Fractional Elements in Abelian Automata Groups

Chris Grossack[0000−0001−7520−9747]

Carnegie Mellon University, Pittsburgh, USA
cgrossac@andrew.cmu.edu

Abstract. TODO: write this

Keywords: Abelian Automata · Transducer · Module Theory

TODO: change the reference Sutner18 to the published version

1 Background

Finite State Automata are combinatorial objects which encode relations between words over some alphabet. Automata provide deep connections between combinatorics, algebra, and logic, and are essential tools in contemporary computer science. One such link is in the decidability of truth in a structure whose relations are all computable by automata. One can combine these automata into more complicated automata representing logical sentences in such a way that a sentence is true if and only if a simple reachability condition holds [1]. This gives a simple proof that the theory of $\mathbb{N}$ with $+$ and $<$, for example, is decidable.

Different kinds of automata encode different kinds of information, and in this article we will be interested in **Mealey Automata** which encode functions from a set of words to itself. Indeed, the functions we consider will all be invertible (and the inverses are comutable by automata as well), and thus they will generate **Automata Groups**. These groups are surprisingly complicated, and a classification of all groups generated by three state automata over the alphabet $\mathbf{2} = \{0, 1\}$ is an extremely difficult problem, though much impressive progress has been made [3]. This complexity can be useful, as automata groups have become a rich source of examples and counterexamples in group theory [11, 15, 4]. Most notably, automata groups provide examples of finitely generated infinite torsion groups, with application to Burnside's Problem [7], and automata groups have provided the only examples of groups of intermediate growth, providing counterexamples to Milnor's Conjecture regarding the existence of such groups [6]. In fact, one of the simplest conceivable automata (shown below) already generates the lamplighter group $\mathbb{Z}/2\mathbb{Z} \wr \mathbb{Z}$ [5].

1.1 Some Important Definitions

Recall $\mathbf{2} = \{0, 1\}$. For our purposes, a **Mealey Automaton** is a tuple $\mathcal{A} = (S, \tau)$ where S is the **State Set**, and $\tau : S \times \mathbf{2} \to S \times \mathbf{2}$ is the **transition function**.

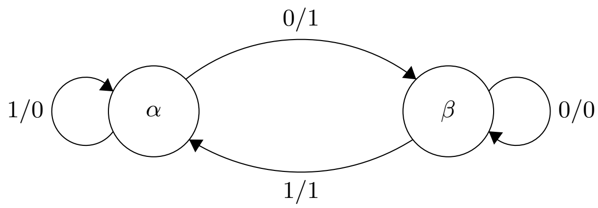


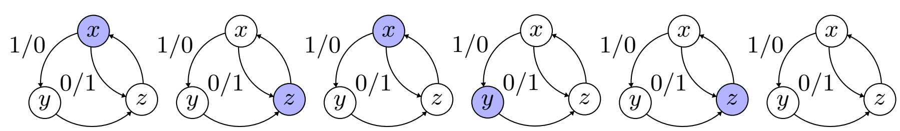
Fig. 1: An automaton generating the lamplighter group

We represent $\mathcal{A}$ as a (directed, multi-)graph with S as vertices and a labelled edge from s_1 to s_2 by a/b exactly when $\tau(s_1, a) = (s_2, b)$. Following Sutner, we remove clutter by writing a single unlabled edge in place of two parallel edges labelled 0/0 and 1/1. We write $(\partial_a s, \underline{s}(a)) = \tau(s, a)$ and call $\partial_0 s$ (resp. $\partial_1 s$) the **0-residual** (resp. **1-residual**) of s.

We can extend $\underline{s}$ to a length preserving endofunction on the free monoid $\mathbf{2}^*$ as follows (here juxtaposition is concatenation, and the empty word ε is the identity):

$\underline{s} : \mathbf{2}^* \to \mathbf{2}^*$

$$\begin{aligned} \underline{s}(\varepsilon) &= \varepsilon \\ \underline{s}(ax) &= a' \underline{s'}(x) \quad \text{(where } (s', a') = \tau(s, a)) \end{aligned}$$



(a) $\underline{x}(0110)$ (b) $1\underline{z}(110)$ (c) $11\underline{x}(10)$ (d) $110\underline{y}(0)$ (e) $1100\underline{z}(\varepsilon)$ (f) 1100

Fig. 2: An example computation – $\underline{x}(0110) = 1100$.

Clearly we can treat $\underline{s}$ as a function on $\mathbf{2}^\omega$, the set of infinite words, instead. In this case, automata provide a computable way of encoding complicated continuous functions from cantor space to itself, with ties to descriptive set theory[16]. If all of these functions are invertible, we let $\mathcal{G}(\mathcal{A})$ denote the group generated by these functions. We write our group additively, and denote the identity by I.

We can extend the definition of residuals to the whole group $\mathcal{G}(\mathcal{A})$ by defining the **0-residual** (resp. **1-residual**) of a function $f \in \mathcal{G}(\mathcal{A})$ as the unique function $\partial_0 f$ such that for all w, $f(0w) = f(0)\partial_0 f(w)$ (resp. $f(1w) = f(1)\partial_1 f(w)$). For a state $s \in S$, it is clear that $\partial_a \underline{s} = \underline{s'}$, where $(s', a') = \tau(s, a)$, so that this extends the old definition.

