Chapter 9

An Ochmański-type theorem

Kleene's Theorem on recognizable languages of finite words has been generalized in several directions, e.g. to formal power series [Sch61] and to infinite words [Büc60]. More recently, rational monoids were investigated [Sak87], in which the recognizable languages coincide with the rational ones. Building on results from [CP85, CM88, Mét86], a complete characterization of the recognizable languages in a trace monoid by c-rational sets was obtained in [Och85]. A further generalization of Kleene's and Ochmański's results to concurrency monoids was given in [Dro95]. In this chapter, we derive such a result for divisibility monoids. The proofs by Ochmański [Och85] and by Droste [Dro95] rely on the *internal* structure of the elements of the monoids. Here, we do not use the internal representation of the monoid elements, but algebraic properties of the monoid itself. The results presented in this chapter were obtained together with Manfred Droste. They appeared in [DK99] and the presentation follows [DK00].

9.1 Complete grids and the rank

In trace theory, the generalized Levi Lemma (cf. [DM97]) plays an important role. It was extended to concurrency monoids in [Dro95]. Here, we develop a further generalization to divisibility monoids using complete grids. This enables us to obtain the concept of the "rank" of a language for these monoids, similar to the one given by Hashigushi [Has91] for trace monoids. Let M be a divisibility monoid and $x, y \in M$. Recall that $r_x(y) = y \uparrow x$. Sometimes (for instance in the following definition), it is more convenient to use this notation for the functions res_u, too. Therefore, we define $v \uparrow u := \mathrm{res}_u(v)$ whenever the latter is defined for $u, v \in T^\star$.

Definition 9.1.1 For $0 \leq i \leq j \leq n$ let $x^i_j, y^j_i \in M$ ($\in T^\star$, respectively). The tuple $(x^i_j, y^j_i)_{0\leq i\leq j\leq n}$ is a *complete grid* in M (in $T^\star$, respectively) provided the following holds for any $0 \leq i < j \leq n$:

$$x^i_j \mathbin{\mathbf{co}} y^{j-1}_i,\ \ x^i_j \uparrow y^{j-1}_i = x^{i+1}_j,\ \text{ and }\ y^{j-1}_i \uparrow x^i_j = y^j_i.$$

A complete grid can be depicted as in Figure 9.1. There, edges depict elements from M ($T^\star$, resp.) and an angle denotes that the two edges correspond to complementary elements. Note that in any of the small squares in Figure 9.1, the lower left corner is marked by an angle. This indicates that $x^i_j y^j_i = y^{j-1}_i x^{i+1}_j$ because of $x^i_j y^j_i = x^i_j(y^{j-1}_i \uparrow x^i_j) = x^i_j \vee y^{j-1}_i = y^{j-1}_i(x^i_j \uparrow y^{j-1}_i) = y^{j-1}_i x^{i+1}_j$. By Lemma 7.2.5 (1)-(3), for any rectangle in the grid $(x^i_j, y^j_i)_{0\le i\le j\le n}$ the bottom and the left side are complementary and their residuum is the top (the right) side, respectively. By induction, it is easy to show that

$$(x^0_1 x^0_2 \ldots x^0_n) \cdot (y^n_0 y^n_1 \ldots y^n_n) = (x^0_0 y^0_0)(x^1_1 y^1_1)(x^2_2 y^2_2) \ldots (x^n_n y^n_n).$$

The right hand side of this equation is the diagonal border of the grid in Figure 9.1.

Let $(M, \cdot, 1)$ be a divisibility monoid and $(x^i_j, y^j_i)_{0\le i\le j\le n}$ a complete grid in M or in $T^\star$. For a sequence $0 = i_0 < i_1 < \cdots < i_{m+1} = n$, we construct a subgrid $(a^k_l, b^l_k)_{0\le k\le l\le m}$ as follows: Define $a^k_l := x^{i_k}_{i_l} x^{i_k}_{i_l+1} x^{i_k}_{i_l+2} \ldots x^{i_k}_{i_{l+1}-1}$ and $b^l_k := y^{i_{l+1}-1}_{i_k} y^{i_{l+1}-1}_{i_k+1} \ldots y^{i_{l+1}-1}_{i_{k+1}-1}$ (for $m = 4$ and $\vec{\imath} = (0, 1, 5, 7, 9)$, this grid is marked by thick lines in Figure 9.1). Then $(a^k_l, b^l_k)_{0\le k\le l\le m}$ is a complete grid in M or $T^\star$. We call it the *subgrid generated by the sequence* $(i_k)_{0\le k\le m}$.

