MACT 4134 FINAL PROJECT
FALL 2018, AUC

ALGEBRAIC THEORY OF AUTOMATA AND SEMIGROUPS

HASHEM ELEZABI

Abstract. In 1965, Kenneth Krohn and John Rhodes proved that every finite automaton can be decomposed into a cascade of very simple automata consisting of two types, reset automata and permutation automata. They proved this using algebraic techniques, marking the first major study of finite automata via modern algebra and earning two simultaneous PhDs for their work. The main connection between automata and algebra stems from the fact that every finite automaton has a corresponding *transformation semigroup*. In fact, the Krohn-Rhodes theorem above can be restated in purely semigroup-theoretic terms: every finite semigroup divides a wreath product of finite simple groups and finite aperiodic semigroups. In this way, the theorem proves a significant statement about semigroups themselves, and reveals a deep connection between finite automata and finite semigroups. In this paper, we briefly introduce finite automata theory, explain the correspondence between automata and semigroups, and prove the Krohn-Rhodes theorem in the language of semigroup theory.

1. Introduction and Motivation

The Krohn-Rhodes theorem started the field sometimes called algebraic automata theory. Finite automata theory is a subfield of computer science dealing with finite automata. A *finite automaton* is a simple abstract model of a computing device. It is one way to mathematically model computation. It is a finite-state machine that accepts and rejects strings of symbols based on a set of rules, called *transitions*. Here, we will only study *determinstic* finite automata, or DFAs, in which there is a unique computation for every input string. A DFA has an input state, where the machine starts its computation, and a set of accept states. If an automaton A stops at one of its accept states after reading an input string w, we say that A accepts w. The typical way to depict finite automata is using a directed graph with labeled vertices and edges. Figure 1.1 [7] shows an automaton that accepts any string w in which the number of occurrences of b is congruent to 2 mod 3. In this example, the input strings can consist only of a's and b's. In the figure, each circle is a state, and a circle with another circle inside it is an accept state. The directed edges define the transitions. The computation works as follows. We start at the initial state, or start state, as indicated in the figure, and read the input string one symbol at a time. Once we read a symbol, we go to the state indicated by the relevant transition, and move on to the next symbol in the string. In this example, if we are in state q_1 and read the symbol a, we stay in q_1, and if we read b we go to state q_2, which is an accept state. Another example is in Figure 1.2 [7]. Here, the automaton accepts any string w that contains aa as a substring.

Figure 1.1

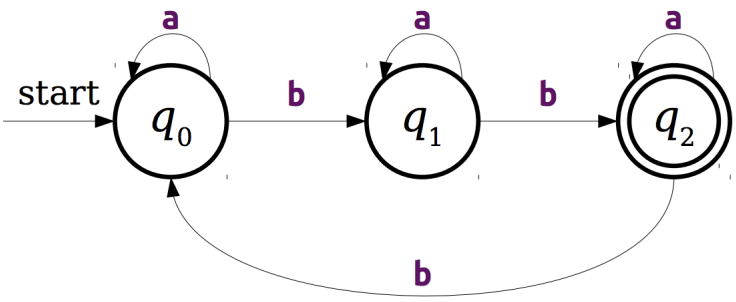
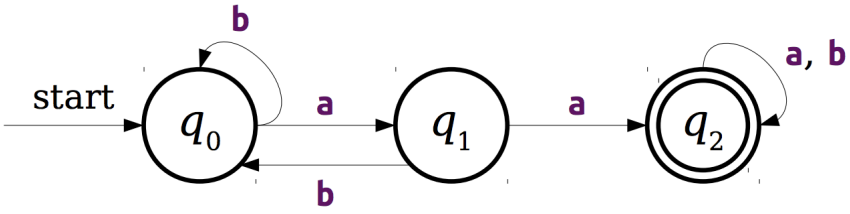


Figure 1.2



We now give some formal definitions. An *alphabet* Σ is a finite, nonempty set of symbols, which we also refer to as letters. A *string*, or *word*, over an alphabet is a finite sequence of letters drawn from Σ. We allow an empty string and denote it by ϵ. Formally, a deterministic finite automaton is a 5-tuple $(Q, \Sigma, \delta, q_0, F)$ consisting of a finite set of states Q; a finite set of input symbols, the alphabet Σ; a transition function $\delta : Q \times \Sigma \to Q$; a start state $q_0 \in Q$; and a set of accept states $F \subseteq Q$. Note that with this definition, every combination of states and symbols has a corresponding transition. An automaton with this property is called *complete*, so according to our definition all DFAs are complete. There is a clear correspondence between the formal definition of a DFA and the directed graph representations shown above. In both examples we have $\Sigma = \{a, b\}$.

Before going further, we take some time to comment on why we study automata. First, we discuss the motivation from the perspective of computer science. Figure 1.3 shows an architecture of a real computer [6]. While this may be a familiar diagram for a computer scientist studying computer architecture and computer systems, it looks quite messy for a theoretical computer scientist, and there is no way to mathematically study such systems. This is where we feel the need for an *abstraction*. Finite automata are simple abstractions that can nevertheless be made into powerful tools which can be used to study what problems we can solve with computers. For example, we can prove that some problems are *undecidable*, which means that it is impossible to construct an algorithm that would yield a yes-no answer to every instance of that problem. Such a statement is quite powerful and holds regardless of technological advancements. These purely theoretical results can then be used to guide concrete applications and decisions. In addition, finite automata have some other direct applications in computer science. For example, Figure 1.4 shows a more elaborate finite automaton that determines whether or not the given string is a valid

Figure 1.3: Example architecture of a computer.

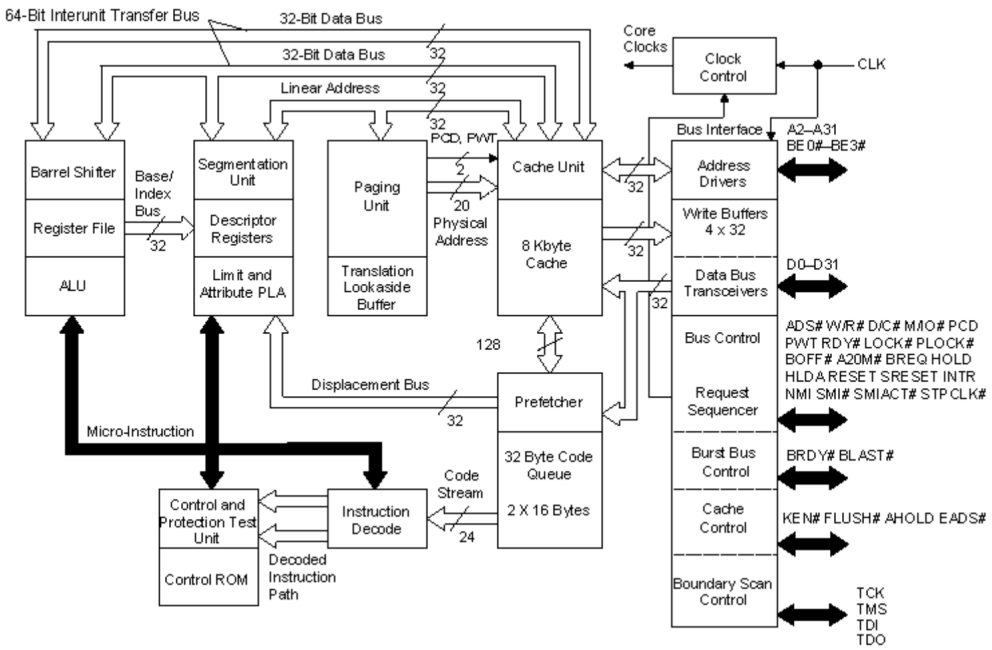
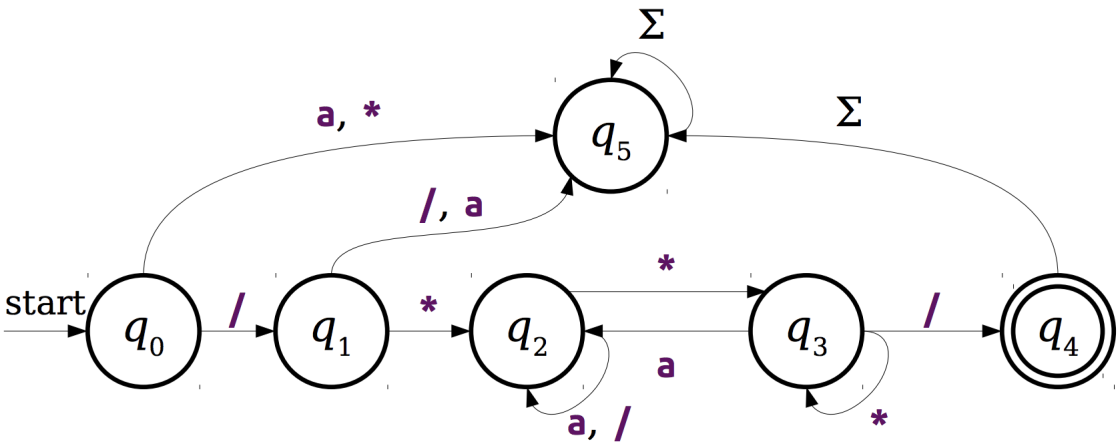


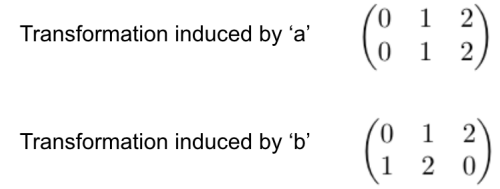
Figure 1.4: A DFA that only accepts valid C-style comments.