Thus each $\mathcal{G}(\mathcal{A})$ can also be viewed as an automaton, by taking $\mathcal{G}(\mathcal{A})$ as states, and defining $\tau(f, i) = \partial_i f$. Now $\mathcal{A}$ is a natural subautomaton of $\mathcal{G}(\mathcal{A})$ by identifying $s \in \mathcal{A}$ with $\underline{s} \in \mathcal{G}(\mathcal{A})$.

We will call a function **Odd** if it flips its first bit, and **Even** otherwise, and we call an automaton **Abelian** or **Trivial** exactly when its group is.

So in figure 1, $\underline{\alpha}$ is odd, $\underline{\beta}$ is even, $\partial_0\underline{\alpha} = \underline{\beta}$, and $\partial_1\underline{\alpha} = \underline{\alpha}$. For a more in depth description of Mealy Automata and their properties, see [14, 8]. (TODO: add Klaus' early paper)

Of great importance to abelian automata theory is the result of Nekrashevich and Sidki that every such group is either torsion free abelian or boolean [12]. Because of this classification, much of the interesting structure of these groups comes from the residuation functions. To that end, for the duration of this paper, homomorphisms and isomorphisms are all restricted to those which preserve the residuation structure in addition to the group structure. It is a theorem by Sutner [17] that $\mathcal{G}(\mathcal{A})$ is abelian iff for even states $\partial_1 f - \partial_0 f = I$ and for odd states $\partial_1 f - \partial_0 f = \gamma$, where γ is independent of f. Moreover, the case $\gamma = I$ corresponds precisely to the case where $\mathcal{G}(\mathcal{A})$ is boolean. We now restrict ourselves further to the case where $\mathcal{G}(\mathcal{A})$ is free abelian, that is to say $\mathcal{G}(\mathcal{A}) \cong \mathbb{Z}^m$ for some m, and $\gamma \neq I$.[1]

1.2 The Complete Automaton

From the discussion above, it follows that $\mathbb{Z}^m \cong \mathcal{G}(\mathcal{A})$ carries a residuation structure, and Nekrashvych and Sidki give a characterization of all posible such structures [12].

Without loss of generality, we can take the odd (resp. even) states to be exactly the vectors with odd (resp. even) first component. The automata structure is given by the following affine maps (which depend on a matrix $\mathbf{A}$ and an odd vector $\bar{e}$):

$$\tau(\bar{v}, 0) = \begin{cases} \mathbf{A}\bar{v} & \bar{v} \text{ even} \\ \mathbf{A}(\bar{v} - \bar{e}) & \bar{v} \text{ odd} \end{cases} \tag{1}$$

$$\tau(\bar{v}, 1) = \begin{cases} \mathbf{A}\bar{v} & \bar{v} \text{ even} \\ \mathbf{A}(\bar{v} + \bar{e}) & \bar{v} \text{ odd} \end{cases} \tag{2}$$

In the above definition, $\mathbf{A}$ is a $\frac{1}{2} - integral$ matrix $\mathbf{A}$ of $\mathbb{Q}$-irreducible character. This automaton group is generated by a *finite* automaton exactly if $\mathbf{A}$ is a contraction (that is, all of its complex eigenvalues have norm < 1). By a $\frac{1}{2} - integral$ matrix, we mean a matrix of the form

$$\begin{pmatrix} \frac{a_{11}}{2} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{a_{n1}}{2} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

[1] For historical reasons we use $\mathbb{Z}^m$ instead of $\mathbb{Z}^n$ because traditionally n is reserved for the size of the state set of an automaton.

where each $a_{ij} \in \mathbb{Z}$. These matrices all have characteristic polynomial $\chi = x^n + \frac{1}{2}g(x)$, where $g \in \mathbb{Z}[x]$ and has constant term ± 1. Without loss of generality we may take $\mathbf{A}$ to be in rational canonical form.

Since $\mathbf{A} : 2\mathbb{Z} \oplus \mathbb{Z}^{m-1} \to \mathbb{Z}^m$, residuation acts as multiplication by $\mathbf{A}$ for even vectors, but for odd vectors we need to first make them even. This is the reason for the **residuation vector** $\bar{e}$ in the above definition. It is easy to see that this definition gives rise to the following residuation structure:

If $\bar{v}$ is even:

$$\partial_0 \bar{v} = \partial_1 \bar{v} = \mathbf{A}\bar{v}$$

If $\bar{v}$ is odd:

$$\partial_0 \bar{v} = \mathbf{A}(\bar{v} - \bar{e})$$

$$\partial_1 \bar{v} = \mathbf{A}(\bar{v} + \bar{e})$$

Following Sutner [17], for a specific matrix $\mathbf{A}$ and residuation vector $\bar{e}$ we define **The Complete Automaton** $\mathfrak{C}(\mathbf{A}, \bar{e}) = (\mathbb{Z}^m, \tau)$, where τ is defined as in equations 1 and 2.

