A

Regular Languages are Church-Rosser Congruential

VOLKER DIEKERT, University of Stuttgart
MANFRED KUFLEITNER, University of Stuttgart
KLAUS REINHARDT, University of Tübingen
TOBIAS WALTER, University of Stuttgart

This paper shows a general result about finite monoids and weight reducing string rewriting systems. As a consequence it proves a long standing conjecture in formal language theory: All regular languages are Church-Rosser congruential. The class of Church-Rosser congruential languages was introduced by McNaughton, Narendran, and Otto in 1988. A language L is Church-Rosser congruential if there exists a finite, confluent, and length-reducing semi-Thue system S such that L is a finite union of congruence classes modulo S. It was known that there are deterministic linear context-free languages which are not Church-Rosser congruential, but the conjecture was that all regular languages are of this form. The paper shows a stronger statement: A language is regular if and only if it is strongly Church-Rosser congruential. It is the journal version of the conference abstract which was presented at ICALP 2012.

Categories and Subject Descriptors: F.4.2 [**Mathematical logic and formal languages**]: Grammars and Other Rewriting Systems—*Thue systems*; F.4.3 [**Mathematical logic and formal languages**]: Formal Languages—*Algebraic language theory*

General Terms: Theory

Additional Key Words and Phrases: Church-Rosser system, local divisor, regular language, semigroup, string rewriting

ACM Reference Format:
Diekert, V., Kufleitner, M., Reinhardt, K., and Walter, T. 2015. Regular Languages are Church-Rosser Congruential. *J. ACM* V, N, Article A (January YYYY), 19 pages.
DOI: http://dx.doi.org/10.1145/0000000.0000000

1. INTRODUCTION

The notion of Church-Rosser congruential language appeared first in Narendran's PhD thesis [Narendran 1984]. The thesis led to a systematic study of Church-Rosser languages in a joint work by McNaughton, Narendran, and Otto which appeared in [McNaughton et al. 1988]. A main motivation to consider Church-Rosser languages is that the word problem can be solved in linear time: This is done by computing normal forms on input words by using a finite, confluent, and length-reducing string rewriting system. Once an irreducible normal form is obtained, membership can be decided by checking whether the irreducible normal form belongs to some finite table. As common, a string rewriting system over a finite alphabet A is called a semi-Thue system. We are interested in finite systems, only. This finiteness assumption is part of the following definition. A language $L \subseteq A^*$ is called Church-Rosser congruential, if there exists a finite, confluent, and length-reducing semi-Thue system $S \subseteq A^* \times A^*$ such

The research on this paper was initiated during the program *Automata Theory and Applications* at the Institute for Mathematical Sciences, National University of Singapore in September 2011. The support and hospitality of NUS is greatly acknowledged.
The second author was supported by the German Research Foundation (DFG) under grant DI 435/5-1.

DOI: http://dx.doi.org/10.1145/0000000.0000000

that L is a finite union of congruence classes modulo S. If, in addition, the index of S is finite (i.e., the monoid A^*/S of all congruence classes is finite) then L is called *strongly* Church-Rosser congruential.

It is not hard to see that $\{a^n b^n \mid n \geq 1\}$ is Church-Rosser congruential, but $\{a^m b^n \mid m \geq n\}$ is not. This led the authors of [McNaughton et al. 1988] to the more technical notion of Church-Rosser languages which captures all deterministic context-free languages. In [Niemann and Otto 2005] Church-Rosser languages were characterized as *deterministic growing context-sensitive* languages. For more results about Church-Rosser languages see e.g. [Buntrock and Otto 1998; Narendran 1984; Woinowski 2001; Woinowski 2003]. Since $\{a^m b^n \mid m \geq n\}$ is deterministic context-free, the class of Church-Rosser languages is strictly larger than the class of Church-Rosser congruential languages. It was conjectured that all regular languages are Church-Rosser congruential. After some significant initial progress towards a solution of this conjecture [Narendran 1984; Niemann 2002; Niemann and Otto 2005; Niemann and Waldmann 2002; Reinhardt and Thérien 2003] there was stagnation.

Before 2011 the most advanced result was the one announced in 2003 by Reinhardt and Thérien [Reinhardt and Thérien 2003]. According to this manuscript the conjecture is true for all regular languages where the syntactic monoid is a group. Unfortunately, the manuscript has never been published as a refereed paper and there are some flaws in its presentation. The main problem with [Reinhardt and Thérien 2003] has however been quite different for us. The statement is too weak to be useful in the induction for the general case. So, instead of being able to use [Reinhardt and Thérien 2003] as a black box, we needed to prove a more general result in the setting of weight-reducing systems. This part about group languages is a cornerstone in our approach.

The other ingredient to our paper has been established only very recently. Knowing that the result is true if the syntactic monoid is a group, it was natural to investigate aperiodic monoids. Finite aperiodic monoids do not have non-trivial groups as subsemigroups. They correspond to star-free languages; and the first two authors together with Weil proved that all star-free languages are Church-Rosser congruential [Diekert et al. 2012b]. It became possible by *loading the induction hypothesis* leading to a much stronger statement: For every star-free language $L \subseteq A^*$ there exists a finite confluent semi-Thue system $S \subseteq A^* \times A^*$ with the following properties. The quotient monoid A^*/S is finite and aperiodic, L is a union of congruence classes modulo S, and moreover all right-hand sides of rules appear as scattered subwords in the corresponding left-hand side. The last property is called *subword-reducing*, and it is obvious that every subword-reducing system is length-reducing. However, we have little hope that such a strong result holds outside aperiodic languages, in general. Indeed, here we step back from subword-reducing to weight-reducing systems. Thus, we combine a stronger result than stated in [Reinhardt and Thérien 2003] together with a weaker result than shown in [Diekert et al. 2012b] for aperiodic languages. The proof in [Diekert et al. 2012b] used crucially the construction of *local divisors*. The same is true here.

Theorem 5.1 states the following result: Let $L \subseteq A^*$ be a regular language and $\|a\| \in \mathbb{N} \setminus \{0\}$ be a positive weight for every letter $a \in A$ (e.g., $\|a\| = |a| = 1$). Then we can construct a finite, confluent and weight-reducing semi-Thue system $S \subseteq A^* \times A^*$ such that the quotient monoid A^*/S is finite and recognizes L. In particular, L is a finite union of congruence classes modulo S. As a consequence, a language is regular if and only if it is strongly Church-Rosser congruential.

This paper therefore solves a problem which was open for about 25 years after the journal publication [McNaughton et al. 1988]. If we consider Nivat's paper in 1970 [Nivat 1970] (where this kind of questions has been initiated) or Narendran's PhD thesis [Narendran 1984] as starting point, one can say it was open for an even longer period. The solution to this problem became possible by proving a general algebraic re-

sult that homomorphisms from free monoids to finite monoids factorize through finite, confluent, and weight reducing semi-Thue systems.

The present paper is the journal version of the conference abstract [Diekert et al. 2012a]. The present paper contains full proofs and improvements concerning the presentation.

2. PRELIMINARIES

Throughout this paper, A is a finite alphabet. An element of A is called a *letter*. The set A^* is the free monoid generated by A. It consists of all finite sequences of letters from A. The elements of A^* are called *words*. The empty word is denoted by 1. The *length* of a word u is denoted by $|u|$. We have $|u| = n$ for $u = a_1 \cdots a_n$ where $a_i \in A$. The empty word has length 0, and it is the only word with this property. The set of words of length at most n is denoted by $A^{\leq n}$, and the set of all nonempty words is A^+. We generalize the length of a word by introducing weights. A *weighted alphabet* $(A, \|\cdot\|)$ consists of an alphabet A equipped with a weight function $\|\cdot\| : A \to \mathbb{N} \setminus \{0\}$. The *weight* of a letter $a \in A$ is $\|a\|$ and the *weight* $\|u\|$ of a word $u = a_1 \cdots a_n$ with $a_i \in A$ is $\|a_1\| + \cdots + \|a_n\|$. The weight of the empty word is 0. Length is the special weight with $\|a\| = 1$ for all $a \in A$.

We use the standard notation from combinatorics on words: A word u is a *factor* of a word v if there exist $p, q \in A^*$ such that $puq = v$, and u is a *proper factor* of v if $pq \neq 1$. The word u is a *prefix* of v if $uq = v$ for some $q \in A^*$, and it is a *suffix* of v if $pu = v$ for some $p \in A^*$. We say that u is a *factor* (resp. *prefix*) of v^+ if u is a factor (resp. prefix) of $v^{|u|}$. We denote the set of factors of a word v by $\mathrm{Factors}(v)$ and the set of factors which are factors of some word in v^+ by $\mathrm{Factors}(v^+)$. Two words $v, w \in A^*$ are *conjugate*, if there exist $p, q \in A^*$ such that $v = pq$ and $w = qp$. Note that if u a is factor of v^+ and if v, w are conjugate, then u is a factor of w^+, too. An integer $m > 0$ is a *period* of a word $u = a_1 \cdots a_n$ with $a_i \in A$ if $a_i = a_{i+m}$ for all $1 \leq i \leq n - m$. A word $u \in A^+$ is *primitive* if there exists no $v \in A^+$ such that $u = v^n$ for some integer $n > 1$. It is a standard fact that a word u is not primitive if and only if $u^2 = puq$ for some $p, q \in A^+$. This follows immediately from the result from combinatorics on words that $xy = yx$ if and only if x and y are powers of a common root; see e.g. [Lothaire 1983, Section 1.3].

An equivalence relation $\sim \subseteq A^* \times A^*$ is called a *congruence* if $u \sim v$ implies $xuy \sim xvy$ for all $u, v, x, y \in A^*$. The set of congruence classes $[u] = \{v \in A^* \mid u \sim v\}$ forms a quotient monoid of A^* by defining $[u] \cdot [v] = [uv]$. For a language $L \subseteq A^*$ the *syntactic congruence* is defined by $u \sim v$ if we have $xuy \in L \Leftrightarrow xvy \in L$ for all $x, y \in A^*$. The quotient monoid is called the *syntactic monoid* of L. It is denoted by $\mathrm{Synt}(L)$ and the canonical homomorphism is denoted by $\pi_L : A^* \to \mathrm{Synt}(L)$. As usual, a language $L \subseteq A^*$ is called *regular* if $\mathrm{Synt}(L)$ is finite. There are various other well-known characterizations of regular languages; e.g., regular expressions, finite automata or monadic second order logic. Here we use another equivalent definition. A monoid M *recognizes* L if there exists a homomorphism $\varphi : A^* \to M$ such that $L = \varphi^{-1}\varphi(L)$. We also say that φ recognizes L in this case. If $\varphi : A^* \to M$ recognizes L, then the syntactic homomorphism π_L factorizes through φ. This classical observation shows that a language $L \subseteq A^*$ is regular if and only if it is recognized by some finite monoid.