Let $(x^i_j, y^j_i)_{0\le i\le j\le n}$ be a complete grid in $T^\star$. Then it is immediate that $(\text{nat}(x^i_j), \text{nat}(y^j_i))_{0\le i\le j\le n}$ is a complete grid in M. The following lemma deals with the converse implication. More precisely, let a complete grid in M be given and suppose that $u^0_j, v^n_j \in T^\star$ are representatives of monoid elements at the left and the upper border of the complete grid. Then the lemma states that this tuple of words can be extended to a complete grid in $T^\star$ that is compatible with the complete grid in M we started with.

Lemma 9.1.2 *Let $(x^i_j, y^j_i)_{0\le i\le j\le n}$ be a complete grid in M and let u^0_j, v^n_j be words from $T^\star$ that satisfy $\text{nat}(u^0_j) = x^0_j$ and $\text{nat}(v^j_n) = y^j_n$ for $0 \le j \le n$. Then there exists a complete grid $(u^i_j, v^j_i)_{0\le i\le j\le n}$ with $\text{nat}(u^i_j) = x^i_j$ and $\text{nat}(v^j_i) = y^j_i$ for $0 \le i \le j \le n$.*

Proof. For $1 \le i \le j \le n$ let $u^i_j := \text{res}_{y^{j-1}_i}(u^{i-1}_j)$. Using Lemma 7.2.7, one can check that $\text{nat}(u^i_j) = x^i_j$ and that therefore $u^{i-1}_j \in \text{dom}(\text{res}_{y^{j-1}_i})$. To construct the elements v^j_i, we use Lemma 7.2.9: Let v^j_i be the unique word over T such that $v^j_i \uparrow u^i_{j+1} = v^{j+1}_i$. Then $\text{nat}(v^j_i) = y^j_i$ is immediate since $y^j_i \in M$ is the unique complement of x^i_{j+1} in the distributive lattice $[x^i_{j+1}, x^i_{j+1} \vee y^j_i]$. It is clear that $(u^i_j, v^j_i)_{0\le i\le j\le n}$ is a complete grid because of $\text{nat}(u^i_j) = x^i_j$ and $\text{nat}(v^j_i) = y^j_i$ for $0 \le i \le j \le n$. $\square$