C-style comment (the letter 'a' represents any character other than '/' and '*'). In fact, finite automata are used in software compilers for similar purposes.

But there is a more general motivation for studying finite automata, beyond applications to computer science and software engineering. These finite-state machines are able to capture the structure and interactions of various real-world systems and how they act in response to environmental and internal changes. A key notion here is the notion of *change* of a system, which is a fundamental concept in science and computation [11]. We can observe an interesting parallel here between mathematical analysis and automata theory [11]. In the former, we study continuous functions, in which our set of states is a continuum. In the latter, we study automata, where we

Figure 1.5



have a discrete set of states. Some applications fit naturally in calculus, while others are more suitable for automata. In light of this observation, the study of automata has applications that greatly transcend theoretical computer science, and while they are already used in various contexts, we might potentially find new groundbreaking applications of this theory in the future.

What Krohn and Rhodes showed in their 1965 paper is that we can describe and study automata algebraically. This is an exciting application of the wide array of techniques and results in modern algebra that have been developed in the 19th and 20th centuries and continue to expand today. This connection comes from the correspondence between automata and semigroups. A *semigroup* is a set of elements along with an associative binary operation. No other conditions exist for a semigroup, so it need not have an identity or inverses. We will discuss semigroups and their properties in more detail in the next section, but here we mention just enough to establish the connection between semigroups and automata. A *transformation* is a function $s : Q \to Q$ over a finite set Q of states. By a function, we mean a total function, so s is defined for every state $q \in Q$. Thus, a transformation simply takes an input state q and sends it to an output state q'. Notice that *permutations* form a subset of transformations. Specifically, permutations are bijective transformations. Now define the *transformation semigroup* (Q, S) as a pair consisting of a set of states Q and a semigroup of transformations S closed under the associative operation of function composition. Transformation semigroups are the semigroup analogues of permutation groups; we will revisit them in the next section. If we ignore the start state and the accept states of a finite automaton (eg. if we define a finite automaton as a 3-tuple (Q, Σ, δ)), we can see that every automaton corresponds to some transformation semigroup whose generators are the transformations induced by the input letters in Σ. Each letter would have a corresponding transformation that sends each state to the state defined by the relevant transition. Figure 1.5 shows the two transformations corresponding to the two input letters in the first example of an automaton in Figure 1.1. The computation of the automaton on any word w is determined by the composition of the transformations of the letters of w in sequence. This shows that transformation semigroups are another way to capture the notion of change in a discrete way. A transformation semigroup is essentially the same concept as a finite automaton in the sense of the correspondence above. But since semigroups are algebraic objects living in the rich world of abstract algebra, we can study them through that lens and ultimately gain new insights into finite automata, as we will show in the next sections.

We further motivate the Krohn-Rhodes decomposition theorem with a final note. The idea of a *decomposition* is a fundamental technique used in numerous areas of science and engineering. It involves identifying the building blocks of a system and how they work together to build the system, where the simpler building blocks are easier to understand. There is in fact an analogue of the Krohn-Rhodes theorem in group theory, the Jordan-Hölder theorem. In one sense, the Krohn-Rhodes theorem generalizes the Jordan-Hölder theorem by showing that a general decomposition exists for semigroups like it exists for groups. Another example of an important decomposition theorem is the fundamental theorem of finite abelian groups, which is a powerful theorem that enables us to describe all finite abelian groups. The general ability to decompose a complicated structure into simpler constituent parts is very helpful. Decompositions based on the Krohn-Rhodes theorem have been computationally implemented [10], and we might find some fascinating uses of these decomposition techniques in the future. Many possible applications have been outlined in [9].

2. Semigroups

In this section, we introduce semigroups and monoids and discuss some basic properties and constructs. We mostly focus on only defining concepts that we will need when proving the Krohn-Rhodes theorem in the next section, while including a few proofs. Almost all of the definitions are from [8].

A *semigroup* S is a non-empty set along with an associative binary operation (under which S is closed, by definition). To formally denote a semigroup, we write $(S, \cdot)$ where S is the underlying set and $\cdot$ is the associative binary operation. We will denote the *multiplication* of two elements $a, b \in$S, which is the result of applying the operation on them, by ab, unless denoting the operation is needed for clarity. An *identity* is an element $e \in$ such that $ae = ea = a$ for all $a \in S$. A *monoid* is a semigroup with an identity. Note that we can always form a monoid from a semigroup S by simply adjoining an identity element and defining its multiplication with all other elements appropriately. We denote this newly formed monoid by S^1, following the notation in [8].

Example 2.1. $\mathbf{Z}^+$, the set of positive integers, is a semigroup under addition. The set of natural numbers $\mathbf{N} = \mathbf{Z}^+ \cup \{0\}$ is a monoid under addition, with identity 0. (In this paper, the natural numbers $\mathbf{N}$ are defined to include 0.)

Example 2.2. The set of all non-empty strings over a finite alphabet Σ is a semigroup under the associative operation of string concatenation. It is called the *free semigroup over* Σ, and is usually denoted by Σ^+. If we include the empty string, we get the *free monoid over* Σ, usually denoted by Σ^*.

Example 2.3. Any ring is a semigroup under multiplication. A ring with unity is a monoid under multiplication. (Multiplication here refers to the second operation defined for every ring. It is different from the general usage of 'multiplication' used throughout this paper as the operation on a semigroup.)

An *inverse* of an element a in a monoid M with identity e is an element a^{-1} such that $aa^{-1} = a^{-1}a = e$. If x has an inverse, we say that x is *invertible*. A monoid in which every element is invertible is a group, as we know. Clearly, we have that $Groups \subset Monoids \subset Semigroups$.

A *subsemigroup* T of a semigroup S is a non-empty subset of S such that for any $a, b \in T$, $ab \in T$. In words, it is closed under multiplication. A *proper* subsemigroup is any subsemigroup except S itself. A *submonoid* is a subsemigroup that is a monoid. A *subgroup* is a subsemigroup that is a group.

Theorem 2.4. *Let T be the set of invertible elements of a monoid M. Then T is a subgroup of M.*

Proof. Note that T is not empty since the identity $e \in M$ is invertible, because $ee = e$ by definition, so $e \in T$. Let $a, b \in T$. First we need to prove that $ab \in T$, meaning that ab is invertible. We multiply the inverses of b and a to get $b^{-1}a^{-1}$. Since $abb^{-1}a^{-1} = b^{-1}a^{-1}ab = ee = e$, $ab \in T$. Thus T is a subsemigroup. Since $e \in T$ serves as an identity in T as well, T is a submonoid. Finally, every element $a \in T$ has an inverse element a^{-1} that is also invertible with inverse a, so $a^{-1} \in T$. Thus, T is a subgroup. □

This subgroup of invertible elements is called the *group of units*, similar to the name given to multiplicatively invertible elements in a ring. We will use this group in part of the proof of the Krohn-Rhodes theorem in Section 3.

Similar to how an ideal is defined for rings, an *ideal* of a semigroup S is a non-empty subset T such that for any $s \in S$ and any $t \in T$, both st and ts are in T. In words, T is closed under both left and right multiplication by any element in S. Of course, any ideal is a subsemigroup of S.