Regular languages L can be classified in terms of structural properties of the monoids recognizing L. In particular, we consider *group languages*; these are languages recognized by finite groups. At the other end in the spectrum of regular languages are *aperiodic* languages. These are languages recognized by finite monoids where all groups which occur as a subsemigroup are trivial.

In this paper, a *semi-Thue system* over A is a finite subset $S \subseteq A^* \times A^*$. The elements of S are called *rules*. We frequently write $\ell \to r$ for the rule (ℓ, r). A system S is called *length-reducing* if we have $|\ell| > |r|$ for all rules $\ell \to r$ in S. It is called *weight-reducing*

with respect to some weighted alphabet $(A, \|\cdot\|)$ if $\|\ell\| > \|r\|$ for all rules $\ell \to r$ in S. Every system S defines a rewriting relation $\underset{S}{\Longrightarrow} \subseteq A^* \times A^*$ by setting $u \underset{S}{\Longrightarrow} v$ if there exist $p, q, \ell, r \in A^*$ such that $u = p\ell q$, $v = prq$, and $\ell \to r$ is in S. By $\overset{*}{\underset{S}{\Longrightarrow}}$ we mean the reflexive and transitive closure of $\underset{S}{\Longrightarrow}$. By $\overset{*}{\underset{S}{\Longleftrightarrow}}$ we mean the symmetric, reflexive, and transitive closure of $\underset{S}{\Longrightarrow}$. We also write $u \overset{*}{\underset{S}{\Longleftarrow}} v$ whenever $v \overset{*}{\underset{S}{\Longrightarrow}} u$. The system S is *confluent* if for all $u \overset{*}{\underset{S}{\Longleftrightarrow}} v$ there is some w such that $u \overset{*}{\underset{S}{\Longrightarrow}} w \overset{*}{\underset{S}{\Longleftarrow}} v$. It is *locally confluent* if for all $v \underset{S}{\Longleftarrow} u \underset{S}{\Longrightarrow} v'$ there exists w such that $v \overset{*}{\underset{S}{\Longrightarrow}} w \overset{*}{\underset{S}{\Longleftarrow}} v'$. It is *terminating* if there are no infinite chains $u_1 \underset{S}{\Longrightarrow} u_2 \underset{S}{\Longrightarrow} u_3 \underset{S}{\Longrightarrow} \cdots$. Weight-reducing systems are terminating since then $u \underset{S}{\Longrightarrow} v$ implies $\|u\| > \|v\|$.

In order to check that a system S is locally confluent, it is enough to show the following two statements.

(i) For all $(\ell, r), (\ell', r') \in S$ where $x\ell = \ell' z$ and $|x| < |\ell'|$ there is some $w \in A^*$ with $xr \overset{*}{\underset{S}{\Longrightarrow}} w \overset{*}{\underset{S}{\Longleftarrow}} r'z$.

(ii) For all $(\ell, r), (\ell', r') \in S$ where $\ell = y\ell' z$ there is some $w \in A^*$ with $r \overset{*}{\underset{S}{\Longrightarrow}} w \overset{*}{\underset{S}{\Longleftarrow}} yr'z$.

The corresponding pairs $(xr, r'z)$ and $(r, yr'z)$ are so-called *critical pairs*. The important property is that a finite system has only finitely many critical pairs. They arise from words where left-hand sides ℓ and ℓ' overlap as follows:

overlap critically factor critically

Fig. 1: Sources of critical pairs

Throughout we use the classical result that terminating and locally confluent systems are confluent. The proof is easy and can be found in textbooks, see e.g. [Book and Otto 1993] or [Jantzen 1988].

The relation $\overset{*}{\underset{S}{\Longleftrightarrow}} \subseteq A^* \times A^*$ is a congruence. The monoid of congruence classes $[u]_S = \{v \in A^* \mid u \overset{*}{\underset{S}{\Longleftrightarrow}} v\}$ is denoted by A^*/S. The size of A^*/S is called the *index* of S. A finite semi-Thue system S can be viewed as a finite set of defining relations. By $\mathrm{IRR}_S(A^*)$ we denote the set of irreducible words in A^*, i.e., the set of words where no left-hand side occurs as a factor.

Assume that $L \subseteq A^*$ is recognized by some homomorphism $\varphi : A^* \to M$ where M is finite. As every finite monoid is finitely presented there exists a finite semi-Thue system S such that $M = A^*/S$. The finite system S can be chosen to be terminating and confluent. However, it is not true that S can be chosen length- or weight-reducing, in general. Whenever the weighted alphabet $(A, \|\cdot\|)$ is fixed, a finite semi-Thue system $S \subseteq A^* \times A^*$ is called a *weighted Church-Rosser system* if it is finite, weight-reducing for $(A, \|\cdot\|)$, and confluent. Hence, a finite semi-Thue system S is a weighted Church-Rosser system if and only if (1) we have $\|\ell\| > \|r\|$ for all rules $\ell \to r$ in S and (2) every congruence class has exactly one irreducible element. In particular, for weighted Church-Rosser systems S there is a one-to-one correspondence between A^*/S

and $\mathrm{IRR}_S(A^*)$. A *Church-Rosser system* is a finite, length-reducing, and confluent semi-Thue system. In particular, every Church-Rosser system is a weighted Church-Rosser system for $(A, |\cdot|)$. A language $L \subseteq A^*$ is called a *Church-Rosser congruential language* if there is a finite Church-Rosser system S such that L can be written as a finite union of congruence classes $[u]_S$.

Example 2.1. Consider the following three languages $L_1 = \{a^n b^n \mid n \geq 1\}$, $L_2 = \{a^m b^n \mid m \geq n\}$, and $L_3 = a(ba)^*$ over the two-letter alphabet $A = \{a, b\}$. The languages L_1 and L_2 are non-regular but deterministic context-free whereas L_3 is regular. The first language L_1 is Church-Rosser congruential. The corresponding system is $S_1 = \{a^2b^2 \to ab\}$; and we have $L_1 = [ab]_{S_1}$.

The complement of L_1 is not Church-Rosser congruential. Indeed assume that S' is a Church-Rosser system such that we can write $A^* \setminus L_1$ as a finite union of congruence classes. Then some congruence class must contain words $a^k b$ and $a^m b$ with $k > m \geq 1$. But then $a^k b^m$ and $a^m b^m$ share the same class, too. This is impossible since $a^k b^m \notin L_1$ and $a^m b^m \in L_1$. A very similar argument shows that neither L_2 nor its complement $A^* \setminus L_2$ are Church-Rosser congruential.

Consider $L_3 = a(ba)^*$. It is Church-Rosser congruential due to the system $S = \{aba \to a\}$. With respect to S all words a^n are irreducible. In particular, the monoid A^*/S is infinite. Hence, S has infinite index. An explicit Church-Rosser system T for L_3 of finite index has been constructed in [Diekert et al. 2012b]. It is given by

$$\begin{aligned} T = \{\ & bbb \to bb,\ bba \to bb,\ abb \to bb,\ bab \to b, \\ & aaa \to bb,\ aab \to bb,\ baa \to bb,\ aba \to a\}. \end{aligned}$$

The monoid $\{a, b\}^*/T$ has seven elements: $[1]_T$, $L_3 = [a]_T$, $[b]_T$, $[ab]_T$, $[ba]_T$, $[aa]_T$, and $[bb]_T$. It is not the smallest monoid recognizing L_3, because aa and bb behave as a "zero" and could be identified. The smallest monoid recognizing L_3 is its syntactic monoid and has 6 elements.

The observation in Example 2.1 leads to the notion of *strongly Church-Rosser congruential language* [Niemann 2002] and the corresponding class sCRCL. Let CRCL denote the class of Church-Rosser congruential languages. The subclass sCRCL is defined as the family of languages $L \subseteq A^*$ which can be written as union of congruence classes w.r.t some Church-Rosser system of finite index. In particular, such a congruence refines the syntactic congruence of L; as a consequence L is regular. Thus, $\mathrm{sCRCL} \subseteq \mathrm{REG}$. Here and in the following REG denotes the class of all regular languages. By Example 2.1:

$$\begin{aligned} & L_1 \in \mathrm{CRCL} \setminus \mathrm{sCRCL}, \\ & L_2 \notin \mathrm{CRCL}, \\ & L_3 \in \mathrm{sCRCL} \subseteq \mathrm{REG}. \end{aligned}$$

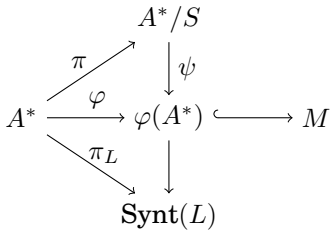
Our goal is to show $\mathrm{REG} \subseteq \mathrm{sCRCL}$, thus $\mathrm{REG} = \mathrm{sCRCL}$. This is stated in Corollary 5.2. Actually, we shall prove a statement about finite monoids rather than on regular languages in which the following definition is central.

Definition 2.2. Let $\varphi : A^* \to M$ be a homomorphism and let S be a semi-Thue system. We say that φ *factorizes through* S if for all $u, v \in A^*$ we have:

$$u \underset{S}{\Longrightarrow} v \quad \text{implies} \quad \varphi(u) = \varphi(v).$$

As a special case, we obtain that $L \in \mathrm{sCRCL}$ if and only if the syntactic homomorphism of L factorizes through some Church-Rosser system S of finite index. More generally, let $S \subseteq A^* \times A^*$ be any semi-Thue system such that $\varphi : A^* \to M$ factorizes through S,

then $\psi(\pi(u)) = \varphi(u)$ is well-defined, where $\pi(u) = [u]_S$ is the canonical homomorphism. Moreover, if φ recognizes L, then we obtain the following commutative diagram.



Our goal is to show that every homomorphism $\varphi : A^* \to M$ to a finite monoid factorizes through a Church-Rosser system of finite index. This aim is achieved by Theorem 5.1.

3. FINITE GROUPS

Our main result is that every homomorphism $\varphi : A^* \to M$ to a finite monoid M factorizes through a Church-Rosser system S. We distinguish whether or not M is a group; and we first prove this result for groups. Before we turn to the general group case, we show that proving the claim for some particular groups is easy. The techniques developed here will also be used when proving the result for arbitrary finite groups.