Nekrashevich and Sidki's theorem gives us a purely linear algebraic method of discussing these automaton groups, since a restatement of their theorem says that every torsion free abelian automaton group $\mathcal{G}(\mathcal{A})$ is isomorphic to $\mathfrak{C}(\mathbf{A}, \bar{e})$ for some $\mathbf{A}$ and $\bar{e}$, and therefore every abelian automaton $\mathcal{A}$ is a subautomaton of some $\mathfrak{C}(\mathbf{A}, \bar{e})$. Nekrashevych and Sidki show that each $\mathcal{A}$ has a unique matrix $\mathbf{A}$ (up to $\mathrm{GL}(\mathbb{Q})$ similarity) which works, though their proof is nonconstructive. We call this $\mathbf{A}$ (in rational canonical form) the **Associated Matrix** of $\mathcal{A}$.

We say a function $f \in \mathcal{G}(\mathcal{A})$ is **Located at** $\bar{v} \in \mathfrak{C}(\mathbf{A}, \bar{e})$ iff the isomorphism between $\mathcal{G}(\mathcal{A})$ and $\mathfrak{C}(\mathbf{A}, \bar{e})$ sends f to $\bar{v}$. Further, given any state $\bar{v} \in \mathfrak{C}(\mathbf{A}, \bar{e})$, closing $\{\bar{v}\}$ under residuation will result in an automaton $\mathcal{A}_{\bar{v}}$. We say $\mathcal{A}$ is **Located at** $\bar{v} \in \mathfrak{C}(\mathbf{A}, \bar{e})$ iff the isomorphism sends $\mathcal{A}$ to $\mathcal{A}_{\bar{v}} \subseteq \mathfrak{C}(\mathbf{A}, \bar{e})$. Note that the location of a function or an automaton, and indeed whether a location exists or not, will depend on the choice of $\bar{e}$.

Unfortuately, Nekrashevych and Sidki leave entirely open the question of which $\bar{e}$ admit $\mathcal{A}$ as a subautomaton, and moreover where $\mathcal{A}$ is located if an embedding exists. This problem was first acknowledged in a paper of Okano [**?**] (TODO: double check that), and in 2018 Becker found an effective algorithm to compute $\mathbf{A}$ from $\mathcal{A}$ [**?**], solving part of the problem. In this paper we finish the job by fully characterizing the impact of $\bar{e}$ on the residuation structure of $\mathfrak{C}(\mathbf{A}, \bar{e})$. For a more detailed discussion of these linear algebraic methods and their origins, see [11, 12].

1.3 Principal Automata

To each abelian automaton we can associate a matrix as above, however each matrix can be associated to infinitely many automata. It was shown by Okano [13] that there is a distinguished automaton, now called the **Principal Automaton** $\mathfrak{A}$, associated to each matrix. $\mathfrak{A}(\mathbf{A})$ is defined to be $\mathfrak{A} = \mathcal{A}_{\bar{e}_1} \cup \mathcal{A}_{-\bar{e}_1} \subseteq \mathfrak{C}(\mathbf{A}, \bar{e}_1)$, though there is a longstanding conjecture that in most cases this is the same machine as $\mathcal{A}_{\bar{e}_1} \subseteq \mathfrak{C}(\mathbf{A}, \bar{e}_1)$. We will write $\mathfrak{A}$ when the matrix is clear from context.

We shall soon see that the same element of $\mathfrak{A}$ is located at $\bar{e}$ in $\mathfrak{C}(\mathbf{A}, \bar{e})$ for all $\bar{e}$, and so we call this group element $\delta \in \mathcal{G}(\mathfrak{A})$. Notice that for all $\bar{e}$, $\partial_0 \delta = \mathbf{A}(\bar{e} - \bar{e}) = \bar{0} = I$, and so $\partial_1 \delta = \gamma$. $\mathfrak{A}$ is clearly minimal in terms of state complexity, as its states are distinct group elements of $\mathfrak{C}(\mathbf{A}, \bar{e}_1)$ and therefore definitionally have different behavior. However it will be more interesting for us to note that $\mathfrak{A}$ is also minimal in the subgroup relation for nontrivial automata sharing its matrix.

By this we mean for every automaton $\mathcal{A}$ with associated matrix $\mathbf{A}$, $\mathcal{G}(\mathfrak{A}(\mathbf{A})) \leq \mathcal{G}(\mathcal{A})$, and so every function in $\mathfrak{A}$ is a $\mathbb{Z}$-linear combination of functions in $\mathcal{A}$. In particular, we see $\mathfrak{A}(\mathbf{A})$ is a subautomaton of $\mathcal{G}(\mathcal{A})$ for every $\mathcal{A}$ with matrix $\mathbf{A}$. While there are proofs of this claim which rely heavily on the ambient linear algebraic structure [13], we present here a difference construction which uses only the given automaton $\mathcal{A}$ to construct $\mathfrak{A}$. Thus every $s \in \mathfrak{A}$ is already in $\mathcal{G}(\mathcal{A})$, and the subgroup relation follows.

Theorem 1. *For each nontrivial $\mathcal{A}$ with associated matrix $\boldsymbol{A}$, $\mathcal{G}(\mathfrak{A}(\boldsymbol{A})) \leq \mathcal{G}(\mathcal{A})$.*

Proof. Let $\mathcal{A}$ be an abelian automaton with at least one odd state. Note that if $\mathcal{A}$ has no odd states, its group is trivial, so we may safely ignore it.