Example 2.5. (from [8])
Let S be the semigroup $(\mathbf{Z}^+, +)$. For a positive integer n, let $I_n = \{m \in \mathbf{Z}^+ \mid m \geq n\}$. Then I_n is an ideal of S.

We will now define the concept of *generating* a subsemigroup. Let X be a non-empty subset of a semigroup S and let T be the set of subsemigroups of S that contain X. This set T has at least one element, the semigroup S itself. Now take the intersection of all the subsemigroups in T, $\bigcap T$. Since every subsemigroup in T contains X, the intersection $\bigcap T$ is not empty. Also, it is a subsemigroup. It is easy to see this: for any $a, b \in \bigcap T$, a and b are in every subsemigroup in T, so ab is in every subsemigroup in T, thus $ab \in \bigcap T$. In fact, $\bigcap T$ is the smallest subsemigroup of S that contains X. We denote this subsemigroup by $\langle X \rangle$, and we call it the *subsemigroup generated by* X. If X is a subset of S such that $\langle X \rangle = S$, then X is a *generating set* for S and we say that X *generates* S.

Theorem 2.6. *Let X be a non-empty subset of a semigroup S. Then $\langle X \rangle = \{x_1 x_2 \dots x_n \mid n \in \mathbf{Z}^+, x_i \in X\}$.*

Proof. Let $Y = \{x_1 x_2 \dots x_n \mid n \in \mathbf{Z}^+, x_i \in X\}$. By its definition, Y is closed under multiplication and so is a subsemigroup of S. We have that $X \subseteq Y$, since n can be 1 in the definition of Y. Then Y must be one of the subsemigroups in T that contain X, so by definition we have $\langle X \rangle \subseteq Y$. Every element in Y is a combination of one or more elements from X. Since $X \subseteq \langle X \rangle$ and $\langle X \rangle$ is closed under multiplication, $Y \subseteq \langle X \rangle$. Thus, $\langle X \rangle = Y$. □

If a semigroup S is generated by a single element, which means that $S = \langle\{x\}\rangle$ (which we denote by $\langle x\rangle$) for some $x \in S$, then S is called a *monogenic semigroup*. By Theorem 2.6, $S = \{x^n \mid n \in \mathbf{Z}^+\}$. This is the same concept as a cyclic group, and a monogenic semigroup is sometimes called a cyclic semigroup. We will call it a monogenic semigroup following [8]. This property will be used in the proof of the Krohn-Rhodes theorem.

Recall that a *transformation* is a function $s : Q \to Q$ over a finite set Q. A *transformation semigroup* is a semigroup of transformations over Q, denoted (Q, S) where S is the semigroup of transformations. When dealing with semigroup theory, we will not use this notation (Q, S) since we are not concerned with what the elements of the set Q are; we can just call them $0, 1, \ldots, n-1$ where $n = |Q|$. When dealing with finite automata, the set Q would be the set of states, but that will come later. We denote the semigroup of all transformations on n elements under function composition as T_n. Note that T_n is actually a monoid since the identity transformation that sends each element to itself is included in T_n. Also, T_n has n^n elements, since there are n choices of where to map each of the n elements. Since permutations are bijective transformations, the symmetric group S_n is a subset of T_n. In fact, it's a subgroup of T_n, since 1) the composition of two permutations is a permutation, 2) the identity transformation is a permutation and so is in S_n, and 3) each permutation has an inverse permutation formed by the inverse bijection.

We introduce matrix notation for transformations similar to that used for permutations in group theory. Specifically, the transformation takes the element in entry $(1, j)$ to the element in entry $(2, j)$. Consider the semigroup of transformations T_2, which has $n^n = 2^2 = 4$ elements, each a transformation on $\{0, 1\}$. We write down all its elements in $2 \times n = 2 \times 2$ matrix form below.

$$\begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}.$$

Composing transformations is straightforward: we simply start with the element and follow wherever it ultimately leads to in the chain of compositions. If $a \in \{0, 1, \ldots, n-1\}$ and ϕ and ψ are two transformations on $\{0, 1, \ldots, n-1\}$, then $(\phi\psi)a = \phi(\psi(a))$, where we apply ψ first then ϕ.

Homomorphisms and isomorphisms are defined just as they are in group theory. Specifically, a *semigroup homomorphism* is a mapping $\phi : S \to T$, where S and T are two semigroups, such that $\phi(ab) = \phi(a)\phi(b)$ for any $a, b \in S$. An *isomorphism* is a bijective homomorphism. If two semigroups S and T are isomorphic, we denote this by $S \approx T$. We can prove an analogue of Cayley's theorem for semigroups, which shows the importance of transformation semigroups as more concrete representations of abstract semigroups similarly to permutations and abstract groups. Recall that Cayley's theorem states that every group of order n is isomorphic to a subgroup of S_n, the symmetric group of order n. We can slightly modify the proof of Cayley's theorem by first adjoining an identity to the semigroup in order to prove the following theorem. The statement of the theorem below is taken from [1]. We need to adjoin the identity to be able to prove that the homomorphism we construct is injective in

the same way we do in the original proof. We don't present the proof here (which is very similar to the original); this theorem is important but not relevant to our purposes in this paper.

Theorem 2.7. *(Semigroup analogue of Cayley's theorem)*
Every semigroup is isomorphic to a subsemigroup of T_{n+1}, the semigroup of transformations on $n+1$ elements.

A *left zero* in a semigroup S is an element x such that $xy = x$ for all $y \in S$. A *right zero* is an element x such that $yx = x$ for all $y \in S$. An element x is a zero of S if $xy = yx = x$ for all $y \in S$. If every element of S is a left zero, which means that $xy = x$ for all $x, y \in S$, then S is called a *left zero semigroup*. A *right zero semigroup* is defined analogously.

A semigroup is *simple* if it contains no proper ideals, which means that the only ideal of S is S itself. A *left simple* semigroup is one with no proper *left ideals*, which are subsemigroups closed under only *left* multiplication by every element in S. *Right simple semigroups* and *right ideals* are defined analogously. Note that this is a different concept from simple groups.

We now define the semigroup analogue of factor groups. We need another definition first. An equivalence relation ρ on a semigroup S is a *left congruence* if for all $x, y, z \in S$, $x \ \rho \ y \implies zx \ \rho \ zy$; a *right congruence* if for all $x, y, z \in S$, $x \ \rho \ y \implies xz \ \rho \ yz$; and a *congruence* if it is both a left congruence and a right congruence. Let $1_S = \{(s, s) \mid s \in S\}$ be the identity relation on S. Let I be an ideal of S. Then $\rho_I = (I \times I) \cup 1_S$ is a congruence on S.

Proof. For any $(a, b) \in \rho_I$, either $a = b$ or $a, b \in I$. In the first case, we have $xa = xb$ and $ax = bx$ so $(xa, xb), (ax, bx) \in \rho_I$. In the second case, we have $xa, xb, ax, bx \in I$ so $(xa, xb), (ax, bx) \in \rho_I$ as well. □

We can denote the quotient set S/ρ_I by S/I, and we call S/I the *Rees factor semigroup by I*. The elements of S/I are the ρ_I-classes, which comprise I and singleton sets $\{x\}$ for each $x \in S - I$, according to the definition of ρ_I above. For two elements $[x]_{\rho_I}$ and $[y]_{\rho_I}$ of S/I, where x and y are representatives of their respective ρ_I-classes, multiplication is defined by $[x][y] = [xy]$. This multiplication is well defined.

Proof. Let $[x]_{\rho_I} = [x']_{\rho_I}$ and $[y]_{\rho_I} = [y']_{\rho_I}$. In words, x and x' are possibly different representatives of the same ρ_I-class, and similarly for y and y'. We will prove that $[xy]_{\rho_I} = [x'y']_{\rho_I}$. First, note that if ρ is some congruence on S, then for all $x, y, a, b \in S$, $(x \ \rho \ y) \wedge (a \ \rho b) \implies (xa \ \rho \ yb)$. This is because we have (1) $xa \ \rho \ ya$ from $x \ \rho \ y$ since ρ is a right congruence, and (2) $ya \ \rho \ yb$ from $a \ \rho \ b$ since ρ is a left congruence. Then since ρ is transitive, (1) and (2) yield $xa \ \rho \ yb$, as required. Now by our assumption we have $x \ \rho_I \ x'$ and $y \ \rho_I \ y'$. Then by the property just proven, we have $xy \ \rho_I \ x'y'$, so $[xy]_{\rho_I} = [x'y']_{\rho_I}$. Thus, the multiplication is well defined. □

Note that I is a zero of S/I. Figure 2.1 (from [8]) shows how one forms S/I from S by merging elements of I to form a zero.

