordinal Numbers.* Warsaw: PWN, 1958.

[25] T. Wilke. An Eilenberg theorem for ∞-languages. In *Proc. of 18th ICALP Conference*, number 510 in LNCS, pages 588–599. Springer-Verlag, 1991.

[26] T. Wilke. An algebraic theory for regular languages of finite and infinite words. *Internat. J. Algebra Comput.*, 3:447–489, 1993.

[27] J. Wojciechowski. Finite automata on transfinite sequences and regular expressions. *Fundamenta Informaticae*, 8(3–4):379–396, 1985.