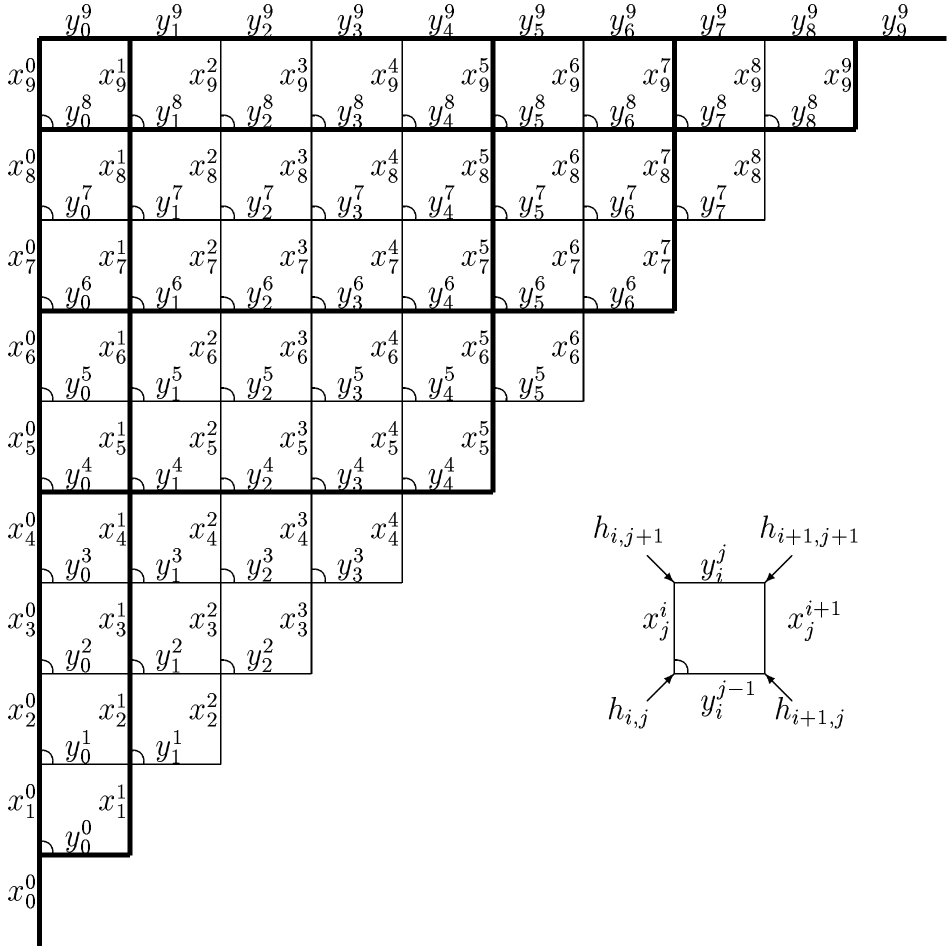

Figure 9.1: A complete grid

The elements $h_{i,j}$ from M marking the corners of the small square in Figure 9.1 are defined by

$$h_{i,j} := (x_0^0 x_1^0 \dots x_{j-1}^0) \cdot (y_0^{j-1} y_1^{j-1} \dots y_{i-1}^{j-1}).$$

Then $h_{i,j+1} = h_{i,j} \cdot x_j^i$ and $h_{i+1,j} = h_{i,j} \cdot y_i^{j-1}$. Hence by Lemma 7.1.1, $h_{i,j} = h_{i,j+1} \wedge h_{i+1,j}$ and $h_{i+1,j+1} = h_{i,j+1} \vee h_{i+1,j}$. These relations will be used in the proof of the following lemma. It can be read as the converse of the equality $(x_0^0 x_1^0 \dots x_n^0) \cdot (y_0^n y_1^n \dots y_n^n) = (x_0^0 y_0^0)(x_1^1 y_1^1)(x_2^2 y_2^2) \dots (x_n^n y_n^n)$: Whenever the product of two elements of M equals the product of finitely many elements, there exists a corresponding complete grid. This lemma is the announced extension of Levi's Lemma from trace theory into our setting of divisibility monoids.

Lemma 9.1.3 *Let $z_0, z_1, \ldots, z_n, x, y \in M$ with $x \cdot y = z_0 \cdot z_1 \cdots z_n$. Then there exists a complete grid $(x^i_j, y^j_i)_{0 \leq i \leq j \leq n}$ in M such that*

- $x = x^0_0 x^0_1 \ldots x^0_n$,
- $y = y^n_0 y^n_1 \ldots y^n_n$, *and*
- $z_i = x^i_i y^i_i$ *for* $i = 0, 1, \ldots, n$.

Proof. Let $h_{0,j} := x \wedge z_0 z_1 \ldots z_{j-1}$ and $h_{j,j} = z_0 \cdot z_1 \ldots z_{j-1}$ (for $0 \leq j \leq n+1$). Note that $h_{0,j} \leq h_{0,j+1}$ and $h_{j,j} \leq h_{j+1,j+1}$. Furthermore, let $h_{i,j} := h_{0,j} \vee h_{i,i}$ for $0 < i < j \leq n+1$ (this supremum exists since $\{h_{0,j}, h_{i,i}\}$ is bounded by xy). Then $h_{i+1,j+1} = h_{0,j+1} \vee h_{i+1,i+1} = h_{0,j+1} \vee h_{0,j} \vee h_{i+1,i+1} \vee h_{i,i} = h_{i,j+1} \vee h_{i+1,j}$.