A semigroup E is an *ideal extension* of a semigroup S by a semigroup T if S is an ideal of E and $E/S \approx T$. Note that since E/S contains a zero (S), T must contain a zero for this ideal extension to exist.

Figure 2.1

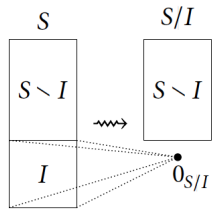


FIGURE 1.5
Forming S/I from S by merging elements of I to form a zero.

2.1. **Semidirect product.** Let S and T be semigroups. The following definition of a left action and a semidirect product is partially taken from [4]. For clarity, denote the multiplication in S by $+$, and denote the multiplication in T normally by t_1t_2. A *left action* ϕ *of* T *on* S is a function $\phi : T \times S \to S$ defined by $(t, s) \to t \cdot s$, where $t \cdot s$ denotes some $s' \in S$, such that for all $s, s_1, s_2 \in S$ and all $t, t_1, t_2 \in T$, we have **(1)** $t_1 \cdot (t_2 \cdot s) = (t_1t_2) \cdot s$, and **(2)** $t \cdot (s_1 + s_2) = t \cdot s_1 + t \cdot s_2$. Intuitively, the left action defines for every $t \in T$ a corresponding *endomorphism* on S, which is a homomorphism from S to itself. It is a homomorphism because of rule **(2)** above.

The *semidirect product of* S *and* T *with respect to* ϕ, which we denote by $S \times_\phi T$, is defined on $S \times T$ by $(s, t)(s', t') = (s + t \cdot s', tt')$. Note that if we take the trivial left action where $t \cdot s = s$ for all $t \in T$ and $s \in S$, we get $(s, t)(s', t') = (s + s', tt')$, so the semidirect product generalizes the direct product. Also, the semidirect product has cardinality $|S||T|$ just like the direct product. Finally, note that this multiplication defined for the semidirect product is associative (proof left as an exercise), and since the result of the multiplication is also in $S \times T$, $S \times_\phi T$ is closed under this multiplication and so it is a semigroup.

2.2. **Wreath product.** We finally start defining the wreath product! We first introduce new notation. If we have a direct product of n copies of a set X, which would be $X \otimes X \cdots \otimes X$ n times, we can denote this by X^A where $A = \{1, \ldots, n\}$. In this way, the copies of the set X are *indexed* by A. We write this instead of X^n to differentiate between this direct product and the set $\{x_1x_2 \ldots x_n \mid x_i \in X\}$ where the multiplication is defined for the set X (which may be a semigroup or another similar object). We can define the set X^A more formally, and more generally, by defining it as the set of functions from A to X. This corresponds to the direct product definition for the following reason. For any element $x \in X^A$, we have $x = (x_1, x_2, \ldots, x_n)$ for $x_i \in X$ as per the direct product definition. One can look at each x_i as an element in X selected as an output of a function f from A to X. In this way, x_1 would be the output $f(1)$, and any x_i would be the output $f(i)$. Thus, any element $x = (x_1, x_2, \ldots, x_n)$ corresponds to a unique function f from A to X. We will use this second definition of X^A as the set of functions from A to X in the following, when we define the wreath product. Specifically, we will consider every element of X^A as a function f. Note that $|X^A| = |X|^{|A|}$, since for every element ("index") $i \in A$ there are $|X|$ choices for a corresponding element $f(i)$.

Let S and T be semigroups. Define a left action ϕ of T on S^T by letting $y \cdot f$ for $y \in T$ and $f \in S^T$ be such that $(y \cdot f)(x) = f(xy)$ for all $x \in T$. (Recall that f is a function from T to S and $(y \cdot f)$ is a function $f' \in S^T$ that is also from T to S.) First, we show that this satisfies the definition of a left action. The proof of rule (1) is taken from [8], while the proof of rule (2) is our own. ([8] defined left actions differently.)

Proof. Let $y, z \in T$ and $f \in S^T$. For rule (1), we need to prove that $z \cdot (y \cdot f) = (zy) \cdot f$. Since these expressions are functions, we start with any $x \in T$ and show that the left-hand side applied on x gives the right-hand side applied on x. In particular, $(z \cdot (y \cdot f))(x) = (y \cdot f)(xz) = f(xzy) = ((zy) \cdot f)(x)$, where we used the way the left action was defined.

Let $y \in T$. For rule (2), we need to prove that $(y \cdot (f_1 f_2))(x) = (y \cdot f_1)(x)(y \cdot f_2)(x)$ for all $x \in T$. By the definition of this left action, we equivalently need to prove that $(f_1 f_2)(xy) = f_1(xy) f_2(xy)$. Note that each of f_1, f_2, and $f_3 = f_1 f_2$ corresponds to an elemnet in the direct product S^T, where f_3 is the product of pointwise multiplication between f_1 and f_2. Applying these functions to x gives an element in S. Let $f_3(x) = s_3$. Then $s_3 = s_1 s_2$ (multiplication in S), where $s_1 = f_1(x)$ and $s_2 = f_2(x)$; these elements are at position x in their respective $|T|$-tuples in the direct product. From $s_3 = s_1 s_2$ we have $f_3(x) = (f_1 f_2)(x) = f_1(x) f_2(x)$, from which rule (2) follows. □

The *wreath product of S and T*, denoted $S \wr T$, is the semidirect product $S^T \times_\phi T$, where ϕ is the left action defined earlier. So the product in $S \wr T$ is defined by $(f_1, t_1)(f_2, t_2) = (f_1(t_1 \cdot f_2), t_1 t_2)$. (We don't use the clarifying + notation here for the operation of S^T, but it is clear which multiplication applies to which elements.) This multiplication is associative and closed since the semidirect product multiplication is associative and closed. Thus, $S \wr T$ is a semigroup. Note, however, that the wreath product as an operation on semigroups is not associative [8]. We leave it as a simple exercise to prove this (hint: starting with three semigroups, consider only their cardinalities). The cardinality of $S \wr T$ is clearly $|S|^{|T|}|T|$.

One might find it difficult to form an intuition for what the wreath product is or what it does, just from this definition. As we start proving some results involving the wreath product, we can hopefully become more familiar with it.

2.3. **Division.** In this section we introduce the concept of *division* of semigroups. Most of the proofs we give are taken from [8]; we sometimes clarify steps, fill in details, and fix mathematical typos.

A semigroup S *divides* a semigroup T, denoted $S \preceq T$, if S is a homomorphic image of a subsemigroup of T. In other words, there exists a subsemigroup T' of T such that there is a surjective homomorphism from T' to S. Clearly the divisibility relation $\preceq$ is reflexive. It is also transitive. Before proving this, we prove a basic result on subsemigroups under homomorphisms in the following lemma.

Lemma 2.8. *Let ϕ be a surjective homomorphism from a semigroup S to a semigroup T. If T' is a subsemigroup of T, then $\phi^{-1}(T') = \{s \in S \mid \phi(s) \in T'\}$ is a subsemigroup of S.*

Proof. Let $S' = \phi^{-1}(T')$. S itself is a semigroup so we know that multiplication in S' is associative. Assume that S' is not closed under this multiplication, so that there exist two elements $s_1, s_2 \in S'$ such that $s_1 s_2 \notin S'$. Then $\phi(s_1 s_2) \notin T'$, since S' is the pullback of T' under ϕ. Since $\phi(s_1), \phi(s_2) \in T'$, we have $\phi(s_1)\phi(s_2) \in T'$ since T' is a subsemigroup. But $\phi(s_1)\phi(s_2) = \phi(s_1 s_2)$ because ϕ is a homomorphism, which is a contradiction. Thus, S' is closed and so is a subsemigroup of S. □

The following proof follows the one in [8] and fixes a small mathematical typo (the restriction is to U'' not U').

Theorem 2.9. ***(Transitivity of divisibility. Proposition 7.8 in*** [8]***)*** *The divisibility relation $\preceq$ is transitive.*

Proof. Let S, T, U be semigroups with $S \preceq T$ and $T \preceq U$. Then there are subsemigropus T' of T and U' of U and surjective homomorphisms $\phi : T' \to S$ and $\psi : U' \to T$. Let $U'' = \psi^{-1}(T')$, the pullback of T' under ψ. Since T' is a subsemigroup of T, U'' is a subsemigroup of U' by Lemma 2.8. Thus U'' is a subsemigroup of U. Let Φ be the restriction of ψ to U'', i.e. its domain is U''. Then the function composition $\Phi\phi : U'' \to S$ is a surjective homomorphism. Thus $S \preceq U$. □

The following theorem will be useful in the proof of the Krohn-Rhodes theorem in Section 3.