Put $\gamma = \partial_1 f - \partial_0 f$ for $f \in \mathcal{A}$ odd, and construct a new automaton by closing γ under residuation. Note that this can be done using only information contained in $\mathcal{A}$, since it is easy to check that:

$$\partial_0(f+g) = \begin{cases} \partial_0 f + \partial_1 g & \text{both odd} \\ \partial_0 f + \partial_0 g & \text{otherwise} \end{cases}$$

$$\partial_1(f+g) = \begin{cases} \partial_1 f + \partial_0 g & \text{both odd} \\ \partial_1 f + \partial_1 g & \text{otherwise} \end{cases}$$

$$\partial_0(-f) = -\partial_1 f$$

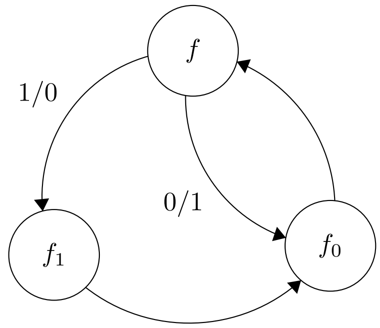
$$\partial_1(-f) = -\partial_0 f$$

Thus using the characterization by Sutner [17], that a state is odd iff it has distinct residuals, we can close γ under residuation using only information in $\mathcal{A}$. Since $\gamma \in \mathcal{G}(\mathcal{A})$ and $\mathcal{G}(\mathcal{A})$ is residuation closed, this entire closure is a subset of $\mathcal{G}(\mathcal{A})$.

Another theorem by Sutner [17] says that adding a state which residuates into an existing automaton does not change the group. To that end, the above closure generates the same group as the above closure with an additional state δ residuating into γ and a self loop I. This new machine is exactly $\mathcal{A}_{\bar{e}_1} \subseteq \mathfrak{C}(\mathbf{A}, \bar{e}_1)$. Any state in $\mathcal{A}_{e_1}$ is the negation of a state in $\mathcal{A}_{e_1}$, and so $\mathfrak{A}(\mathbf{A}) = \mathcal{A}_{\bar{e}_1} \cup \mathcal{A}_{-\bar{e}_1} \subseteq \mathcal{G}(\mathcal{A})$. Then $\mathcal{G}(\mathfrak{A}(\mathbf{A})) \leq \mathcal{G}(\mathcal{A})$, as desired. $\square$

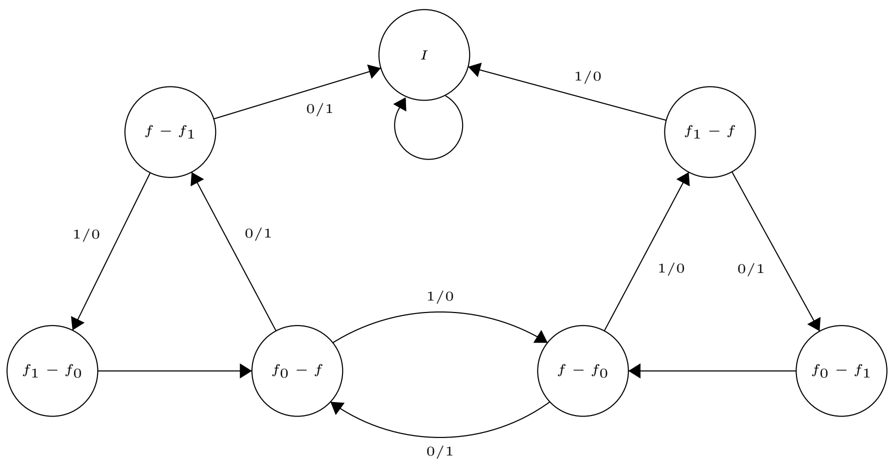
1.4 An Example

Consider the following machine, $\mathcal{A}_2^3$:



Here the unlabeled transitions both copy the input bit, however these have been omitted for cleanliness.

Then by letting $\gamma = \partial_1 f - \partial_0 f = f_1 - f_0$, and closing under residuation using the above algorithm, we construct the following machine (γ is shown at the bottom left):



It is easy to check that this is the principal machine for $\mathbf{A} = \begin{pmatrix} -1 & 1 \\ -\frac{1}{2} & 0 \end{pmatrix}$, and further that $\mathcal{A}_2^3$ as above is located at $\bar{e}_1 \in \mathfrak{C}\left(\mathbf{A}, \begin{pmatrix} 3 \\ 2 \end{pmatrix}\right)$

When running the algorithm in this case, we do not need to separately add $\pm\delta$ or the inverse machine. Here $\delta = f - f_1$, and the machine is already closed under negation. The Strongly Connected Component Conjecture predicts that this will be the case whenever $\mathbf{A}$ has characteristic polynomial other than $x^m - \frac{1}{2}$, which corresponds to the so called sausage automata. Unfortunately, however, this conjecture is yet unproven, and so in the above proof we had to explicitly add in these extra states.

2 Group Extensions

Going forward, $\mathcal{G} = \mathbb{Z}^m$ will denote $\mathcal{G}(\mathfrak{A})$ for some principal machine $\mathfrak{A}$.