Now let $0 \leq i < j \leq n+1$. Then $h_{0,j+1} \wedge h_{i+1,i+1} = x \wedge z_0 z_1 \cdots z_j \wedge z_0 z_1 \cdots z_i = h_{0,i+1} \leq h_{0,j}$, and so

$$\begin{aligned} h_{i,j+1} \wedge h_{i+1,j} &= (h_{0,j+1} \vee h_{i,i}) \wedge (h_{0,j} \vee h_{i+1,i+1}) \\ &= (h_{0,j+1} \wedge h_{0,j}) \vee (h_{0,j+1} \wedge h_{i+1,i+1}) \\ &\quad \vee (h_{i,i} \wedge h_{0,j}) \vee (h_{i,i} \wedge h_{i+1,i+1}) \\ &= h_{0,j} \vee h_{0,i+1} \vee (h_{i,i} \wedge h_{0,j}) \vee h_{i,i} \\ &= h_{0,j} \vee h_{i,i} \qquad \text{(since } h_{0,i+1} \leq h_{0,j} \text{ and } h_{i,i} \wedge h_{0,j} \leq h_{i,i}) \\ &= h_{i,j}. \end{aligned}$$

Now the elements x^i_j and y^j_i for $0 \leq i \leq j \leq n$ of the complete grid are the monoid elements uniquely determined by $h_{i,j+1} = h_{i,j} \cdot x^i_j$ and $h_{i+1,j+1} = h_{i,j+1} \cdot y^j_i$.

Since $h_{i,j+1} \wedge h_{i+1,j} = h_{i,j}$ and $h_{i+1,j+1} = h_{i,j+1} \vee h_{i+1,j}$, Lemma 7.1.1 implies $x^i_j \wedge y^{j-1}_i = 1$ and $x^i_j \vee y^{j-1}_i = x^i_j y^j_i = y^{j-1}_i x^{i+1}_j$. Hence $x^i_j \,\mathbf{co}\, y^{j-1}_i$, $x^i_j \uparrow y^{j-1}_i = x^{i+1}_j$ and $y^{j-1}_i \uparrow x^i_j = y^j_i$. Thus, we showed that $(x^i_j, y^j_i)_{0 \leq i \leq j \leq n}$ is a complete grid in M.

Note that $x^0_0 x^0_1 \ldots x^0_n = h_{0,n+1} = x \wedge z_0 z_1 \ldots z_n = x \wedge xy = x$. Furthermore, $h_{j,j} \cdot z_j = h_{j+1,j+1} = h_{j,j} \cdot x^j_j \cdot y^j_j$ implies $z_j = x^j_j \cdot y^j_j$ since M is cancellative. Hence we have

$$\begin{aligned} x(y^n_0 y^n_1 \cdots y^n_n) &= (x^0_0 x^0_1 \ldots x^0_n)(y^n_0 y^n_1 \cdots y^n_n) \\ &= (x^0_0 y^0_0)(x^1_1 y^1_1)(x^2_2 y^2_2) \ldots (x^n_n y^n_n) \\ &= z_0 z_1 \ldots z_n = xy. \end{aligned}$$

This equality implies $y = y^n_0 y^n_1 \ldots y^n_n$. □

It is reasonable that the elements y^j_i in a complete grid do not completely determine the elements x^i_j. The following lemma describes the freedom we have in choosing these elements: as long as we keep the residuum functions of the elements in the first column, we can complete the complete grid.

Lemma 9.1.4 *Let* $(x^i_j, y^j_i)_{0\leq i\leq j\leq n}$ *be a complete grid in* M*, and, for* $0 \leq j \leq n$*, let* $w^0_j \in M$ *with* $r_{x^0_j} = r_{w^0_j}$*. Then there exists a complete grid* $(w^i_j, y^j_i)_{0\leq i\leq j\leq n}$ *in* M*.*