Theorem 2.10. ***(Proposition 7.9 in*** [8]***)*** *Let S and T be semigroups. Then S, T, and their direct product $S \otimes T$ divide their wreath product $S \wr T$.*

Proof. First, we have that $S \preceq S \otimes T$ and $T \preceq S \otimes T$ since S and T are homomorphic images of $S \otimes T$ under the projection maps $\pi_S : S \otimes T \to S$ and $\pi_T : S \otimes T \to T$. Since division is transitive by Lemma 2.9, it is enough to prove that $S \otimes T \preceq S \wr T$.

For every $s \in S$, define $f_s \in S^T$ to be the element of the direct product with all components equal to s. Define a function $\phi : S \otimes T \to S \wr T$ by $\phi((s,t)) = (f_s, t)$. We first prove that this function is a homomorphism. Let $(s,t), (s',t') \in S \otimes T$. By definition, we have $\phi((s,t))\phi((s',t')) = (f_s, t)(f_{s'}, t')$. Since we are now in the wreath product, we perform the multiplication to get $(f_s, t)(f_{s'}, t') = (f_s(t \cdot f_{s'}), tt')$. This is equal to $(f_{ss'}, tt')$. To see this, take any $x \in T$. First note that $(f_s(t \cdot f_{s'}))(x) = f_s(x)(t \cdot f_{s'})(x)$, as we showed when we proved the validity of the left action we defined when defining the wreath product. Now $f_s(x)(t \cdot f_{s'})(x) = ss'$, since each component in f_s is s and each component in $f_{s'}$ is s'. Finally $ss' = f_{ss'}(x)$ for the arbitrary x we picked, by definition. Thus we have $(f_{ss'}, tt')$. In turn this is equal to $\phi((ss', tt'))$ by definition. Since the product given as an argument to ϕ is in the direct product $S \otimes T$, this is equal to $\phi((s,t)(s',t'))$. So ϕ is a homomorphism.

We now show that ϕ is injective. Assume that $\phi((s,t)) = \phi((s',t'))$ for some $(s,t), (s',t') \in S \otimes T$. Then $(f_s, t) = (f_{s'}, t')$, so we have $s = s'$ and $t = t'$. Thus $(s,t) = (s',t')$, as required.

Since ϕ is injective, restricting the codomain of ϕ to its image gives the isomorphism $\phi : S \otimes T \to \phi(S \otimes T)$. So ϕ^{-1} is also an isomorphism, and particularly a surjective homomorphism, from the subsemigroup $\phi(S \otimes T)$ of $S \wr T$ to the semigroup $S \otimes T$. Therefore, $S \otimes T \preceq S \wr T$. □

The following two theorems will also be used in the Krohn-Rhodes theorem proof. We present their statements here, but we refer the interested reader to [8] for their proofs. They are Propositions 7.10 and 7.11 in Chapter 7. We omit them to leave space for the other proofs we will carry out.

Theorem 2.11. ***(Proposition 7.10 in*** [8]***)***
Let M be a monoid and let E be an ideal extension of M by T. Then $E \preceq T \wr M$.

Theorem 2.12. ***(Proposition 7.11 in*** [8]***)***
If $S' \preceq S$ and $T' \preceq T$, then $S' \wr T' \preceq S \wr T$.

We now give a new definition, from [8]. Let S be a semigroup and let S' be a set in bijection with S under the mapping $x \to x'$. Define a multiplication on $S \cup S'$ as follows. Multiplication in S is as before, and for all $x, y \in S$, **(1)** $xy' = x'y' = y'$, and **(2)** $x'y = (xy)'$. One can prove that this multiplication is associative (see Exercise 7.9 in [8]). Thus, the set $S \cup S'$ is a semigroup, called the *constant extension* of S and denoted $C(S)$. The following theorem will also be used in Section 3.

The following proof follows Proposition 7.12 in [8], but we add the details of the four cases when proving that $\hat{\phi}$ is a homomorphism, which are omitted in the proof in [8].

Theorem 2.13. ***(Proposition 7.12 in*** [8]***)***
If $S \preceq T$, then $C(S) \preceq C(T)$.

Proof. Suppose $S \preceq T$. There are three possibly overlapping cases: (1) S is a subsemigroup of T, (2) S is a homomorphic image of T, and (3) S is a homomorphic image of a subsemigroup of T. That the theorem holds for (1) and (2) implies that the theorem holds for (3), as we will show. For case (1), it is easy to see that if S is a subsemigroup of T then $C(S)$ is a subsemigroup of $C(T)$.

For case (2), suppose S is a homomorphic image of T. Then there is a surjective homomorphism $\phi : T \to S$. Define $\hat{\phi} : C(T) \to C(S)$ by $\hat{\phi}(x) = \phi(x)$ and $\hat{\phi}(x') = (\phi(x))'$. We will prove that $\hat{\phi}$ is a homomorphism. Let $x, y \in C(T)$. There are four cases to consider.

Case A: $x, y \in T$. Then $xy \in T$, so $\hat{\phi}(xy) = \phi(xy)$ by the first rule in the definition of $\hat{\phi}$. Then $\phi(xy) = \phi(x)\phi(y)$ since ϕ is a homomorphism. Finally $\phi(x)\phi(y) = \hat{\phi}(x)\hat{\phi}(y)$, again by definition. So $\hat{\phi}(xy) = \hat{\phi}(x)\hat{\phi}(y)$, as required.

Case B: $x \in T, y' \in T'$. Then $\hat{\phi}(xy') = \hat{\phi}(y')$ by the rules of multiplication in $C(T)$. By definition of $\hat{\phi}$, $\hat{\phi}(y') = (\phi(y))'$. By a $C(T)$ multiplication rule, $zw' = w'$ for all $z \in S, w' \in S'$. Specifically, we have $\phi(x)(\phi(y))' = (\phi(y))'$ since $\phi(x) \in S$ and $(\phi(y))' \in S'$. Finally $\phi(x)(\phi(y))' = \hat{\phi}(x)\hat{\phi}(y')$. Thus $\hat{\phi}(xy') = \hat{\phi}(x)\hat{\phi}(y')$, as required.

Case C: $x' \in T', y \in T$. We start from the other direction: take $\hat{\phi}(x')\hat{\phi}(y)$. By definition of $\hat{\phi}$ this is equal to $(\phi(x))'\phi(y)$. By the second rule of multiplication in $C(T)$ $(x'y = (xy)')$, this equals $(\phi(x)\phi(y))'$. Since ϕ is a homomorphism, this equals $(\phi(xy))'$. By definition, $(\phi(xy))' = \hat{\phi}((xy)')$. Using the rule $x'y = (xy)'$ again, we get $\hat{\phi}(x'y)$. Thus $\hat{\phi}(x')\hat{\phi}(y) = \hat{\phi}(x'y)$, as required.

Case D: $x', y' \in T'$. By the multiplication rules, $\hat{\phi}(x'y') = \hat{\phi}(y')$. By definition this equals $(\phi(y))'$. Taking $w' = (\phi(y))' \in S'$, note that $w' = z'w'$ for all $z' \in S'$ by the first multiplication rule. In particular, we have $(\phi(y))' = (\phi(x))'(\phi(y))'$, which equals $\hat{\phi}(x')\hat{\phi}(y')$ by definition. Thus $\hat{\phi}(x'y') = \hat{\phi}(x')\hat{\phi}(y')$, as required.

We've established that $\hat{\phi}$ is a homomorphism. It is also surjective by definition. Thus $C(S)$ is a homomorphic image of $C(T)$.

We need to show that the truth of the theorem for cases (1) and (2) above implies the truth of the theorem in general. Start with $S \preceq T$ and assume neither (1) nor (2) applies, so that S is a homomorphic image of a (proper) subsemigroup T' of T. Apply case (2) to S and T' to get that $C(S) \preceq C(T')$. Then apply case (1) to T' and T to get that $C(T') \preceq C(T)$. By transitivity of divisibility, we have $C(S) \preceq C(T)$, as required. □

The following result and its corollary will also be useful to us in the next section. We skip their proofs in the interest of space. Note that the corollary follows easily from the fact that $C(S)^M$ is a direct product of $|M|$ copies of $C(S)$ along with Theorems 2.10, 2.12, and 2.9 (transitivity of divisibility).