$\mathcal{G}$ clearly admits representation as a $\mathbb{Z}[x]$ module where $x \cdot \bar{v} = \mathbf{A}^{-1}\bar{v}$, extended linearly. Further, since $\mathbf{A}$ has irreducible character so does $\mathbf{A}^{-1}$. Thus this module is cyclic, and is generated by $\bar{e}_1 = \delta$. (Note that since $\mathbf{A}$ sends $2\mathbb{Z} \oplus \mathbb{Z}^{m-1}$ to $\mathbb{Z}^m$, and therefore has multiples of $\frac{1}{2}$ in general, $\mathbf{A}^{-1}$ sends $\mathbb{Z}^m$ to $2\mathbb{Z} \oplus \mathbb{Z}^{m-1}$, and so has only integer entries).

Now for $p \in \mathbb{Z}[x]$ with odd constant term, we write $p \cdot \mathcal{G}$ in place of $\mathcal{G}(\mathfrak{C}(\mathbf{A}, p \cdot \bar{e}_1))$. That is to say, $p \cdot \mathcal{G}$ has as its states $\mathbb{Z}^m$ and as its odd residuations $\partial_0 \bar{v} = \mathbf{A}(\bar{v} - p \cdot \bar{e}_1)$, and $\partial_1 \bar{v} = \mathbf{A}(\bar{v} + p \cdot \bar{e}_1)$. Since this module is cyclic, every $\bar{v}$ arises as $p_{\bar{v}} \cdot \bar{e}_1$ where $p_{\bar{v}} = \bar{v}_0 + \bar{v}_1 x + \ldots + \bar{v}_{m-1} x^{m-1}$. We will only discuss polynomials p with an odd constant term, as this ensures $p \cdot \bar{e}_1$, our residuation vector, is odd.

We call $p \cdot \mathcal{G}$ the **Group Extension** of $\mathcal{G}$ by p. To justify this nomenclature, we first notice $\mathcal{G} \hookrightarrow p \cdot \mathcal{G}$ for all p by the homomorphism $\bar{v} \mapsto p \cdot \bar{v}$. Further, we recognize that if p is not a unit in $\mathrm{End}_{\mathcal{G}} \cong \mathbb{Z}^m/\chi^*$, this homomorphism is *not* surjective. That is to say $\mathcal{G}$ is a proper subgroup of $p \cdot \mathcal{G}$. Here, $\mathrm{End}_{\mathcal{G}}$ is the ring of all group endomorphisms, not necessarily preserving the residuation structure, and χ^* is the characteristic polynomial of $\mathbf{A}^{-1}$. This observation is true in more generality, as shown below.

Theorem 2. *If $rp = q$ in $\mathbb{Z}[x]$, then $p \cdot \mathcal{G} \hookrightarrow q \cdot \mathcal{G}$, with a canonical injection $\varphi_r : \bar{v} \mapsto r \cdot \bar{v}$. In particular, if r is a unit, then $p \cdot \mathcal{G} \cong q \cdot \mathcal{G}$.*

Proof. Let $rp = q$, $f \in p \cdot \mathcal{G}$ located at $\bar{v}$. Consider $f' \in q \cdot \mathcal{G}$ located at $r \cdot \bar{v}$.

First note f and f' have the same parity, since r has odd constant term, and so $\bar{v}$ and $r \cdot \bar{v}$ have the same parity. Now, consider the residuals of f and f'.

If f is even, then

$$\partial_0 f' = \mathbf{A}(r \cdot \bar{v}) = r \cdot \mathbf{A}\bar{v} = r \cdot \partial_0 f$$

If f is odd, then

$$\partial_0 f' = \mathbf{A}(r \cdot \bar{v} - q \cdot \bar{e}_1) = r \cdot \mathbf{A}(\bar{v} - p \cdot \bar{e}_1) = r \cdot \partial_0 f$$

A similar argument shows $\partial_1 f' = r \cdot \partial_1 f$

If r is a unit, then r^{-1} also has odd constant term (since $r * r^{-1} = 1$ has odd constant term) and so φ_r is an isomorphism with inverse $\varphi_{r^{-1}}$. □

3 Fractional Elements

As the previous proof shows, $p \cdot \bar{v} \in p \cdot \mathcal{G}$, computes exactly the same function as $\bar{v} \in \mathcal{G}$. However, most vectors cannot be written as $p \cdot \bar{v}$. What do they do as functions? We call such vectors (and their corresponding functions) **Fractional**, due to the following observation and theorem:

Consider $\bar{e}_1 \in 3 \cdot \mathcal{G}$. By the above theorem, $3\bar{e}_1 = \delta$, and so we should expect $\bar{e}_1$ to behave like "$\frac{1}{3}\delta$", and in fact it does.

In general, $\bar{v} \in p \cdot \mathcal{G}$ behaves like $p^{-1} \cdot \bar{v} \in \mathcal{G}$, (where p^{-1} comes from $\mathbb{Q}[x]$ and so $p^{-1} \cdot \bar{v} \in \mathbb{Q}^m$) and so Group Extensions give us access to fractions of functions from our base group $\mathcal{G}$.