Proof. By Lemma 7.2.5(1), $x^0_j \,\mathbf{co}\, y^{j-1}_0 y^{j-1}_1 \dots y^{j-1}_{j-1}$ for any $0 \leq j \leq n$. Since the residuum functions of x^0_j and w^0_j coincide, this implies $w^0_j \,\mathbf{co}\, y^{j-1}_0 y^{j-1}_1 \dots y^{j-1}_{i-1}$ for $0 \leq i \leq j \leq n$. Hence $w^i_j := w^0_j \uparrow (y^{j-1}_0 y^{j-1}_1 \dots y^{j-1}_{i-1})$ is defined. Using Lemma 7.2.5(1), we get $w^i_j \,\mathbf{co}\, y^{j-1}_i$. By Lemma 7.2.5(2), $w^{i+1}_j = w^i_j \uparrow y^{j-1}_i$. Since $x^i_j = x^0_j \uparrow y^{j-1}_0 y^{j-1}_1 \dots y^{j-1}_{i-1}$ and $r_{x^0_j} = r_{w^0_j}$, Lemma 7.2.5(4) implies $r_{x^i_j} = r_{w^i_j}$. Hence $y^j_i = r_{x^i_j}(y^{j-1}_i) = r_{w^i_j}(y^{j-1}_i) = y^{j-1}_i \uparrow w^i_j$. □

Recall that $\mathrm{nat}(\mathrm{res}_v(u)) = r_{\mathrm{nat}(v)}(\mathrm{nat}(u))$ for any words u, v by Lemma 7.2.7. Using Lemma 9.1.2, one gets as a direct consequence

Corollary 9.1.5 *Let* $(u^i_j, v^j_i)_{0\leq i\leq j\leq n}$ *be a complete grid in* $T^\star$*, and, for* $0 \leq j \leq n$*, let* $w^0_j \in T^\star$ *with* $\mathrm{res}_{u^0_j} = \mathrm{res}_{w^0_j}$*. Then there exists a complete grid* $(w^i_j, v^j_i)_{0\leq i\leq j\leq n}$ *in* $T^\star$*.*

Similarly as above, a direct consequence of Lemma 9.1.3 is the existence of complete grids in $T^\star$:

Corollary 9.1.6 *Let* $z_0, z_1, \dots, z_n, u, v \in T^\star$ *with* $\mathrm{nat}(uv) = \mathrm{nat}(z_1 z_2 \dots z_n)$*. Then there exists a complete grid* $(u^i_j, v^j_i)_{0\leq i\leq j\leq n}$ *in* $T^\star$ *such that*

1. $\mathrm{nat}(u) = \mathrm{nat}(u^0_0 u^0_1 \dots u^0_n)$,
2. $\mathrm{nat}(v) = \mathrm{nat}(v^n_0 v^n_1 \dots v^n_n)$, *and*
3. $\mathrm{nat}(z_i) = \mathrm{nat}(u^i_i v^i_i)$ *for* $i = 0, 1, \dots, n$.

Note that the equations in the corollary above do not hold in the free monoid $T^\star$ but only in the divisibility monoid M. But if the words z_i are actually from $T \cup \{\varepsilon\}$ (i.e. their length is at most 1), we can replace the third statement by

3'. $z_i = u^i_i v^i_i$ for $i = 0, 1, \dots, n$.

Now we can introduce the notion of rank in the present context. For traces, it was defined and shown to be very useful by Hashigushi [Has91], cf. [DR95, Ch. 6] and [DM97]. Recall that $\mathrm{nat}(X) = \{\mathrm{nat}(w) \mid w \in X\}$ for any set $X \subseteq T^\star$ of words over T.

Definition 9.1.7 Let $u, v \in T^\star$ and $X \subseteq T^\star$ such that $\mathrm{nat}(uv) \in \mathrm{nat}(X)$. Let $\mathrm{rk}(u, v, X)$, the *rank of* u *and* v *relative to* X, denote the minimal integer n such that there exists a complete grid $(u_j^i, v_i^j)_{0\leq i\leq j\leq n}$ in $T^\star$ with

1. $\mathrm{nat}(u) = \mathrm{nat}(u_0^0 u_1^0 \dots u_n^0)$,
2. $\mathrm{nat}(v) = \mathrm{nat}(v_0^n v_1^n \dots v_n^n)$,
3. $u_0^0 v_0^0 u_1^1 v_1^1 \dots u_n^n v_n^n \in X$.