Theorem 2.14. ***(Proposition 7.13 in*** [8]***)***
Let M be a monoid and S a semigroup. Then $C(S \wr M) \preceq C(S)^M \wr C(M)$.

Corollary 2.15. ***(Corollary 7.14 in*** [8]***)***
Let M be a finite monoid and S a semigroup. Then $C(S \wr M)$ divides divides a wreath product of copies of $C(S)$ and $C(M)$.

Now we are ready to start proving the Krohn-Rhodes theorem!

3. The Krohn-Rhodes decomposition theorem in semigroup theory

In this section, we will prove the semigroup theory version of the Krohn-Rhodes decomposition theorem. The proofs in this section follow almost exactly the proofs in [8], which presents the topic in the clearest and most accessible way out of the sources we have found. The proofs are reproduced in my own words, where I sometimes clarify steps, fill in details, fix mathematical typos, and use different notation that is more familiar. Note that the notation used throughout this paper follows the notation used in [12], except for things defined in [8] but not in [12].

First, a couple of definitions. A semigroup S is *aperiodic* if for every $x \in S$, there exists a positive integer n such that $x^n = x^{n+1}$. Now let U_3 be the monoid obtained by adjoining an identity to a two-element right zero semigroup $\{a, b\}$. Thus U_3 has elements $\{1, a, b\}$, and its multiplication table (taken from [8]) is shown in Figure 3.1.

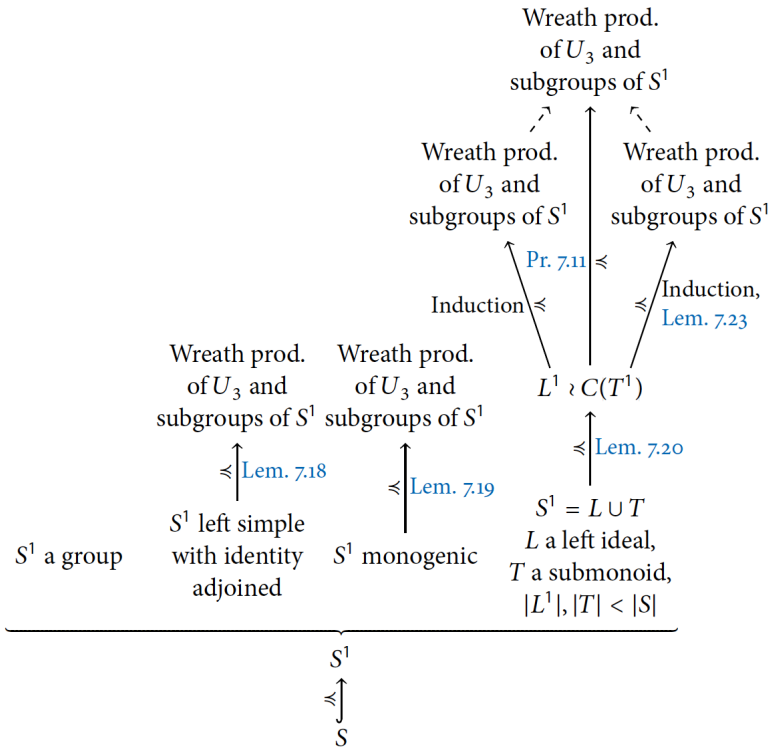
U_3 plays a key role in the Krohn-Rhodes decomposition. Notice that since $x^2 = x$ for all $x \in U_3$, U_3 is aperiodic. Cain, the author of [8], proves a stronger version of the Krohn-Rhodes theorem than the one we introduced in the abstract. This is the version we will prove. It states that every finite semigroup divides a wreath product of its own subgroups and copies of U_3. First, note that it is enough to prove the theorem for monoids, since $S \preceq S^1$ (trivially) and division is transitive. The proof will proceed by induction on the number of elements in the monoid. The main idea

Figure 3.1 [8]

	1	a	b
1	1	a	b
a	a	a	b
b	b	a	b

TABLE 7.1
Multiplication table of U_3.

Figure 3.2 [8]



of the induction is Lemma 3.3, which shows that a monoid S is either a group, a left simple semigroup with an identity adjoined, monogenic, or can be decomposed as $S = L \cup T$ where L is a left ideal and T is a submonoid and L^1 and T each have fewer elements than S. If S is a group, a valid decomposition into its own subgroups would be formed by only S itself, so the theorem is trivial for groups. The base cases of the induction are (1) left simple semigroups with identities adjoined (this case is proved in Lemma 3.7) and (2) monogenic semigroups (this case is proved in Lemma 3.8). The inductive step is used for the fourth case of decomposition as $S = L \cup T$. Figure 3.2 [8] shows the roles of the various lemmas, which can be helpful for the reader to keep track of the proof.

The following lemma is needed for the proof of Lemma 3.3.

Lemma 3.1. ***(Lemma 7.15 in*** [8]***)***
Let S be a finite semigroup. Then at least one of the following is true:

(1) *S is trivial;*
(2) *S is left simple;*
(3) *S is monogenic;*
(4) *$S = L \cup T$, where L is a proper left ideal of S and T is a proper subsemigroup of S.*

Proof. Suppose that none of (1), (2), and (3) are true. We will prove (4). Since S is not left simple, it contains proper left ideals. Since it is finite, it has a maximal proper left ideal K. A maximal ideal is defined as it is defined in [12] for rings. Let $x \in S - K$. Define the set $S^1x = \{sx \mid s \in S^1\}$. Then $K \cup S^1x$ is a left ideal that strictly contains K, since it clearly contains K and there is at least one element, namely x, in $K \cup S^1x$ but not in K (S^1x includes x because $1x = x \in S^1x$ where 1 is the identity in S^1). Since K is maximal, $K \cup S^1x = S$. If $S^1x \neq S$, let $L = K$ and $T = S^1x$ and the proof is complete.

So assume $S^1x = S$. Then $S = Sx \cup \{x\}$, since removing the identity from S^1 might remove x, and only x, from S^1x. Clearly $Sx \cup \{x\} \subseteq Sx \cup \langle x \rangle$. Since we have $Sx \cup \langle x \rangle \subseteq S$ by closure, $S = Sx \cup \langle x \rangle$. If $S \neq Sx$, then let $L = Sx$ and $T = \langle x \rangle$ and the proof is complete, since $T \neq S$ because S is not monogenic. (Note that L would be a left ideal because for any $s \in S$ and $a \in Sx$, $sa = ss'x$ for some $s' \in S$, and $ss'x \in Sx$ since $ss' \in S$.)

So assume $S = Sx$. Then every element in S is of the form sx for some $s \in S$. Let $M = \{y \in S \mid yx \in K\}$. Since every element in K has the form yx for some $y \in S$, M is nonempty (in fact $Mx = K$). Moreover, M is a left ideal of S. To see why, let $m \in M$ and $s \in S$. Then $mx \in K$, so $smx \in K$ since K is a left ideal. Since $(sm)x \in K$, $sm \in M$. Furthermore, M is a proper left ideal. To see why, assume not, so that $M = S$. But then $K = Sx = S$, contradicting that it is a proper left ideal. If $M \not\subseteq K$, then $M \cup K$ is a left ideal of S strictly containing the maximal left ideal K, so $M \cup K = S$. In this case, let $L = M$ and $T = K$ and the proof is complete.

So assume $M \subseteq K$. This means that for any $y \in S$, $yx \in K \implies y \in K$. Now repeat the process above, starting from selecting an $x \in S - K$, for all $x \in S - K$. Either some x allows us to complete the proof in any one of the steps above, or the implication $y \in S$, $yx \in K \implies y \in K$ holds for all $x \in S - K$. In the latter case, we can take the contrapositive of the implication to get that $y \in S - K \implies yx \in S - K$ for all $x \in S - K$. This means that $S - K$ is a subsemigroup! In this case, let $L = K$ and $T = S - K$, and the proof is complete. □

Before we prove Lemma 3.3, we describe a result that we will need, which is a version of Proposition 7.1 in [8]. Its proof relies on Green's relations, and is not given here.