We will consider $p^{-1} \cdot \mathbb{Z}^m = \{p^{-1} \cdot \bar{v} \mid \bar{v} \in \mathbb{Z}^m\}$ as a subgroup of $\mathbb{Q}^m$. Residuation in this setting is given by $\partial_0 \bar{v} = \mathbf{A}(\bar{v} - \bar{e}_1)$ and $\partial_1 \bar{v} = \mathbf{A}(\bar{v} + \bar{e}_1)$. Here, instead of scaling *up* our residuation vector, we scale *down* all of our other vectors. Then we have access to certain elements of $\mathbb{Q}^m$, which are exactly the factional elements as noted before. Now δ is always located at $\bar{e}_1$.

Morally, however, this is just a different way of looking at the group extension construction. We justify this with the following theorem:

Theorem 3. *For $p \in \mathbb{Z}[x]$ with odd constant term, $p^{-1} \cdot \mathbb{Z}^m \cong p \cdot \mathcal{G}$.*

Proof. Consider $\varphi : p^{-1} \cdot \mathbb{Z}^m \to p \cdot \mathcal{G}$ by $\varphi(p^{-1} \cdot \bar{v}) = \bar{v}$.
φ is clearly bijective, and is a homomorphism since:

$$\begin{aligned}
\varphi(p^{-1} \cdot \bar{v}_1 + p^{-1} \cdot \bar{v}_2) &= \varphi(p^{-1} \cdot (\bar{v}_1 + \bar{v}_2)) \\
&= \bar{v}_1 + \bar{v}_2 \\
&= \varphi(p^{-1} \cdot \bar{v}_1) + \varphi(p^{-1} \cdot \bar{v}_2)
\end{aligned}$$

Further, if $\bar{v}$ is even, then:

$$\begin{aligned}
\varphi(\partial_0(p^{-1} \cdot \bar{v})) &= \varphi(\mathbf{A}(p^{-1} \cdot \bar{v})) \\
&= \varphi(p^{-1} \cdot \mathbf{A}\bar{v}) \\
&= \mathbf{A}\bar{v} \\
&= \partial_0(\varphi(p^{-1} \cdot \bar{v}))
\end{aligned}$$

If $\bar{v}$ is odd, then:

$$\begin{aligned}
\varphi(\partial_0(p^{-1} \cdot \bar{v})) &= \varphi(\mathbf{A}(p^{-1} \cdot \bar{v} - \bar{e}_1)) \\
&= \varphi(p^{-1} \cdot \mathbf{A}(\bar{v} - p \cdot \bar{e}_1)) \\
&= \mathbf{A}(\bar{v} - p \cdot \bar{e}_1) \\
&= \partial_0(\varphi(p^{-1} \cdot \bar{v}))
\end{aligned}$$

The proof for ∂_1 is similar. □

Thus we can view functions in $p \cdot \mathcal{G}$ as fractions of functions in $\mathcal{G}$. It is a natural question to ask which fractions are attainable in this way.

Clearly, for any $f \in \mathcal{G}$, we can attain $\frac{1}{k} f$ for any odd k. Simply take $\bar{v} \in k \cdot \mathcal{G}$ for f located at $\bar{v}$. However, fractions with even denominator are, in general, unattainable. $2\frac{1}{2}\delta = \delta$ should be an odd function, but no function, when doubled, is odd.

4 Characterizing Automata

Since each automaton $\mathcal{A}$ is a subautomaton of some $\mathfrak{C}(\mathbf{A}, \bar{e})$, equivalently some $p \cdot \mathcal{G}$, there should be a minimal $\bar{e}$ (up to multiplication by units) which still has $\mathcal{A}$ as a subautomaton.

Notice that if we locate $\mathcal{A}$ at $\bar{e}_1 \in p \cdot \mathcal{G}$, then there can be no smaller polynomial q (in the division ordering) which also places $\mathcal{A}$ at an integral position. The following theorem shows this is always possible.

Theorem 4. *Every nontrivial abelian automaton $\mathcal{A}$ can be located at $\bar{e}_1$ in $p \cdot \mathcal{G}$ for some p.*

Proof. It is a theorem by Sutner [17] that every finite state abelian automaton residuates into a strongly connected component, and further that this component generates the same group as the entire machine. So we may, with no loss of generality, assume our machine is strongly connected (that is, every state except possibly I has a path to every other state).

Let f be an odd state in $\mathcal{A}$. Then at least one of $\partial_0 f$ and $\partial_1 f$ is not equal to f. So there is some nontrivial cycle from f to itself, which we can represent by a matrix equation relating $\bar{v}_f$, and $\bar{e}$. (Here $\bar{v}_f$ is where f will be located, and $\bar{e}$ will be the residuation vector). We can then rearrange this equation to obtain $p_1(\mathbf{A})\bar{v}_f = p_2(\mathbf{A})\bar{e}$.