3.1. Groups without proper cyclic quotient groups

The aim of this section is to show that finding a Church-Rosser system is easy for some cases. This list includes presentations of finite (non-cyclic) simple groups, but it goes beyond this. Let $\varphi : A^* \to G$ be a homomorphism to a finite group, where $(A, \|\cdot\|)$ is a weighted alphabet. This defines a regular language $L_G = \{w \in A^* \mid \varphi(w) = 1\}$. Let us assume that the greatest common divisor $\gcd\{\|w\| \mid w \in L_G\}$ satisfies $\gcd\{\|w\| \mid w \in L_G\} = \gcd\{\|a\| \mid a \in A\}$. This happens e.g. if $\{6, 10, 15\} \subseteq \{\|w\| \mid w \in L_G\}$, because then $\gcd\{\|w\| \mid w \in L_G\} = \gcd\{\|a\| \mid a \in A\} = 1$. Then there are two words $u, v \in L_G$ such that $\|u\| - \|v\| = q$, where $q = \gcd\{\|a\| \mid a \in A\}$. We can use these words u and v to find a constant $d \in q\mathbb{N}$ such that all $g \in G$ have a representing word v_g with the exact weight $\|v_g\| = d$. To see this, start with some arbitrary set of representing words v_g. We multiply words v_g with minimal weight by u and all other words v_g by v until all weights are equal. The final step is to define the following weight-reducing system

$$S_G = \left\{ w \to v_{\varphi(w)} \;\middle|\; w \in A^* \text{ and } d < \|w\| \le d + \max\{\|a\| \mid a \in A\} \right\}.$$

Confluence of S_G is straightforward: Let $w \in A^*$. If $\|w\| \le d$, then no rule applies to w and w is irreducible. Next, we prove by induction that for all $w \in A^*$ with $\|w\| > d$ there exists a derivation $w \overset{*}{\underset{S_G}{\Longrightarrow}} v_{\varphi(w)}$ with $\left\|v_{\varphi(w)}\right\| = d$. Thus we consider $w \in A^*$ with $\|w\| > d$. Then there exists a factorization $w = uv$ with $d < \|u\| \le d + \max\{\|a\| \mid a \in A\}$. Since $u \to v_{\varphi(u)}$ is a rule and $\varphi(v_{\varphi(u)}) = \varphi(u)$, we deduce $w \underset{S_G}{\Longrightarrow} v_{\varphi(u)}v \overset{*}{\underset{S_G}{\Longrightarrow}} v_{\varphi(w)}$ by induction. Moreover, if $w \overset{*}{\underset{S_G}{\Longrightarrow}} \widehat{w}$ is any derivation such that $\widehat{w}$ is irreducible with respect to S_G, then $\|\widehat{w}\| = d$, because $\|w\| > d$. Since $\varphi(w) = \varphi(\widehat{w})$ and $\|\widehat{w}\| = d$, we conclude $\widehat{w} = v_{\varphi(w)}$. Thus, S_G is indeed a Church-Rosser system of finite index such that φ factorizes through S_G.

Now assume $\gcd\{\|w\| \mid w \in L_G\} > \gcd\{\|a\| \mid a \in A\}$. Without loss of generality, we have $\gcd\{\|a\| \mid a \in A\} = 1$. Then there is a prime number p such that p divides $\|w\|$

for all $w \in L_G$. The image $G' = \varphi(A^*)$ is a subgroup of G since G is finite. Define $\varphi' : G' \to \mathbb{Z}/p\mathbb{Z}$ by $\varphi'(g) = \|u\| \bmod p$ if $\varphi(u) = g$. This is well-defined, because for $\varphi(u) = \varphi(v)$ there exists $w \in A^*$ with $\varphi(uw) = \varphi(vw) = 1$. Therefore p divides both $\|uw\|$ and $\|vw\|$. Hence $\|u\| \equiv \|v\| \bmod p$. Since φ' is surjective, we see that $\mathbb{Z}/p\mathbb{Z}$ becomes a quotient group of G'. This can never happen if $\varphi(A^*)$ is a simple and non-cyclic subgroup of G, because a simple group does not have any proper quotient group. But there are many other cases where a natural homomorphism $A^* \to G$ for some weighted alphabet $(A, \|\cdot\|)$ satisfies the property $\gcd\{\|w\| \mid w \in L_G\} = 1$ although the subgroup $\varphi(A^*)$ of G has a non-trivial cyclic quotient group. Just consider the length function and a presentation by standard generators for dihedral groups D_{2n} or the permutation groups $\mathcal{S}_n$ where n is odd.

In order to have a concrete example, let $G = D_6 = \mathcal{S}_3$ be the permutation group of a triangle. The group G is generated by elements τ and ρ with defining relations $\tau^2 = \rho^3 = 1$ and $\tau\rho\tau = \rho^2$. It has $\mathbb{Z}/2\mathbb{Z}$ as a quotient. Still, $\gcd(|\tau^2|, |\rho^3|) = \gcd(2, 3) = 1$; and the following six words of length 3 represent all six group elements:

$$1 = \rho^3,\ \rho = \rho\tau^2,\ \rho^2 = \tau\rho\tau,\ \tau = \tau^3,\ \tau\rho = \rho^2\tau,\ \tau\rho^2.$$

More systematically, one could obtain a normal form of length 5 for each of the group elements in $\{1, \rho, \rho^2, \tau, \tau\rho, \tau\rho^2\}$ by adding factors ρ^3 and τ^2. For example, this could lead to the set of normal forms $\{\tau^2\rho^3, \tau^4\rho, \rho^5, \tau^5, \tau\rho^4, \tau^3\rho^2\}$. We will use this pumping idea in our proof of the general case for finding normal forms of approximately the same size.

It is much harder to find a Church-Rosser system for the homomorphism $\varphi : \{a, b, c\}^* \to \mathbb{Z}/3\mathbb{Z}$ where $\varphi(a) = \varphi(b) = \varphi(c) = 1 \bmod 3$. Restricting φ to the submonoid $\{a, b\}^*$ makes the situation simpler. Still it is surprisingly complicated. A possible Church-Rosser system $S \subseteq \{a, b\}^* \times \{a, b\}^*$ of finite index such that the restriction of φ factorizes through S is given by:

$$S = \left\{ aaa \to 1,\ baab \to b,\ (ba)^3b \to b,\ bb\,u\,bb \to b^{|u|+1} \;\middle|\; 1 \le |u| \le 3 \right\}.$$

There are 273 irreducible elements and the longest irreducible word has length 16. Note that the last set of rules has bb as a prefix and as a suffix on both sides of every rule. The idea of preserving end markers such as $\omega = bb$ in the above example is essential for the solution of the general case, too.

In some sense this phenomenon suggests that finite cyclic groups or more general commutative groups form an obstacle for constructing Church-Rosser systems.

3.2. The general case for group languages

3.2.1. Outline. The proof that group languages are Church-Rosser congruential uses induction on the size of the alphabet. We will show that every homomorphism $\varphi : A^* \to G$ factorizes through a weighted Church-Rosser system S of finite index using the following road map: For $|A| > 1$ we remove some letter c from the alphabet A. This leads to a system R over the remaining letters B. Lemma 3.1 allows us to assume that all words of any given finite set are irreducible. Then we set $K = \mathrm{IRR}_R(B^*)c$ which is a prefix code in A^*. We consider K as a new alphabet. Essentially, it is this situation where weighted alphabets come into play, because we can choose the weight of K such that it is compatible with the weight over the alphabet A. We introduce two sets of rules T_Δ and T_Ω over K. The T_Δ-rules reduce long repetitions of short words Δ, and the T_Ω-rules have the form $\omega\,u\,\omega \to \omega\,v_g\,\omega$. Here, Ω is some finite set of markers, $\omega \in \Omega$ is such a marker and the word v_g is a normal form for the group element g. The T_Ω-rules reduce long words without long repetitions of short words. We show that T_Δ and T_Ω are confluent and that their union has finite index over K^*. The confluence of T_Δ

is Lemma 4.2. The confluence of T_Ω relies on several combinatorial properties of the normal forms v_g and the markers Ω. Using Lemma 3.2, we see that all sufficiently long words are reducible. Since by construction all rules in $T = T_\Delta \cup T_\Omega$ are weight-reducing, the system T is a weighted Church-Rosser system over K^* with finite index such that $\varphi : K^* \to G$ factorizes through T. Since $K \subseteq A^*$, we can translate the rules $\ell \to r$ in T over K^* to rules $c\ell \to cr$ over A^*. This leads to the set of T'-rules over A^*. The letter c at the beginning of the T'-rules is required to shield the T'-rules from R-rules. Finally, we show that $S = R \cup T'$ is the desired system over A^*.

3.2.2. Group languages are Church-Rosser congruential. In this section, we consider the general case of groups.

LEMMA 3.1. *Let $(A, \|\cdot\|)$ be a weighted alphabet, $e \in \mathbb{N}$ be some natural number, and let $S \subseteq A^* \times A^*$ be a weighted Church-Rosser system such that $\mathrm{IRR}_S(A^*)$ is finite. Then*

$$S_e = \{u\ell v \to urv \mid u, v \in A^e \text{ and } \ell \to r \in S\}$$

is a weighted Church-Rosser system satisfying:

- *(i) The mapping $[u]_{S_e} \mapsto [u]_S$ for $u \in A^*$ is well-defined and yields a surjective homomorphism from A^*/S_e onto A^*/S.*
- *(ii) All words of length at most $2e$ are irreducible with respect to S_e.*
- *(iii) The set $\mathrm{IRR}_{S_e}(A^*)$ is finite.*

PROOF. Since S is weight-reducing, the system S_e is weight-reducing, too. For all $w, w' \in A^*$ and $u, v \in A^e$, we have $w \overset{*}{\underset{S}{\Longrightarrow}} w'$ if and only if $uwv \overset{*}{\underset{S_e}{\Longrightarrow}} uw'v$. Moreover, rules of S_e apply only to words of length more than $2e$ and an application leaves the prefix and suffix of length e invariant. Hence, confluence of S transfers to confluence of S_e. Thus, S_e is indeed a weighted Church-Rosser system. Since $w \overset{*}{\underset{S_e}{\Longleftrightarrow}} w'$ implies $w \overset{*}{\underset{S}{\Longleftrightarrow}} w'$ for all $w, w' \in A^*$, we obtain $[u]_{S_e} \subseteq [u]_S$ and thus assertion (i) holds.