Lemma 3.2. *Let M be a finite monoid and let G be its group of units. Then $M - G$ is either empty or an ideal of M.*

Lemma 3.3. ***(Lemma 7.16 in*** [8]***)***
Let S be a finite monoid. Then at least one of the following is true:

(1) *S is a group;*
(2) *S is a left simple semigroup with an identity adjoined;*
(3) *S is monogenic;*
(4) *$S = L \cup T$, where L is a left ideal of S and T is a submonoid of S, and L^1 and T each have fewer elements than S.*

Proof. Suppose that none of (1), (2), and (3) hold. We will prove (4). Let G be the group of units of S. Clearly 1, the identity of S, is in G. If 1 is the only element of G, we say that G is trivial. There are two cases to consider.

Case A: G is trivial. Then $S - G = S - \{1\}$ is an ideal by Lemma 3.2 since $S - G$ cannot be empty (if it is, S is a group, which we assumed it is not). Thus $S - G$ is a subsemigroup of S. Since S is not left simple with an identity adjoined, the semigroup $S - \{1\}$ is not left simple. Now we can apply Lemma 3.1 to $S - \{1\}$ to get that $S - \{1\} = L \cup Q$ where L is a proper left ideal of $S - \{1\}$ and Q a proper subsemigroup of $S - \{1\}$. Since $L \neq S - \{1\}$, we know that $L \cup \{1\} \neq S$, so the condition that L^1 has fewer elements than S is satisfied. Let $T = Q \cup \{1\}$. Since we had that Q is a proper subsemigroup of $S - \{1\}$, T is a proper submonoid of S and $S = L \cup T$.

Case B: G is nontrivial. Then let $L = S - G$ and let $T = G$. Again by Lemma 3.2, L is an ideal of S, since $S - G$ cannot be empty because S is not a group by our assumption. We have $S = L \cup T$. Since G is nontrivial, $L \cup \{1\} \neq S$. Since S is not a group, $T \neq S$. So the conditions are satisfied in this case as well.

We have thus proven (4) for both possible cases. □

In Lemma 3.7 we will prove the first base case of the induction: monoids consisting of a left simple semigroup with an identity adjoined. To do so, we first need the following lemma, which shows that the theorem holds for left zero semigroups. We don't include its proof.

Lemma 3.4. ***(Lemma 7.17 in*** [8]***)***
Every finite left zero semigroup divides a wreath product of copies of U_3.

We also need two more lemmas before we prove Lemma 3.7. Recall that an *idempotent* is an element $x \in S$ such that $x^2 = x$, where in our case S is a semigroup. The proof for the first one, Lemma 3.5, is from Mikko Korhonen's answer in [3]. We skip the proof of the second one, Lemma 3.6. It follows from Theorem 4.19 in [8], the proof of which uses Green's relations and other concepts we have not defined here.

Lemma 3.5. *If S is a finite semigroup, then S contains an idempotent.*

Proof. We need to find an idempotent $a \in S$. First, note that it is enough to prove that $a^k = a$ for some $k \geq 2$. To see why, start with $a^k = a$ for some $k > 2$ and multiply both sides by a^{k-2} to get $(a^{k-1})^2 = a^{k-1}$. Now let $x \in S$ and consider the sequence $x, x^2, x^4, x^8, x^{16}, \ldots$. Since S is finite, this sequence must have some repetition. Specifically, we must have $x^{2^i} = x^{2^j}$ for some $j > i \geq 1$. So we have

$x^{2^j} = (x^{2^i})^{2^{j-i}} = x^{2^i}$. Now take $a = x^{2^i}$ and $k = 2^{j-i}$. This yields $a^k = a$, and since $k \geq 2$ because $j > i$, the proof is complete. □

Lemma 3.6. *Let S be a semigroup. If S is left simple and contains an idempotent, then $S \approx Z \otimes G$, where Z is a left zero semigroup and G is a subgroup of S.*

We now prove the first base case of the induction.

Lemma 3.7. ***(Lemma 7.18 in** [8]**)***
Let S be a finite left simple semigroup. Then S^1 divides the wreath product of a subgroup of S and copies of U_3.

Proof. Since S is finite, it contains an idempotent by Lemma 3.5. So by Lemma 3.6, S is isomorphic to $Z \otimes G$, where Z is a left zero semigroup and G is a subgroup of S. By Lemma 3.4, Z divides a wreath product of copies of U_3. Thus by Theorems 2.10 and 2.12, $Z \otimes G$ divides a wreath product of G and copies of U_3. □

Now we prove the other base case: monogenic monoids. We fix a minor typo from [8] ($x^k = x^{k+1}$ instead of $x^k = x^k + 1$).

Lemma 3.8. ***(Lemma 7.19 in** [8]**)***
Let S be a finite monogenic monoid. Then S divides the wreath product of a subgroup of S and copies of U_3.

Proof. Consider a monogenic monoid $S = \{1, x, \ldots, x^k, \ldots, x^{k+m-1}\}$, where $x^{k+m} = x^k$. This way of writing a semigroup, and the fact that there are $k + m - 1$ distinct positive powers of x, are introduced in Chapter 1 of [8] when the author introduces the *periodicity* of a semigroup (in this case, S has period m). We don't need to fully cover these tangential concepts here, but note that the subset $\{x^k, \ldots, x^{k+m-1}\}$ is closed under multiplication by any element in S and is thus an ideal (from $x^{k+m} = x^k$). It is in fact a subgroup of S. Now we have that S is an ideal extension of the subgroup $G = \{x^k, \ldots, x^{k+m-1}\}$ by the monogenic monoid $C_k = \{1, x, \ldots, x^k\}$, with $x^k = x^{k+1}$. It is easy to verify that this is indeed an ideal extension, i.e. that $S/G \approx C_k$. As we know, the zero of S/G is the ideal G. The corresponding zero of C_k is x^k, which can be seen from $x^k = x^{k+1}$.

We use induction on k to show that C_k divides a subsemigroup of $C_{k-1} \wr C_1$. As for our base case, it is proven by observing that $C_1 = \{1, x\}$ (with $x^2 = x$) divides U_3, since it is isomorphic to the subsemigroup $\{1, a\}$ of U_3.

For $i \in \mathbf{Z}^+$, define $f_i : C_1 \to C_{k-1}$ by $f_i(1) = x^{i-1}$ and $f_i(0) = x^i$. First, notice that we denote $x \in C_1$ here more conveniently by 0, since it is a zero of C_1. We have defined this function to construct a subset U of $C_{k-1} \wr C_1$ which we will prove is a submonoid that is isomorphic to C_k, thus proving our desired result. Specifically, let $U = \{1\} \cup \{(f_i, 0) \mid i \in \mathbf{Z}^+\} \subseteq C_{k-1} \wr C_1$. Recall that this is a subset of $C_{k-1} \wr C_1$ because each element $(f, a) \in C_{k-1} \wr C_1$ is such that f is a function from C_1 to C_{k-1} and a is an element of C_1, which is always 0 in the case of U.

We will now prove that U is a submonoid. Take two elements $(f_i, 0), (f_j, 0) \in U$. Then $(f_i, 0)(f_j, 0) = (f_i(0 \cdot f_j), 0) = (f_{i+j}, 0)$. The first equality follows from the definition of the wreath product multiplication. As for the second, we show it by showing that $f_i(0 \cdot f_j) = f_{i+j}$, by showing that the two functions give equal values

for all possible input values. The possible inputs are the elements of C_1, 0 (x) and 1. For 0, we have $(f_i(0 \cdot f_j))(0) = f_i(0)f_j(0)$, which we get first by distributing then by using the rule of the left action corresponding to the wreath product. Now this is equal to $x^i x^j$ by definition, which in turn equals $x^{i+j} = f_{i+j}(0)$. As for 1, we have $(f_i(0 \cdot f_j))(1) = f_i(1)f_j(0) = x^{i-1}x^j = x^{i+j-1} = f_{i+j}(1)$. Thus we have proven the equality $f_i(0 \cdot f_j) = f_{i+j}$. Therefore, U is a submonoid of $C_{k-1} \wr C_1$.

In fact, notice that $(f_i, 0) = (f_1, 0)^i$ for all positive integers i, so U is the monogenic submonoid of $C_{k-1} \wr C_1$ generated by $(f_1, 0)$. We have that $(f_1, 0)^{k+1} = (f_{k+1}, 0) = (f_k, 0) = (f_1, 0)^k$. The second equality holds because we are in $C_{k-1} \wr C_1$, so $x^{k-1} = x^k = x^{k+1}$. Thus $f_{k+1}(1) = x^{k+1-1} = x^k = x^{k-1} = f_k(1)$, and the equality holds similarly for 0. On the other hand, $(f_1, 0)^k = (f_k, 0) \neq (f_{k-1}, 0) = (f_1, 0)^{k-1}$. Thus $(f_1, 0)^k$ is the largest distinct power, and so $U = \{1, (f_1, 0), (f_1, 0)^2, \ldots, (f_1, 0)^k\}$. From this it is clear that $U \approx C_k$. Therefore, $C_k \preceq C_{k-1} \wr C_1$, and by induction this holds for all k.