$p_1, p_2 \in \mathbb{Z}[x]$, and $\mathbf{A}$ has irreducible character over $\mathbb{Z}$. It is well known that the eigenvalues of $p(\mathbf{A})$ are precisely $p(\lambda)$ where λ is an eigenvalue of $\mathbf{A}$, so $\mathbf{A}$'s invertibility implies the invertibility of both $p_1(\mathbf{A})$ and $p_2(\mathbf{A})$. Thus

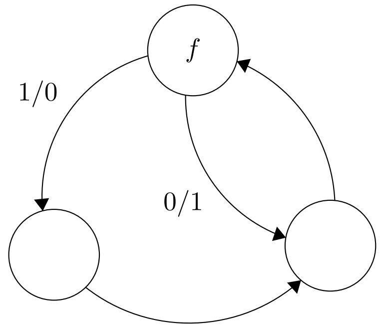
$$\bar{e} = p_2(\mathbf{A})^{-1} p_1(\mathbf{A}) \bar{v}_f$$

Choosing $\bar{v}_f = \bar{e}_1$ gives a value for the residuation vector $\bar{e}$, and (since $\mathcal{G}$ is cyclic as a $\mathbb{Z}[x]$ module) a value $\bar{e}$ induces a polynomial $p_{\bar{e}}$ such that $p_{\bar{e}} \cdot \bar{e}_1 = \bar{e}$. Then, by construction, $\mathcal{A}$ is a subautomaton of $p_e \cdot \mathcal{G}$, and is anchored with f at $\bar{e}_1$. As desired. □

For any automaton $\mathcal{A}$, we can now completely characterize in which $\mathfrak{C}(\mathbf{A}, \bar{e})$ it can be located, and at what vectors. Simply locate $\mathcal{A}$ at $\bar{e}_1 \in p \cdot \mathcal{G}$, and then to locate it at any odd vector $\bar{v}$, scale both sides by $p_{\bar{v}}$ to see $\mathcal{A}$ located at $\bar{v} \in p_{\bar{v}} p \cdot \mathcal{G}$. In the above proof, the choice of $\bar{v}_f = \bar{e}_1$ was arbitrary, and we can directly locate $\mathcal{A}$ at a different odd vector $\bar{v}'$ by setting $\bar{v}_f = \bar{v}'$. This will give the same result as locating it at $\bar{e}_1$ and then multiplying by $p_{\bar{v}'}$, again, by cyclicity. The same observation shows that, given some polynomial q (equivalently some vector $q \cdot \bar{e}_1$) $\mathcal{A}$ is located somewhere in $q \cdot \mathcal{G} = \mathfrak{C}(\mathbf{A}, q \cdot \bar{e}_1)$ if and only if $p \mid q$. Further, it will be located at exactly $p^{-1} q \cdot e_1$.

4.1 An Example

Recall the abelian automaton $\mathcal{A}_2^3$ from earlier in the paper:



Say we want to find $\bar{v}$ and $\bar{e}$ such that $\mathcal{A}_2^3$ is located at $\bar{v} \in \mathfrak{C}(\mathbf{A}, \bar{e})$.

Using the algorithm described by Becker [2] gives $\mathbf{A} = \begin{pmatrix} -1 & 1 \\ -\frac{1}{2} & 0 \end{pmatrix}$.

Then notice $\partial_0 \partial_0 f = f$. So $\mathbf{A}^2(\bar{v}_f - \bar{e}) = \bar{v}_f$, and $\mathbf{A}^2 \bar{v}_f - \bar{v}_f = \mathbf{A}^2 \bar{e}$. Thus

$$\bar{e} = \mathbf{A}^{-2}(\mathbf{A}^2 - I)\bar{v}_f$$

Choosing $\bar{v}_f = \bar{e}_1$ gives $\bar{e} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$.

Then $f = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \in (3 + 2x) \cdot \mathcal{G}$

4.2 Limiting Object

Since each $p \cdot \mathcal{G}$ can be viewed as $p^{-1} \cdot \mathbb{Z}^m \subseteq \mathbb{Q}^m$ with residuation vector $\bar{e}_1$, it is reasonable to consider the subgroup of $\mathbb{Q}^m$

$$\widetilde{\mathcal{G}} = \bigcup_p p^{-1} \cdot \mathbb{Z}^m$$

(Recall we only include polynomials with odd constant term in this union)

Notice that this group is universal, in the sense that it contains as subgroups each $p \cdot \mathcal{G}$. Further, it concretely shows the relationships between the various automata. In this setting, we see exactly why automata show up in multiple group extensions, and why the division ordering of polynomials is the characterizing factor. $p \cdot \mathcal{G}$ is an approximation of $\widetilde{\mathcal{G}}$, where we scale up by a factor of p and take only the integral vectors (residuation is necessarily scaled up to match). In this structure, then, there is no unnecessary duplication of the location of automata, and there is no extra parameter $\bar{e}$.

5 Conclusion

We have shown that the residuation vector $\bar{e}$ corresponds to how fine an approximation of $\widetilde{\mathcal{G}}$ one wants. This is because each $\mathfrak{C}(\mathbf{A}, \bar{e})$ corresponds to $p_{\bar{e}} \cdot \mathcal{G}$, with progressively larger $\bar{e}$ corresponding to progressively more complicated fractional elements, which approximate $\widetilde{\mathcal{G}}$. Thus, the parameter really provides a way of interacting with these elements living in $\mathbb{Q}^m$ as though they were in $\mathbb{Z}^m$, and so by computing in $\widetilde{\mathcal{G}}$ (or a suitably large approximation) directly, we can remove the need for this parameter.