All words of length at most $2e$ belong to $\mathrm{IRR}_{S_e}(A^*)$. This yields assertion (ii). More precisely, we can write $\mathrm{IRR}_{S_e}(A^*)$ as a disjoint union

$$\mathrm{IRR}_{S_e}(A^*) = A^{<2e} \cup A^e \cdot \mathrm{IRR}_S(A^*) \cdot A^e.$$

Since $\mathrm{IRR}_S(A^*)$ is finite by hypothesis, the set $\mathrm{IRR}_{S_e}(A^*)$ is finite, too. This shows assertion (iii). □

LEMMA 3.2. *Let $d \geq 1$, $u \in K^*$, and $F = \cup_{\delta \in \Delta} \mathrm{Factors}(\delta^+)$ where K is a finite alphabet and $\Delta \subseteq K^+$ is a set of words of length at most d. Then the following assertion holds. If u has the property that $\mathrm{Factors}(u) \cap K^{\leq 2d} \subseteq F$, then we have $u \in F$. (This means: If $u \in K^*$ is a word such that every factor of u of length at most $2d$ appears as a factor of δ^+ for some $\delta \in \Delta$, then u itself is a factor of δ^+ for some $\delta \in \Delta$.)*

PROOF. We may assume that $|u| > 2d$. Write $u = awb$ for $a, b \in K$. Then, by induction on $|u|$, the prefix aw is a factor of δ^+ and wb is a factor of η^+ for some $\delta, \eta \in \Delta$. Let $p = |\delta|$ and $q = |\eta|$. Note that p is a period of aw and q is a period of wb. Thus p and q are both periods of w. Since $|w| \geq 2d - 1 \geq p + q - \gcd(p, q)$, we see that $\gcd(p, q)$ is also a period of w by the Periodicity Lemma of Fine and Wilf, see e.g. [Lothaire 1983, Section 1.3]. By symmetry we may assume $p \leq q$; and we can write $q + 1 = p + 1 + kr$ for some $k \in \mathbb{N}$ with $r = \gcd(p, q)$. Since the $(p + 1)$-th letter in aw is a, the $(q + 1)$-th letter in aw is a, too. By replacing, if necessary, η by some conjugate we may actually assume that wb is a prefix of η^+. The $(q + 1)$-th letter in aw becomes the last letter of η, because $q = |\eta|$. It follows that awb is a factor of η^+. □

THEOREM 3.3. *Let $(A, \|\cdot\|)$ be a weighted alphabet and let $\varphi : A^* \to G$ be a homomorphism to a finite group G. Then there exists a weighted Church-Rosser system S of finite index such that φ factorizes through S.*

We reduce the proof of Theorem 3.3 to the proof of Proposition 3.4 stated below. The proof of Proposition 3.4 is given in Section 4. We do not pay too much attention to finding a "small" Church-Rosser system S. Even in its present form, the pure existence proof (without optimization on the system size) is rather technical. We decided therefore to prefer conceptual simplicity over system size.

Note that $\varphi : A^* \to G$ factorizes through S if and only if $\varphi : A^* \to \varphi(A^*)$ factorizes through S. Therefore we may assume that $G = \varphi(A^*)$, G is non-trivial, and $|A| \geq 1$. In particular, it is enough to show Theorem 3.3 under the assumption that φ is surjective. In the following n denotes the exponent of G; this is the least positive integer n such that $g^n = 1$ for all $g \in G$. The proof is by induction on the size of the alphabet A. Choose some letter $c \in A$. If $A = \{c\}$, then we set $S = \{c^n \to 1\}$. Let now $B = A \setminus \{c\}$ and $B = \{a_0, \ldots, a_{s-1}\}$ with $s \geq 1$. We choose a_0 to have minimal weight among the letters of B. For $i \in \mathbb{N}$ define words γ_i by

$$\gamma_i = a_{i \bmod s}^{n+\lfloor i/s \rfloor}\, c. \tag{1}$$

In particular, $\gamma_0 = a_0^n c$, $\gamma_1 = a_0^{n+1} c$ for $s = 1$, $\gamma_1 = a_1^n c$ for $s \geq 2$, $\gamma_s = a_0^{n+1} c$, and for $k \geq 0$ we have $\gamma_{ks} = a_0^{n+k} c$. The weight of every γ_i is larger than $n \|a_0\|$. This fact will be used later, e.g., in the proof of Lemma 4.4. The set $\{c, a_0 c, \ldots, a_{s-1} c\}$ generates G; and all group elements $\varphi(c)$ and $\varphi(a_j c)$ with $0 \leq j < s$ occur infinitely often as some $\varphi(\gamma_i)$ (e.g., $\varphi(c) \in G$ occurs for $i = kns$ and $\varphi(a_j c) \in G$ occurs for $i = (kn+1)s + j$ and $k \geq 0$). Hence, there exists m with $1 \leq m \leq |G| \cdot n \cdot |A|$ such that for every $g \in G$ there exists a word

$$v_g = \gamma_0 \gamma_0^{n_0} \gamma_1^{n_1} \cdots \gamma_m^{n_m} \gamma_m \gamma_0 \tag{2}$$

with $n_i \geq 0$ satisfying $\varphi(v_g) = g$ and $\|v_g\| - \|v_h\| < n \|a_0\|$ for all $g, h \in G$. The latter property relies on $\|\gamma_0\| + \|a_0\| = \|\gamma_s\|$ and that we may choose $m \geq s$. Indeed, assume $\|v_g\| - \|v_h\| \geq n \|a_0\|$ for some $g, h \in G$. For those v_g with maximal weight replace the exponent n_0 of γ_0 by $n_0 + n$; for all other words v_h replace the exponent n_s of γ_s by $n_s + n$. After that, the maximal difference $\|v_g\| - \|v_h\|$ has decreased at least by 1. (The decrease is at most $n \|a_0\|$. The decrease does not exceed 1 in general, because there might have been a word v_f with $\|v_g\| = \|v_f\| + 1$.) The image in G did not change since $\varphi(\gamma_0^n) = \varphi(\gamma_s^n) = 1$. We iterate this procedure until the weights of all v_g differ by less than $n \|a_0\|$. In the following we fix the number m and we let

$$\Gamma = \{\gamma_0, \ldots, \gamma_m\}.$$

By induction on the size of the alphabet there exists a weighted Church-Rosser system $R \subseteq B^* \times B^*$ of finite index such that the restriction $\varphi : B^* \to G$ factorizes through R. Note that induction applies to $\varphi : B^* \to G$ whether or not the restriction of φ to B^* is surjective. By Lemma 3.1, we may choose R such that $\Gamma \subseteq \mathrm{IRR}_R(B^*)\, c$. Let

$$K = \mathrm{IRR}_R(B^*)\, c.$$

The set K is a finite prefix code in A^* with $\Gamma \subseteq K$. We consider K as an extended alphabet and its elements as extended letters. The free monoid K^* is viewed as the subset $K^* \subseteq A^*$. The weight $\|u\|$ of $u \in K$ is its weight as a word over A. Each word $\gamma_i \in \Gamma$ is a letter in K. The restriction of the homomorphism $\varphi : A^* \to G$ to K^* induces a homomorphism $\psi : K^* \to G$; it is given by $\psi(u) = \varphi(u)$ for $u \in K$. We define a lexical order on A by $a_0 < \cdots < a_{s-1} < c$ which leads to the length-lexicographic order on $B^* c$. (Words are compared first by length, and if they have equal length, they are compared

in lexicographic order.) The length-lexicographic order induces a linear order $\leq$ on $\mathrm{IRR}_R(B^*)c$ and hence also a linear order on the extended alphabet K. Equations (1) and (2) show that the words v_g satisfy as words over the weighted alphabet $(K, \|\cdot\|)$ the following five properties:

(i) Each word v_g starts with the extended letter γ_0.
(ii) The last two extended letters of v_g are $\gamma_m\gamma_0$.
(iii) From left to right all extended letters in v_g are in non-decreasing order with respect to $\leq$ with the sole exception of the last letter γ_0, which is smaller than its predecessor γ_m.
(iv) All extended letters in v_g have a weight greater than $n\,\|a_0\|$.
(v) All differences $\|v_g\| - \|v_h\|$ are smaller than $n\,\|a_0\|$.

PROPOSITION 3.4. *There exists a weighted Church-Rosser system $T \subseteq K^* \times K^*$ of finite index such that $\psi : K^* \to G$ factorizes through T.*

Let us postpone the proof of Proposition 3.4 to Section 4 and finish the proof of Theorem 3.3 first. Recall that every element in K^* can be read as a sequence of elements in A^*. Thus, every element $u \in K^*$ can be interpreted as a word $u \in A^*$ when applying rules in T to words in A^* (which are in fact irreducible with respect to R). Since we must not destroy K-letters, we guard the first K-letter of every T-rule by prepending the letter c. This leads to the system

$$T' = \{c\ell \to cr \in A^* \times A^* \mid \ell \to r \in T\}\,.$$

Combining the rules R over the alphabet B with the T'-rules yields

$$S = R \cup T'.$$

Since left-hand sides of R-rules and of T'-rules do not overlap, the system S is confluent. By definition, each S-rule is weight-reducing. This means that S is a weighted Church-Rosser system. The sets $\mathrm{IRR}_S(A^*)$ and A^*/S are finite. Since $\ell \to r$ in S satisfies $\varphi(\ell) = \varphi(r)$, the homomorphism φ factorizes through S. Thus, the system S satisfies the assertion of Theorem 3.3. This reduces the proof of Theorem 3.3 to the proof of Proposition 3.4.

4. PROOF OF PROPOSITION 3.4

The difference between Proposition 3.4 and Theorem 3.3 is that the (much larger) alphabet K satisfies more hypotheses than A. We show Proposition 3.4 from an abstract viewpoint. An overview of some notation which will be used in this section is summarized in Table I.

Table I: Overview of some notation in this section

d, m, n, κ	positive natural numbers
$(K, \|\cdot\|)$	finite weighted alphabet with linear order $<$
$\psi : K^* \to G$	homomorphism, w.l.o.g. surjective
$v_g \in K^+$	normal form for $g \in G$
$\Gamma = \{\gamma_0, \ldots, \gamma_m\} \subseteq K$	with $\gamma_0 < \cdots < \gamma_m$ and $\|\gamma_i\| > \kappa$
$\Delta = K \cup \{\delta \in K^+ \mid \|\delta\| \leq \kappa\}$	in particular, $\Gamma \subseteq K \subseteq \Delta \subseteq K^{\leq d}$
F	set of factors of all δ^+ with $\delta \in \Delta$
$J \subseteq K^{\leq 2d}$	minimal such that $K^*JK^* = K^* \setminus F$ (i.e., J is a basis of the ideal $K^* \setminus F$)
$\Omega \subseteq J$	maximal such that $\Omega \cap \Gamma K^* = \{\gamma\gamma' \mid \gamma, \gamma' \in \Gamma, \gamma > \gamma'\}$
	and linear ordered such that $\gamma_m\gamma_0 \prec b\gamma_0 \prec \omega$ for $b \neq \gamma_m$ and $\omega \in \Omega \setminus K^+\gamma_0$
$t < t_0 < \ldots < t_{\|\Omega\|} < t_\Omega$	"threshold" values
$T_\Delta, T_\Omega \subseteq T'_\Omega, T = T_\Delta \cup T_\Omega$	semi-Thue systems
	T_Δ, T_Ω and T are weighted Church-Rosser systems

In a first step we fix $\kappa = n\,\|a_0\|$, and we view κ as a constant which is attached to the finite weighted alphabet $(K, \|\cdot\|)$. The set K contains a linearly ordered subset $\Gamma = \{\gamma_0, \ldots, \gamma_m\}$ with $\gamma_0 < \cdots < \gamma_m$ such that $\|\gamma\| > \kappa$ for all $\gamma \in \Gamma$. In addition we require that there exists a homomorphism $\psi : K^* \to G$ and a subset $\widehat{G} \subseteq \Gamma^*$ with the following properties.