Now we have a chain of divisions where each division subtracts 1 from k until we ultimately reach a wreath product of copies of U_3 (recall that the base case was $C_1 \preceq U_3$). More concretely, every C_k divides a wreath product of copies of U_3 by Theorems 2.10 and 2.12. Since S is an ideal extension of G by C_k, it divides a wreath product of G and copies of U_3 by Theorem 2.11. □

Now we start working on the induction step, in the case of decomposition of the monoid S as $L \cup T$ where L is a left ideal and T is a subsemigroup. We list four lemmas and prove one of them, Lemma 3.10. Of course, we refer the interested reader to [8] for the proofs of the others.

Lemma 3.9. ***(Lemma 7.20 in*** [8]***)***
Let S be a semigroup and suppose $S = L \cup T$, where L is a left ideal of S and T is a subsemigroup of S. Then $S \preceq L^1 \wr C(T^1)$.

Cain says that the "following result is essentially a more precise verison of Lemma 7.20 that holds when we decompose a monoid into the union of its group of units and the set of remaining elements." [8] In the proof, we fix a small typo towards the end; it should be "for all $x \in I^1$" not "for all $x \in S$".

Lemma 3.10. ***(Lemma 7.21 in*** [8]***)***
Let S be a monoid and let G be its group of units. Then $I = S - G$ is an ideal of S and $S \preceq I^1 \wr G$.

Proof. By Lemma 3.2, $I = S - G$ is an ideal of S. For each $x \in I^1$, define a function $f_x : G \to I^1$ by $f_x(g) = gxg^{-1}$ for all $g \in G$. Notice that $f_1(g) = gg^{-1} = 1$ for all $g \in G$. Let $V = \{(f_x, g) \mid x \in I^1, g \in G\}$. We will prove that V is a subsemigroup of $I^1 \wr G$. Let $(f_x, g), (f_y, h) \in V$. We need to prove that the product of the multiplication $(f_x, g)(f_y, h) = (f_x(g \cdot f_y), gh) \in V$. Clearly $gh \in G$. Now take any $k \in G$. Then $(f_x(g \cdot f_y))(k) = f_x(k)f_y(kg)$. By definition, $f_x(k)f_y(kg) = kxk^{-1}kgyg^{-1}k^{-1} = k(xgyg^{-1})k^{-1} = f_{xgyg^{-1}}(k)$. Note that $xgyg^{-1} \in I^1$ since $x \in I^1$ and I is an ideal. So we have $(f_x, g)(f_y, h) = (f_{xgyg^{-1}}, gh) \in V$. Thus V is a subsemigroup of $I^1 \wr G$.

Define a function $\phi : V \to S$ by $\phi((f_x, g)) = xg$. This function is well defined since $f_x = f_y \implies f_x(1) = f_y(1) \implies x = y$. We will now prove that it is a homomorphism. We have $\phi((f_x, g)(f_y, h)) = \phi((f_{xgyg^{-1}}, gh)) = xgyg^{-1}gh = xgyh =$

$\phi((f_x, g))\phi((f_y, h))$, as required. Now note that $G \subseteq \phi(V)$, the image of ϕ, since $\phi((f_1, g)) = g$ for all $g \in G$. Also, $I^1 \subseteq \phi(V)$ since $\phi((f_x, 1)) = x$ for all $x \in I^1$. Because $S = I \cup G$, this means that ϕ is surjective. Thus, $S \preceq I^1 \wr G$. □

The following lemma shows that the theorem holds for right zero semigroups. We use it to prove the next lemma [8].

Lemma 3.11. ***(Lemma 7.22 in*** [8]***)***
Every finite right zero semigroup, and every finite right zero semigroup with an identity adjoined, divides a wreath product of copies of U_3.

Lemma 3.12. ***(Lemma 7.23 in*** [8]***)***
Let S be a finite semigroup. If S divides a wreath product of groups and copies of U_3, then $C(S)$ divides a wreath product of copies of those same groups and copies of U_3.

We are finally ready to prove the Krohn-Rhodes decomposition theorem. As we prove it, the role each lemma plays, in conjunction with the induction, will become apparent, and Figure 3.2 will hopefully make more sense.

Theorem 3.13. ***(Krohn-Rhodes Theorem. Theorem 7.24 in*** [8]***)***
Let S be a finite semigroup. Then S divides a wreath product of subgroups of S and copies of U_3.

Proof. Let S be a semigroup. Since $S \preceq S^1$, we may assume that S is a monoid. We will induct on the number of elements in S. The base case of the induction is when S has one element. In this case, S is trivial, so it is a group and the result holds immediately.

Now assume that the theorem holds for all monoids with fewer elements than S. The result clearly holds if S is trivial or if it is any group. By Lemma 3.7, it holds if S is a left simple semigroup with an identity adjoined. By Lemma 3.8, it holds if S is monogenic.

So assume that S is not trivial, not a group, not a left simple semigroup with an identity adjoined, and not monogenic. Then by Lemma 3.3, $S = L \cup T$, where L is a left ideal and T is a submonoid of S and each of L^1 and T has fewer elements than S. So by our induction hypothesis, L^1 divides a wreath product of subroups of L^1 (which are also subgroups of S) and copies of U_3, and similarly T divides a wreath product of subgroups of T (which are also subgroups of S) and copies of U_3. By Lemma 3.12, $C(T^1)$ also divides a wreath product of subgroups of S and copies of U_3. By Lemma 3.9, $S = L \cup T$ divides $L^1 \wr C(T^1)$. Then by Theorem 2.12, S divides a wreath product of subgroups of S and copies of U_3, as shown in Figure 3.2. This completes the induction, so the result holds for all monoids S. □

We briefly discuss the connection to automata. The semigroup U_3 corresponds to a reset automaton, also called a flip-flop, referring to flip-flop circuits used in electronics. Its three elements can be described as "set", "reset", and "do nothing" [2]. The subgroups in the wreath product correspond to so-called permutation automata, whose corresponding transformation semigroups are actually permutation groups (the subgroups in the wreath product). These are easier to deal with, one reason being that they represent reversible computations: no two states q_1 and q_2 transition to the same state q_3 with the same input letter. Reset automata and permutation automata

are very simple automata, so Theorem 3.13 gives a powerful result about the ability to decompose any complicated automaton into these two simple types. We refer the reader to [5] for an explanation of the Krohn-Rhodes theorem in automata theory, along with an algorithm to actually perform the decomposition, named the holonomy decomposition.

References

[1] https://cameroncounts.files.wordpress.com/2017/03/pgts.pdf.

[2] https://en.wikipedia.org/wiki/semigroup$_{w}ith_{t}hree_{e}lements$.

[3] https://math.stackexchange.com/questions/353028/is-there-an-idempotent-element-in-a-finite-semigroup.

[4] https://math.stackexchange.com/questions/985857/whats-a-semidirect-product-of-semigroups.

[5] http://www-verimag.imag.fr/ maler/papers/kr-new.pdf.

[6] http://www.intel.com/design/intarch/prodbref/272713.html.

[7] Stanford cs 103 lecture. http://web.stanford.edu/class/cs103/lectures/14/small14.pdf.

[8] Alan J. Cain. Nine chapters on the semigroup art. 2013.

[9] Attila Egri-Nagy. Applications of automata theory and algebra via the mathematical theory of complexity to biology, physics, psychology, philosophy, and games. john rhodes. chrystopher l. nehaniv (ed.). foreword by morris w. hirsch. (2009, world scientific books.) isbn: 978-981-283-696-0, us$65 (hardcover); isbn: 978. *Artificial Life*, 17:141–143, 2011.

[10] Attila Egri-Nagy, James D. Mitchell, and Chrystopher L. Nehaniv. Sgpdec: Cascade (de)compositions of finite transformation semigroups and permutation groups. In *ICMS*, 2014.

[11] Attila Egri-Nagy and Chrystopher Nehaniv. Computational holonomy decomposition of transformation semigroups. 08 2015.

[12] Joseph Gallian. Contemporary abstract algebra / joseph a. gallian. *SERBIULA (sistema Librum 2.0)*, 12 2018.

E-mail address: helezabi@aucegypt.edu