Further, the existence of the universal object $\widetilde{\mathcal{G}}$ sheds new light on the connection between affine tiles [9, 10] and abelian automata noted by Sutner [17]. Indeed it is easy to see that in $\widetilde{\mathcal{G}}$ every strongly connected component (and thus every subautomaton of interest) has each vector in the attractor of the iterated function system given by the residuation functions $\{\bar{v} \mapsto \mathbf{A}\bar{v}, \bar{v} \mapsto \mathbf{A}(\bar{v} \pm \bar{e}_1)\}$. Thus, in particular, the size of the principal machine is bounded by the number of integral points in this attractor. Even in $\mathbb{Z}^2$, however, there are examples where this bound is not tight.

The relation between automata and polynomials discussed in this paper also provides a new take on a proof technique for the longstanding Strongly Connected Component Conjecture. This conjecture asserts that principal machines $\mathfrak{A}$ have only one strongly connected component (plus the self looping identity state) whenever their matrix has a characteristic polynomial that is *not* of the form $x^n + \frac{1}{2}$. The new way of looking at residuation vectors allows us to rewrite the residual functions as $\partial_i \bar{v} = \mathbf{A}(\bar{v} - (-1)^i \delta)$ for $\bar{v}$ odd. It is easy to see, then, that the following polynomials correspond to paths ending in δ, since they undo residuation:

$$\begin{aligned} P_\epsilon(x) &= 1 \\ P_{w0}(x) &= xP_w(x) + 0 \\ P_{w1}(x) &= xP_w(x) + 1 \\ P_{w\bar{1}}(x) &= xP_w(x) - 1 \end{aligned}$$

Sutner made a similar observation, and described Path Polynomials [17] which allow us to reason about the existence of directed paths between states in an automaton by purely algebraic means. However, these polynomials are clunky and not always defined, since they correspond to paths *starting* at δ, and so $P'_{w0} \cdot \delta$ is only well defined if $P'_w \cdot \delta$ is even (and P'_{w1} and $P'_{w\bar{1}}$ are only well defined if $P_w \cdot \delta$ is odd). Since the polynomials defined above move *backwards* along transitions instead of forwards, they are always well defined.

The existence of a path polynomial p which is congruent to $-1 \bmod \chi^*$ then shows the existence of a path from $-\delta$ to δ. Then to prove the SCC conjecture, it suffices to prove that whenever $\mathbf{A}$ does not have characteristic $x^n + \frac{1}{2}$ there is a polynomial $p \in \{-1, 0, 1\}[x]$ which is congruent to $-1 \bmod \chi^*$. Efforts are underway to use this method to actually prove the conjecture.

Acknowledgements

This paper would not exist without the advice of my advisor Klaus Sutner. There aren't enough thanks for the hours of conversation I enjoyed.

References

1. Bárány, V.: Automatic presentations of infinite structures (2007)
2. Becker, T.: Representations and complexity of abelian automaton groups. Senior Thesis, CMU. http://tjbecker.me/files/thesis.pdf (may 2018)
3. Bondarenko, I., Grigorchuk, R., Kravchenko, R., Muntyan, Y., Nekrashevych, V., Savchuk, D., Šunić, Z.: Groups generated by 3-state automata over a 2-letter alphabet. ii. Journal of Mathematical Sciences **156**(1), 187–208 (Jan 2009), https://doi.org/10.1007/s10958-008-9262-5
4. Grigorchuk, R.R., Nekrashevich, V.V., Sushchanski, V.I.: Automata, dynamical systems and groups. Proc. Steklov Institute of Math. **231**, 128–203 (2000)
5. Grigorchuk, R.R., Zuk, A.: The lamplighter group as a group generated by a a 2-state automaton, and its spectrum. Geom. Dedicata **87**, 209–244 (2001)
6. Grigorchuk, R.: Milnor's problem on the growth of groups and its consequences (2011)
7. Gupta, N., Sidki, S.: On the burnside problem for periodic groups. Mathematische Zeitschrift **182**(3), 385–388 (Sep 1983). https://doi.org/10.1007/BF01179757, https://doi.org/10.1007/BF01179757
8. Holcombe, W.M.L.: Algebraic Automata Theory. Cambridge University Press (1982)
9. Lagarias, J.C., Wang, Y.: Self-affine tiles in $\mathbb{R}^n$. Adv. Math. **121**, 21–49 (1996)
10. Lagarias, J.C., Wang, Y.: Integral self-affine tiles in $\mathbb{R}^n$ ii. lattice tilings. J. Fourier Anal. Appl. **3**(1), 83–102 (1997)
11. Nekrashevych, V.: Self-Similar Groups, Math. Surveys and Monographs, vol. 117. AMS (2005)
12. Nekrashevych, V., Sidki, S.: Automorphisms of the binary tree: state-closed subgroups and dynamics of 1/2-endomorphisms. Cambridge University Press (2004)
13. Okano, T.: Invertible binary transducers and automorphisms of the binary tree. MS Thesis, CMU (may 2015)
14. Sakarovitch, J.: Elements of Automata Theory. Cambridge University Press (2009)
15. Sidki, S.: Automorphisms of one-rooted trees: Growth, circuit structure, and acyclicity. J. Math. Sciences **100**(1), 1925–1943 (2000)
16. Skrzypczak, M.: Descriptive Set Theoretic Methods in Automata Theory. Springer-Verlag Berlin Heidelberg (2016)
17. Sutner, K.: Abelian Invertible Automata