(i) We have $\widehat{G} \subseteq \gamma_0\gamma_0^*\gamma_1^* \cdots \gamma_m^*\gamma_m\gamma_0$.
(ii) For each $g \in G$ there is exactly one word $v_g \in \widehat{G}$ with $\psi(v_g) = g$.
(iii) For all $g, h \in G$ we have $\|v_g\| - \|v_h\| < \kappa$.

Note that (ii) implies that ψ is surjective which we assume without restriction. Let us define a subset $\Delta \subseteq K^+$ and a parameter d as follows.

$$\Delta = K \cup \left\{\delta \in K^+ \;\middle|\; \|\delta\| \leq \kappa\right\} \text{ and } d = \max\{|\delta| \mid \delta \in \Delta\}. \tag{3}$$

The set Δ is closed under conjugation, i.e., if $uv \in \Delta$ for $u, v \in K^*$, then $vu \in \Delta$. We let $F \subseteq K^*$ be the set of all factors of δ^+ where $\delta \in \Delta$, i.e., we set

$$F = \left\{u \in K^* \;\middle|\; u \text{ is factor of } \delta^+ \text{ for some } \delta \in \Delta\right\}.$$

Let $u\gamma v \in F \cap K^*\gamma K^*$, i.e. $u\gamma v$ is a factor of δ^+ for some $\delta \in \Delta$. Since $\|\gamma\| > \kappa$, we conclude $\delta = \gamma$ and $u\gamma v \in \gamma^+$. Thus, we obtain

$$K^*\gamma K^* \cap F = \gamma^+ \text{ for all } \gamma \in \Gamma. \tag{4}$$

By definition, F is closed under taking factors. Hence there exists a uniquely defined minimal set $J \subseteq K^+$ such that $K^* \setminus F = K^*JK^*$. By Lemma 3.2, we have $J \subseteq K^{\leq 2d}$. In particular, J is finite. Since J and Δ are disjoint, all words in J have a weight greater than κ. Let Ω contain all $\omega \in J$ such that $\omega \in \Gamma K^*$ implies $\omega = \gamma\gamma'$ for some $\gamma, \gamma' \in \Gamma$ with $\gamma > \gamma'$, i.e.,

$$\Omega = J \cap \{\omega \in K^* \mid \omega \notin \Gamma K^* \text{ or } \omega = \gamma\gamma' \text{ for } \gamma, \gamma' \in \Gamma \text{ with } \gamma > \gamma'\}. \tag{5}$$

We have $\Gamma \subseteq \Delta$ and $\Omega \subseteq J$. In particular Ω is finite and every word in Ω has length at most $2d$.

Remark 4.1. We claim $K^*\Gamma K^* \cap J \subseteq K\Gamma \cup \Gamma K$. In particular, for $\omega \in K^*\Gamma \cap \Omega$ we obtain $\omega = b\gamma$ with $b \in K$, $\gamma \in \Gamma$ and $b \neq \gamma$. In order to see the claim, we show that every word in $K^*\Gamma K^* \cap J$ has length 2. Words in J have length at least 2, hence (by left-right symmetry) it is enough to consider words $w = bx\gamma y \in J$ with $b \in K$, $x, y \in K^*$ and $\gamma \in \Gamma$. By minimality of J we obtain $x\gamma y \in F$ and hence $x\gamma y \in \gamma^+$ by Equation (4). Thus, we can write $w = bz\gamma$ with $z \in \gamma^*$ and $bz \in F$. If $z \neq 1$, then $b \in \gamma^+$, too. This implies $w \in \gamma^+$, but this is impossible due to $w \in J$. Therefore $w = b\gamma$ and $b \neq \gamma$.

Let us define a "threshold" value $t \in \mathbb{N}$ by

$$t = \max\{\|\omega v_g \omega\| \mid g \in G,\ \omega \in \Omega\}. \tag{6}$$

This is not the optimal bound at this point, but it allows to use the parameter t again later. For the moment we use only the following two properties, which are satisfied by our choice.

(i) If δ^t is a prefix of a word $u\omega$ or a suffix of a word ωu for $\delta \in \Delta$ and $\omega \in \Omega$, then we have $\|u\| > \max\{\|v_g\| \mid g \in G\}$.
(ii) We have $t > 2d$.

Here $t > 2d$ can be seen by $t > 2\max\{\|\omega\| \mid \omega \in \Omega\} > 2\max\{\|\delta\| \mid \delta \in \Delta\} \geq 2\max\{|\delta| \mid \delta \in \Delta\} = 2d$. The first set of rules over the extended alphabet K deals with

long repetitions of short words: The Δ-rules of the system T are

$$T_\Delta = \left\{ \delta^{t+n} \to \delta^t \;\middle|\; \delta \in \Delta \text{ and } \delta \text{ is primitive} \right\}.$$

LEMMA 4.2. *The system T_Δ is a weighted Church-Rosser system.*

PROOF. Every rule in T_Δ is weight-reducing because primitive words are never empty and $n \geq 1$. Thus, it suffices to show that T_Δ is locally confluent. Let $\delta, \eta \in \Delta$ with $|\delta| \geq |\eta|$ and suppose $\delta^{t+n} \to \delta^t$ and $\eta^{t+n} \to \eta^t$ are rules which are part of a critical pair. We have to study the two cases of factor critical and overlap critical pairs.

We cover factor critical pairs first and thus consider the case that η^{t+n} is a factor of δ^{t+n}. Note that $|\eta^t| \geq t > 2d \geq \left|\delta^2\right|$ by property (ii) above. Thus, there is a conjugate $\zeta = \eta^r \eta_1$ of δ such that η_1 is a proper prefix of η and ζ^2 is a prefix of η^t. By canceling the prefix η^r we see that $\eta_1\eta$ is a prefix of η^2. By primitivity of η this implies that η_1 is empty and by primitivity of δ we obtain $\zeta = \eta$. This implies $|\delta| = |\eta|$ and since η^{t+n} is a factor of δ^{t+n} we obtain that $\eta = \delta$.

The second case are overlap critical pairs. Let $\delta^{t+n} = xy$ and $\eta^{t+n} = yz$ for non-empty words x, y, z. If $|y| > |\eta^t|$, then by $|\eta^t| > \left|\delta^2\right|$ we get that δ^2 is a factor of η^t. Using the same argument as above, we conclude that δ is a conjugate of η and the critical pair resolves. Thus, it remains to prove the case for $|y| \leq |\eta^t| \leq |\delta^t|$. As y is small enough we proceed by writing $x = \delta^n x_1$ and $z = z_1 \eta^n$. The critical pair can be resolved as follows:

$$\begin{aligned} xyz = \delta^{t+n} z &\underset{T_\Delta}{\Longrightarrow} \delta^t z = x_1 \eta^{t+n} \underset{T_\Delta}{\Longrightarrow} x_1 \eta^t = x_1 y z_1, \\ xyz = x\eta^{t+n} &\underset{T_\Delta}{\Longrightarrow} x\eta^t = \delta^{t+n} z_1 \underset{T_\Delta}{\Longrightarrow} \delta^t z_1 = x_1 y z_1. \quad \square \end{aligned}$$

Note that closure under conjugation is not sufficient to guarantee the confluence of T_Δ. We exploited the fact that t is large enough. Example 4.3 shows that at least $t > d - 2$ is necessary.

Example 4.3. Let $\Delta = \{a, aab, aba, baa\}$ with $d = 3$. Consider $t = d - 2 = 1$ and the system $S = \left\{(aab)^2 \to aab, (aba)^2 \to aba, (baa)^2 \to baa, a^2 \to a\right\}$. The set Δ is closed under conjugation and all words in Δ are primitive, but S is not confluent. This can be seen by $abab \underset{S}{\overset{*}{\Longleftarrow}} (aab)^2 \underset{S}{\overset{*}{\Longrightarrow}} ab$.

As we will see next, every sufficiently long word without long Δ-repetitions contains a factor $\omega \in \Omega$:

LEMMA 4.4. *There exists a bound $t_0 \in \mathbb{N}$ such that every word $u \in K^*$ with $\|u\| \geq t_0$ contains either a factor $\omega \in \Omega$ or a factor of the form δ^{t+n} for $\delta \in \Delta$ (or both).*

PROOF. Let us first assume that $u \notin \Gamma^*$. Then there exists a factorization $u = xay$ with $x \in \Gamma^*$ and $a \notin \Gamma$. If $ay \notin F$, there exists a prefix of ay which is in J and consequently also in Ω. Thus, we may assume that $ay \in F$, i.e., ay is a factor of δ^+ for some $\delta \in \Delta$. If $\|ay\| \geq (n + t + 2) \max\{\|\delta\| \mid \delta \in \Delta\}$, then ay contains a factor δ^{t+n}. Thus without loss of generality we may assume that $u = u'u''$ with $u' \in \Gamma^*$ and $\|u''\| \leq (n + t + 2) \max\{\|\delta\| \mid \delta \in \Delta\}$ (This obviously also holds in the case $u \in \Gamma^*$). If u' contains a factor $\gamma\gamma'$ with $\gamma > \gamma'$ we are finished. Thus, $u' = \gamma_{j_1} \cdots \gamma_{j_k}$ with $\gamma_{j_i} \in \Gamma$ and $\gamma_{j_{i-1}} \leq \gamma_{j_i}$. Since u' has no factor $\gamma_{j_i}^{t+n}$ we obtain $k \leq |\Gamma| \cdot (t + n - 1)$. This gives some bound on k and therefore on t_0 as well. $\square$

Remark 4.5. Using $|\Gamma| > s \geq 1$ we can choose the value t_0 of Lemma 4.4 to be $t_0 = (|\Gamma| + 2) \cdot (t + n) \cdot \max\{\|\delta\| \mid \delta \in \Delta\}$.

Words in $\mathrm{IRR}_{T_\Delta}(K^*)$ do not contain any factor of the form δ^{t+n} for $\delta \in \Delta$. Every sufficiently long word v can be written as $v = u_1 \cdots u_k$ with $\|u_i\| \geq t_0$ and k sufficiently large. Thus, by repeatedly applying Lemma 4.4, every long enough word $v \in \mathrm{IRR}_{T_\Delta}(K^*)$ contains two occurrences of the same $\omega \in \Omega$ which are far apart. This suggests rules of the form $\omega\, u\, \omega \to \omega\, v_{\psi(u)}\, \omega$; but in order to ensure confluence we have to limit their use. For this purpose, we equip Ω with a linear order $\preceq$ such that $\gamma_m\gamma_0$ is the least element, and every element in $\Omega \cap K\gamma_0$ is less than all elements in $\Omega \setminus K\gamma_0$.

For a word $v \in K^*\Omega K^*$ define the *maximal Ω-factor* to be the maximal $\omega \in \Omega$ with respect to the linear order $\preceq$ such that $v \in K^*\omega K^*$. The following lemma is the principal reason for excluding all words $\omega \in \Gamma K^*$ in the definition of Ω except for $\omega = \gamma\gamma' \in \Gamma^2$ with $\gamma > \gamma'$.

LEMMA 4.6.

(i) Let $v = x\delta^{t+n}y \in K^*\Omega K^*$. Then $x\delta^t y$ has the same maximal Ω-factor as v.
(ii) Let $v = x\omega u\omega y$ with $\omega \in \Omega$ and $v' = x\omega v_{\psi(u)}\omega y$. Then the maximal Ω-factor of v' is not greater than the maximal Ω-factor of v.

PROOF. (i): By definition of t we have $t > 2d$ and by Lemma 3.2 we have $|\omega| \leq 2d$ for all $\omega \in \Omega$. Thus ω does not contain δ^t as a factor and $x\delta^{t+n}y$ and $x\delta^t y$ have the same Ω-factors. Hence the statement in (i) holds.

(ii): As no Ω-factor can contain ω as a proper factor it suffices to show that ω is the maximal Ω-factor of $\omega v_{\psi(u)}\omega$. The normal form $v_{\psi(u)}$ has $\gamma_m\gamma_0$ as a suffix. In addition, the word $\gamma_m\gamma_0$ is the only element in Ω which is a factor of $v_{\psi(u)}$. The reason is that all other letters in $v_{\psi(u)}$ are in non-decreasing order whereas all $\gamma\gamma' \in \Omega$ are in decreasing order. In particular, if $\gamma_m\gamma_0 v_{\psi(u)}\gamma_m\gamma_0 \in K^*\omega' K^*$ for $\omega' \in \Omega$, then $\omega' = \gamma_m\gamma_0$, i.e., $\gamma_m\gamma_0$ is the only factor of $\gamma_m\gamma_0 v_{\psi(u)}\gamma_m\gamma_0$ which is in Ω.

Let now $\omega \in K^+\gamma_0$. As we have noticed in Remark 4.1, this implies $\omega = b\gamma_0$ with $b \in K \setminus \{\gamma_0\}$ .The set of factors of $\omega v_{\psi(u)}\omega$ which are in Ω is therefore $\{\gamma_m\gamma_0, \omega\}$. Since $\gamma_m\gamma_0 \preceq \omega$ we are done in this case, too.

Next, suppose $\omega \in K^+ b$ for $b \in K \setminus \{\gamma_0\}$. Then the set of factors of $\omega v_{\psi(u)}\omega$ which are in Ω is $\{\gamma_m\gamma_0, b\gamma_0, \omega\}$. Since every element ending with γ_0 is smaller than any other element in Ω, the claim holds in this case, too. □

LEMMA 4.7. *There exists a bound $t_\Omega \in \mathbb{N}$ such that every word $v \in \mathrm{IRR}_{T_\Delta}(K^*)$ with $\|v\| \geq t_\Omega$ contains a factor $\omega u\omega$ for some $\omega \in \Omega$ such that:*

— *$\|v_g\| < \|u\| < t_\Omega$ for all $g \in G$,*
— *ω is the maximal Ω-factor of $\omega u\omega$.*

PROOF. Let $\Omega_v = \{\omega \in \Omega \mid v \in K^*\omega K^*\}$ be the set of factors of v in Ω. For each $k \in \mathbb{N}$ we define a number $t_k = 2^k(t_0 + t) - t$. Thus, for $k \geq 1$ we have

$$t_{k-1} = (t_k - t)/2. \tag{7}$$

Here t_0 and $t = \max\{\|\omega v_g \omega\| \mid g \in G,\ \omega \in \Omega\}$ denote the values which are given by Lemma 4.4 and Equation (6), respectively.

Consider $k \in \mathbb{N}$ and $v \in \mathrm{IRR}_{T_\Delta}(K^*)$ with $k \geq |\Omega_v|$ and assume that the weight of v is at least t_k. By induction on k we show that v contains a factor $\omega u\omega$ with

— $\omega \in \Omega$
— $\|v_g\| < \|u\| < t_k$ for all $g \in G$, and
— ω is the maximal Ω-factor of $\omega u\omega$.

Note that we may replace v by its shortest prefix which has weight at least t_k, because if we find a factor $\omega u \omega$ of the desired form in this prefix, then we are done. Hence we may assume that every proper factor of v has weight less than t_k.

The base $k = 0$ is trivial since such an irreducible word v with $\|v\| \geq t_0$ and a factor in Ω cannot exist by Lemma 4.4. Thus, we may assume that $k \geq 1$. Let $\mu \in \Omega$ be the maximal Ω-factor of v and consider a factorization $v = pfq$ with $f \in \mu K^* \cap K^* \mu$ and p, q have no factor μ, i.e., f is the shortest factor of v which contains every occurence of μ in v. If $\|f\| > t = \max\{\|\omega v_g \omega\| \mid g \in G,\ \omega \in \Omega\}$, then we have found a factor $f = \mu u \mu$. Since u is a proper factor of f and hence of v, we obtain $\|u\| < t_k$. We also have $\|f\| = \|\mu u \mu\| > t = \max\{\|\omega v_g \omega\| \mid g \in G,\ \omega \in \Omega\}$ which implies $\|u\| > \max\{\|v_g\| \mid g \in G\}$. Combining these results shows $\max\{\|v_g\| \mid g \in G\} < \|u\| < t_k$, and thus f is a factor with the required properties.

Therefore we may assume $\|f\| \leq t$. By symmetry let $\|p\| \geq \|q\|$. In particular, $\|p\| \geq (t_k - t)/2$. Hence, by Equation (7) we obtain $\|p\| \geq t_{k-1}$. Since p has at least one factor of Ω less than v, we conclude by induction that p contains a factor $\omega u \omega$ with ω the maximal Ω-factor and $\|v_g\| < \|u\| < t_{k-1} \leq t_k$ for all $g \in G$. Thus the assertion holds for k.

Choose

$$t_\Omega = 2^{|\Omega|}(t_0 + t) = t_{|\Omega|} + t \tag{8}$$

and consider a word $v \in \mathrm{IRR}_{T_\Delta}(K^*)$ with $\|v\| \geq t_\Omega$. By the definition of t_Ω there is a prefix v' of v with $t_{|\Omega|} \leq \|v'\| < t_\Omega$. By the induction above we conclude that v', and thus also v, contains a factor $\omega u \omega$ with the desired properties. □

We are now ready to define the second set of rules over the extended alphabet K. These rules reduce long words without long repetitions of words in Δ. We denote

$$T'_\Omega = \left\{ \omega\, u\, \omega \to \omega\, v_{\psi(u)}\, \omega \;\middle|\; \begin{array}{l} \|v_{\psi(u)}\| < \|u\| < t_\Omega,\ \omega \in \Omega \text{ and} \\ \omega \text{ is the maximal } \Omega\text{-factor of } \omega u \omega \end{array} \right\}.$$

Whenever there is a shorter rule in $T'_\Omega \cup T_\Delta$ then we want to give preference to this shorter rule. Thus, the Ω-rules are

$$T_\Omega = \left\{ \ell \to r \in T'_\Omega \;\middle|\; \begin{array}{l} \text{there is no rule } \ell' \to r' \in T'_\Omega \cup T_\Delta \\ \text{such that } \ell' \text{ is a proper factor of } \ell \end{array} \right\}.$$

Let $T = T_\Delta \cup T_\Omega$. The set $\mathrm{IRR}_T(K^*)$ is finite by Lemma 4.7. Our goal is to prove confluence of T over K^*. As an auxiliary result we prove the following lemma, which is of independent interest.

LEMMA 4.8. *Let $\omega \in \Omega$ and $v = \omega\gamma u\omega$ (resp. $v = \omega u \gamma \omega$) be a word with $\gamma \in \Gamma$ and with $\|\gamma u\| > \max\{\|v_g\| \mid g \in G\}$ such that ω is the maximal Ω-factor of v. Then there exists a derivation $v \overset{*}{\underset{T}{\Longrightarrow}} \omega v_{\psi(\gamma u)} \omega$ (resp. $v \overset{*}{\underset{T}{\Longrightarrow}} \omega v_{\psi(u\gamma)} \omega$).*

PROOF. In order to show this, we will first prove three auxiliary claims. It suffices to consider the case $v = \omega\gamma u\omega$ since $v = \omega u \gamma \omega$ is symmetric.

CLAIM 1. *The word v is reducible in T.*

If v is reducible in T_Δ we are finished. Thus assume that v is irreducible in T_Δ. Then either v is the left side of a rule in T'_Ω or $\|v\| > t_\Omega$. If v is the left side of a rule in T'_Ω, then either v is the left side of a rule in T_Ω or it contains a factor which is the left side of such a rule. If $\|v\| > t_\Omega$, then v contains a factor which is a rule in T_Ω by Lemma 4.7. This concludes Claim 1.

CLAIM 2. *If $\omega\gamma u\omega \underset{T}{\Longrightarrow} v'$, then $v' = \omega\gamma' u' \omega$ where $\gamma' = \gamma$ or $\gamma' = \gamma_0$ and $u' \in K^*$.*

There are three cases. The first case is that v' stems from a rule $\delta^{t+n} \to \delta^t \in T_\Delta$ and γ is contained in δ^{t+n}. Note that by $|\omega| \leq 2d < t$ the left side δ^{t+n} cannot be contained in ω. We have $\delta = \gamma$ by Equation (4). By Remark 4.1 the overlap of δ^{t+n} and ω is at most γ. As $t > 2d \geq 2$ this overlap and the γ are preserved and the claim is clear.

The second case is that v' stems from a rule $\delta^{t+n} \to \delta^t \in T_\Delta$ and γ is not contained in δ^{t+n}. Again we have that δ^{t+n} cannot be contained in ω. Also δ^{t+n} can at most overlap at the right ω. The overlap with ω is still in δ^t as $t > 2d \geq |\omega|$ and therefore the claim holds in this case. Thus, in the first and second case we have $\gamma' = \gamma$.

In the third case v' stems from a rule $\ell = \omega' u' \omega' \to \omega' v_{\psi(u')} \omega' = r$ in T_Ω. If ℓ is a prefix of v, then $v' = \omega\gamma_0 u'\omega$ and $v_{\psi(u')}$ is a prefix of $\gamma_0 u'$. Hence the claim holds in this case. If the factor γ (in $v = \omega\gamma u\omega$) is not a factor of ℓ, then the claim is trivial. Hence let γ be a factor of ℓ. Then γ is a factor of ω' by minimality of J. As ω is preserved at the use of the rule $\ell \to r$, the claim holds in this case too. Therefore, $v' = \omega\gamma' u'\omega$ for $\gamma' = \gamma$ or $\gamma' = \gamma_0$. This concludes Claim 2.

Note that if v is the left side of a rule, the statement of the lemma is clear. Thus we have to study the case that v is not a left side of a rule.

CLAIM 3. $v \underset{T}{\Longrightarrow} v' = \omega\gamma' u'\omega \neq \omega v_{\psi(\gamma u)}\omega$ *implies* $\|\gamma' u'\| > \max\{\|v_g\| \mid g \in G\}$.

We therefore may assume that v is reduced to v' by some rule $\ell \to r \in T$ with $\ell \neq v$. We again use case-by-case analysis for T_Δ and T_Ω rules.

The first case is that $\ell \to r \in T_\Delta$. By definition of T_Δ we have $|r| \geq t$ and thus by $\ell \neq v$ this implies $\|v'\| \geq |v'| = |\omega\gamma' u'\omega| > t = \max\{\|\omega v_g \omega\| \mid g \in G,\ \omega \in \Omega\}$. By cancelation of ω this implies $\|\gamma' u'\| > \max\{\|v_g\| \mid g \in G\}$.

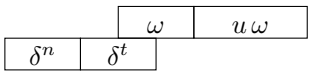
The second case is that $\ell \to r \in T_\Omega$. Thus, we have $\ell = \omega' u'' \omega'$. If the rule does not apply to a prefix, then u' must contain some factor v_g and we obtain $\|u'\| \geq \|v_g\|$ for some $g \in G$. This is large enough since $\|\gamma' u'\| \geq \|v_g\| + \kappa > \max\{\|v_g\| \mid g \in G\}$. The remaining case is that the rule $\ell \to r \in T$ applies to a prefix of v. But then we must have $\omega = \omega'$. Thus, $v = \omega u''\omega x$ with $\omega x = x'\omega$ where $x' \neq 1$. This implies $\|x'\| > \kappa$ since otherwise ω would be a factor of x'^+. This is large enough since $v' = \omega v_{\psi(u'')} x'\omega$ in this case. This concludes Claim 3.

Using these claims we proceed using Noetherian induction on the weight of γu. By Claim 1 the word v is reducible. Thus let $v \underset{T}{\Longrightarrow} v'$. By Claim 2 we obtain $v' = \omega\gamma' u'\omega$ for some $\gamma \in \Gamma$. If $\gamma' u' \neq v_{\psi(\gamma u)}$ we obtain $\|\gamma' u'\| > \max\{\|v_g\| \mid g \in G\}$ by Claim 3. As the weight of $\gamma' u'$ is smaller than the weight of γu, we have $\psi(\gamma' u') = \psi(\gamma u)$ by construction of the rules and v' satisfies the requirements of the lemma by Claim 2 and Claim 3, we can use induction. This process stops as soon as $\gamma' u' = v_{\psi(\gamma u)}$ which concludes the proof. □

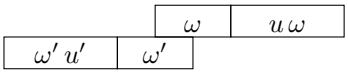
For showing the following lemma we reuse some arguments from the proof of Lemma 4.8. We think however that this overlap may improve the clarity of presentation.

LEMMA 4.9. *The system T is locally confluent over K^*.*

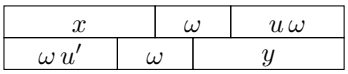
PROOF. The system T_Δ is confluent by Lemma 4.2. Suppose we can apply the rules $\ell \to r \in T_\Omega$ and $\ell' \to r' \in T_\Delta$. Then ℓ' is not a proper factor of ℓ by definition of T_Ω. Moreover no ω is a factor of any δ^+, hence ℓ is not a factor of ℓ'. Thus, there are no factor critical pairs in this case. Next we consider overlap critical pairs. Let $\ell = \omega u\omega$ and $\ell' = \delta^{t+n}$. The maximal overlap between ℓ and ℓ' is a prefix or suffix of ω. By the choice of t we have $t \geq 2d$, hence neither the application of $\ell \to r$ nor the application of $\ell' \to r'$ changes any overlap. Therefore we can apply the rules in any order and we obtain the same result:



It remains to show that T_Ω is locally confluent. By minimality of J, no $\omega \in \Omega$ is a proper factor of another word $\omega' \in \Omega$. Let $\omega u \omega \to r$ and $\omega' u' \omega' \to r'$ be two Ω-rules and first assume $\omega \neq \omega'$. By construction of T'_Ω, the left sides of both rules can overlap at most $\min\{|\omega|, |\omega'|\} - 1$ positions. Thus, the two rules can always be applied independently of one another.



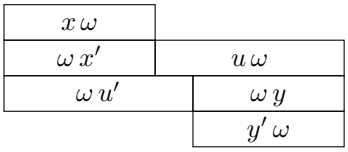
Let finally $\omega u \omega \to \omega v_g \omega$ and $\omega u' \omega \to \omega v_h \omega$ be two Ω-rules. By construction of T_Ω, the left-hand side $\omega u' \omega$ is neither a proper factor of $\omega u \omega$ nor vice versa. Suppose that these two rules are applied to $x\omega u\omega = \omega u' \omega y = \omega u'' \omega$ for $x, y \in K^+$. If $|x| \geq |\omega u'|$, then the two rules can be applied independently of one another.



Thus, we may assume $|x| < |\omega u'|$. We will show

$$x\,\omega\,v_g\,\omega \underset{T}{\overset{*}{\Longrightarrow}} \omega\,v_{\psi(u'')}\,\omega \underset{T}{\overset{*}{\Longleftarrow}} \omega\,v_h\,\omega\,y. \tag{9}$$

Let $x\omega = \omega x'$ for some $x' \in K^+$. If $\|x\| \leq \kappa$, then x is a prefix of ω since $\|\omega\| > \kappa$ and ω becomes a prefix of x^+. Due to $\|x\| \leq \kappa$ we have $x \in \Delta$, hence $\omega \in F$. This is a contradiction since $\Omega \subseteq J$. We obtain $\|x\| > \kappa$. Analogously, we also have $\|y\| > \kappa$ and $\omega y = y'\omega$ for some $y' \in K^+$.



Because of $|x| < |\omega u'|$, the definition of T_Ω, and the fact that different words in Ω are not factors of one another, we have that ω is the maximal Ω-factor in $x\omega u\omega = \omega u'\omega y$. Hence, ω is still the maximal Ω-factor in $x\omega v_g\omega$ and in $\omega v_h \omega y$ by Lemma 4.6. Moreover, since $\|x'\| = \|x\| > \kappa$, we have $\|x' v_g\| > \kappa + \|v_g\| > \max\{\|v_g\| \mid g \in G\}$. The last letter of $x' v_g$ is in Γ since v_g ends in γ_0. Thus, the requirements of Lemma 4.8 are satisfied and we obtain $x\,\omega\,v_g\,\omega = \omega x' v_g \omega \underset{T}{\overset{*}{\Longrightarrow}} \omega\,v_{\psi(x'v_g)}\,\omega$. Similarly, $\omega\,v_h\,\omega y = \omega\,v_h\,y'\omega \underset{T}{\overset{*}{\Longrightarrow}} \omega\,v_{\psi(v_h y')}\,\omega$. Finally, $\psi(x'v_g) = \psi(v_h y') = \psi(u'')$ which shows Equation (9). □

Since all rules in T are weight-reducing, it follows from Lemma 4.9 that T is confluent. Moreover, all rules $\ell \to r$ in T satisfy $\psi(\ell) = \psi(r)$. We conclude that T is a weighted Church-Rosser system such that K^*/T is finite and $\psi : K^* \to G$ factorizes through T. This finishes the proof of Proposition 3.4.

5. ARBITRARY FINITE MONOIDS

This section contains the main result of this paper. We show that every homomorphism $\varphi : A^* \to M$ to a finite monoid M factorizes through a weighted Church-Rosser system S of finite index. The proof relies on Theorem 3.3 and on a construction called local divisors.

5.1. Local divisors

Local divisors are a powerful tool when using inductive proofs for finite monoids. In finite semigroup theory it was used first in [Diekert and Gastin 2006]; in associative algebra the concept goes back to the notion of *local algebra* introduced by Meyberg in the technical report [Meyberg 1972]. The definition of a local divisor is as follows: Let M be a monoid and let $c \in M$. We equip $cM \cap Mc$ with a monoid structure by introducing a new multiplication $\circ$ as follows:

$$xc \circ cy = xcy.$$

It is straightforward to see that $\circ$ is well-defined and $(cM \cap Mc, \circ)$ is a monoid with neutral element c. The following observation is crucial. If $1 \in cM \cap Mc$, then c is a unit. Thus if the monoid M is finite and c is not a unit, then $|cM \cap Mc| < |M|$. The set $M' = \{x \mid cx \in Mc\}$ is a submonoid of M, and $c\cdot : M' \to cM \cap Mc$ with $x \mapsto cx$ is a surjective homomorphism. Since $(cM \cap Mc, \circ)$ is the homomorphic image of a submonoid, it is a divisor of M. We therefore call $(cM \cap Mc, \circ)$ the *local divisor* of M at c.

5.2. The main result

We are now ready to prove our main result. Let $(A, \|\cdot\|)$ be a weighted alphabet. Then every homomorphism $\varphi : A^* \to M$ to a finite monoid M factorizes through a weighted Church-Rosser system S of finite index. The proof uses induction on the size of M and the size of A. If $\varphi(A^*)$ is a group, then we apply Theorem 3.3; and if $\varphi(A^*)$ is not a group, then we find a letter $c \in A$ which is not a unit. Thus in this case we can use local divisors.

THEOREM 5.1. *Let $(A, \|\cdot\|)$ be a weighted alphabet and let $\varphi : A^* \to M$ be a homomorphism to a finite monoid M. Then there exists a weighted Church-Rosser system S of finite index such that φ factorizes through S.*

PROOF. The proof is by induction on $(|M|, |A|)$ with lexicographic order. This means that $(|M'|, |A'|)$ is less than $(|M|, |A|)$ if either $|M'| < |M|$ or both $|M'| = |M|$ and $|A'| < |A|$. If $\varphi(A^*)$ is a group, then the claim follows by Theorem 3.3. If $\varphi(A^*)$ is not a group, then there exists $c \in A$ such that $\varphi(c)$ is not a unit. Let $B = A \setminus \{c\}$. By induction on the size of the alphabet there exists a weighted Church-Rosser system R for the restriction $\varphi : B^* \to M$ satisfying the statement of the theorem. Let

$$K = \mathrm{IRR}_R(B^*)c.$$

We consider the prefix code K as a weighted alphabet. The weight of a letter $uc \in K$ is the weight $\|uc\|$ when read as a word over the weighted alphabet $(A, \|\cdot\|)$. Let $M_c = \varphi(c)M \cap M\varphi(c)$ be the local divisor of M at $\varphi(c)$. We let $\psi : K^* \to M_c$ be the homomorphism induced by $\psi(uc) = \varphi(cuc)$ for $uc \in K$. By induction on the size of the monoid there exists a weighted Church-Rosser system $T \subseteq K^* \times K^*$ for ψ satisfying the statement of the theorem. Suppose $\psi(\ell) = \psi(r)$ for $\ell, r \in K^*$ and let $\ell = u_1c \cdots u_jc$ and $r = v_1c \cdots v_kc$ with $u_i, v_i \in \mathrm{IRR}_R(B^*)$. Then

$$\begin{aligned}\varphi(c\ell) &= \varphi(cu_1c) \circ \cdots \circ \varphi(cu_jc)\\ &= \psi(u_1c) \circ \cdots \circ \psi(u_jc)\\ &= \psi(\ell) = \psi(r) = \varphi(cr).\end{aligned}$$

This means that every T-rule $\ell \to r$ yields a φ-invariant rule $c\ell \to cr$. We can transform the system $T \subseteq K^* \times K^*$ for ψ into a system $T' \subseteq A^* \times A^*$ for φ by

$$T' = \{c\ell \to cr \in A^* \times A^* \mid \ell \to r \in T\}.$$

The system T' is obviously weight-reducing over A^*. Let us show that T' is locally confluent. Consider any derivation $u \xRightarrow[T']{*} \widehat{u}$ such that $\widehat{u} \in \mathrm{IRR}_{T'}(A^*)$. We have to show that $\widehat{u}$ is uniquely defined by these conditions. It suffices to check this for critical pairs. We either have $u = \ell y = x\ell'$ in the case of an overlap critical pair or $u = \ell = x\ell' y$ in the case of a factor critical pair for $x, y \in A^*$ and ℓ, ℓ' left sides of rules in T'. Note that the rules in T' are in K^*. In particular, this implies $u = cv \in cK^*$ and $cv \xRightarrow[T']{*} c\widehat{v}$ such that $\widehat{u} = c\widehat{v} \in K^*$. Using factorizations as words over the free monoid K^* we see that we have $v \xRightarrow[T]{*} \widehat{v}$. Moreover, since $\widehat{u} = c\widehat{v} \in \mathrm{IRR}_{T'}(A^*)$ we have $\widehat{v} \in \mathrm{IRR}_T(K^*)$. Since T is confluent, $\widehat{v}$ is uniquely defined and so is $c\widehat{v}$. Thus, T' is confluent and weight-reducing over A^*. Combining R and T' leads to $S = R \cup T'$. The left sides of a rule in R and a rule in T' cannot overlap. Therefore, S is a weighted Church-Rosser system such that φ factorizes through A^*/S. Suppose that every word in $\mathrm{IRR}_T(K^*)$ has length at most k. Here, the length is over the extended alphabet K. Similarly, let every word in $\mathrm{IRR}_R(B^*)$ have length at most m. Then

$$\mathrm{IRR}_S(A^*) \subseteq \{u_0 c u_1 \cdots c u_{k'+1} \mid u_i \in \mathrm{IRR}_R(B^*),\ k' \leq k\}$$

and every word in $\mathrm{IRR}_S(A^*)$ has length at most $(k+2)(m+1)$. In particular $\mathrm{IRR}_S(A^*)$ and A^*/S are finite. □

The following corollary is a straightforward translation of the result in Theorem 5.1.

COROLLARY 5.2. *A language $L \subseteq A^*$ is regular if and only if there exists a Church-Rosser system S of finite index such that $L = \bigcup_{u \in L} [u]_S$. In particular, all regular languages are strongly Church-Rosser congruential.*

PROOF. If L is regular, then there exists a homomorphism $\varphi : A^* \to M$ recognizing L. By Theorem 5.1 there exists a Church-Rosser system S of finite index such that φ factorizes through S. The latter property implies $\varphi^{-1}(x) = \bigcup_{u \in \varphi^{-1}(x)} [u]_S$ for every $x \in M$. Thus $L = \bigcup_{x \in \varphi(L)} \varphi^{-1}(x) = \bigcup_{u \in L} [u]_S$. The converse is clear because L is recognized by the finite monoid A^*/S. □

6. CONCLUSION AND OPEN PROBLEMS

We have shown that all regular languages are Church-Rosser congruential. The proof was achieved by loading the induction hypothesis. Our result says that for all $\varphi : A^* \to M$ to a finite monoid M and all weights $\|\cdot\| : A \to \mathbb{N} \setminus \{0\}$ there exists a weighted Church-Rosser system S of finite index such that φ factorizes through S. An interesting question is whether we can change quantifiers. Is it true that for all such φ there exists a finite confluent system S of finite index such that φ factorizes through S and which is weight-reducing for all weights? Note that whether a system is weight-reducing for all weights is a natural condition on the number of letters in the Parikh image. Thus, we can call such a system *Parikh-reducing*. This result is true for aperiodic monoids [Diekert et al. 2012b], because every subword-reducing system is Parikh-reducing.

Another problem for future research is which algebraic invariants of M can be maintained in A^*/S. For example, if M satisfies the equation $x^{t+p} = x^t$, then our construction yields that A^*/S satisfies an equation $x^{s+p} = x^s$ for some large enough s. We conjecture that we must choose $s > t$ in general. In particular, we doubt that we can choose A^*/S to be a group, even if M is a (cyclic) finite group. However proving such a *lower bound result* seems to be a hard task.

If a language is a finite union of congruence classes w.r.t. some finite confluent and weight-reducing system, then it has essentially the same nice properties as a Church-Rosser congruential language. Considering weights instead of lengths does not in-

crease the expressive power of the class of Church-Rosser languages, see [Niemann and Otto 2005]. It is however not clear from that result that the same is true for Church-Rosser congruential languages.

Finally, in our construction, the size of the monoid A^*/S is huge compared to M and A. Obvious open problems are lower bounds on the size of the system S and A^*/S, as well as reasonable upper bounds.

ACKNOWLEDGMENT

We thank the anonymous referees for various helpful suggestions which significantly improved the presentation.

References

BOOK, R. AND OTTO, F. 1993. *String-Rewriting Systems*. Springer-Verlag.

BUNTROCK, G. AND OTTO, F. 1998. Growing context-sensitive languages and Church-Rosser languages. *Information and Computation 141*, 1–36.

DIEKERT, V. AND GASTIN, P. 2006. Pure future local temporal logics are expressively complete for Mazurkiewicz traces. *Information and Computation 204*, 1597–1619. Conference version in LATIN 2004, LNCS 2976, 170–182, 2004.

DIEKERT, V., KUFLEITNER, M., REINHARDT, K., AND WALTER, T. 2012a. Regular languages are Church-Rosser congruential. In *International Colloquium Automata, Languages and Programming (ICALP) 2012, Conference Proceedings, Part II*, A. Czumaj, K. Mehlhorn, A. Pitts, and R. Wattenhofer, Eds. Lecture Notes in Computer Science Series, vol. 7392. Springer-Verlag, 177–188.

DIEKERT, V., KUFLEITNER, M., AND WEIL, P. 2012b. Star-free languages are Church-Rosser congruential. *Theoretical Computer Science 454*, 129–135.

JANTZEN, M. 1988. *Confluent String Rewriting*. EATCS Monographs on Theoretical Computer Science Series, vol. 14. Springer-Verlag.

LOTHAIRE, M. 1983. *Combinatorics on Words*. Encyclopedia of Mathematics and its Applications Series, vol. 17. Addison-Wesley, Reading, MA. Reprinted by *Cambridge University Press*, 1997.

MCNAUGHTON, R., NARENDRAN, P., AND OTTO, F. 1988. Church-Rosser Thue systems and formal languages. *J. ACM 35,* 2, 324–344.

MEYBERG, K. 1972. Lectures on algebras and triple systems. Tech. rep., University of Virginia, Charlottesville.

NARENDRAN, P. 1984. Church-Rosser and related Thue systems. Ph.D. thesis, Dept. of Mathematical Sciences, Rensselaer Polytechnic Institute, Troy, NY, USA.

NIEMANN, G. 2002. *Church-Rosser Languages and Related Classes*. Kassel University Press. PhD thesis.

NIEMANN, G. AND OTTO, F. 2005. The Church-Rosser languages are the deterministic variants of the growing context-sensitive languages. *Inf. Comput. 197*, 1–21.

NIEMANN, G. AND WALDMANN, J. 2002. Some regular languages that are Church-Rosser congruential. In *DLT'01, Proceedings*. LNCS Series, vol. 2295. Springer, 330–339.

NIVAT, M. 1970. On some families of languages related to the Dyck language. In *Proceedings of the second annual ACM symposium on Theory of computing*. STOC '70. ACM, New York, NY, USA, 221–225.

REINHARDT, K. AND THÉRIEN, D. 2003. Some more regular languages that are Church Rosser congruential. In *13. Theorietag, Automaten und Formale Sprachen, Herrsching, Germany*. 97–103.

WOINOWSKI, J. R. 2001. Church-Rosser languages and their application to parsing problems. Ph.D. thesis, TU Darmstadt.

WOINOWSKI, J. R. 2003. The context-splittable normal form for Church-Rosser language systems. *Inf. Comput. 183*, 245–274.