Let $u, v \in T^\star$ and $X \subseteq T^\star$ such that $\mathrm{nat}(uv) \in \mathrm{nat}(X)$. Then there exists $z \in X$ with $\mathrm{nat}(z) = \mathrm{nat}(uv)$. Let $z = z_1 z_2 \dots z_n$ with $z_i \in T$. Then $n = |uv|$ and by Corollary 9.1.6 (with 3' instead of 3) we find an appropriate complete grid. Hence $\mathrm{rk}(u, v, X) \leq |uv|$. If not only $\mathrm{nat}(uv) \in \mathrm{nat}(X)$ but even $uv \in X$ we can choose $n = 0$, $u_0^0 = u$ and $v_0^0 = v$ and obtain $\mathrm{rk}(u, v, X) = 0$. We define the *rank* $\mathrm{rk}(X)$ *of* X by

$$\mathrm{rk}(X) := \sup\{\mathrm{rk}(u, v, X) \mid u, v \in T^\star, \mathrm{nat}(uv) \in \mathrm{nat}(X)\} \in \mathbb{N} \cup \{\infty\}.$$

A word language $X \subseteq T^\star$ is *closed* if $\mathrm{nat}(u) \in \mathrm{nat}(X)$ implies $u \in X$ for any $u \in T^\star$. Since $\mathrm{rk}(u, v, X) = 0$ whenever $uv \in X$, the rank of a closed language equals 0.

Theorem 9.1.8 *Let* $(M, \cdot, 1)$ *be a divisibility monoid with finitely many residuum functions (i.e. the monoid* $\mathbb{R}_M$ *is finite). Let* $X \subseteq T^\star$ *be a recognizable language of finite rank. Then* $\mathrm{nat}(X)$ *is recognizable in* M.

Proof. Let $n = \mathrm{rk}(X) \in \mathbb{N}$ be the rank of X. Since X is recognizable, there is a finite monoid S and a homomorphism $\eta : T^\star \to S$ that recognizes X. Since the mapping $u \mapsto \mathrm{res}_u$ is an antihomomorphism from $T^\star$ into the finite monoid $(\mathbb{D}_M, \circ, \mathrm{res}_\varepsilon)$, we may assume that $\eta(u) = \eta(v)$ implies $\mathrm{res}_u = \mathrm{res}_v$.

For $x \in M$, let $R(x)$ denote the set

$$\{(\eta d(u_0), \eta d(u_1) \dots \eta d(u_n))_{d\in\mathbb{D}_M} \mid u_0, u_1, \dots, u_n \in T^\star \text{ and } x = \mathrm{nat}(u_0 u_1 \dots u_n)\}.$$

Then $R(x)$ is a subset of $(S^{n+1})^{|\mathbb{D}_M|}$. Since $\mathbb{D}_M$ and S are finite, there are only finitely many sets $R(x)$. Once we will have shown

$$R(x) = R(z) \Rightarrow x^{-1}\,\mathrm{nat}(X) = z^{-1}\,\mathrm{nat}(X),$$

we thus have that $\{x^{-1}\,\mathrm{nat}(X) \mid x \in M\}$ is finite. Hence $\mathrm{nat}(X)$ is recognizable.

So let $R(x) = R(z)$ and let $y \in x^{-1}\,\mathrm{nat}(X)$, i.e. $xy \in \mathrm{nat}(X)$. Since $\mathrm{rk}(X) = n$, there exists a complete grid $(u_j^i, v_i^j)_{0\leq i\leq j\leq n}$ in $T^\star$ such that

- $x = \mathrm{nat}(u_0^0 u_1^0 \dots u_n^0)$,

- $y = \mathrm{nat}(v_0^n v_1^n \dots v_n^n)$, and
- $u_0^0 v_0^0 u_1^1 v_1^1 \dots u_n^n v_n^n \in X$.

Then $(\eta d(u_0^0), \eta d(u_1^0) \dots \eta d(u_n^0))_{d \in \mathbb{D}_M} \in R(x) = R(z)$. Hence there exist words $w_j^0 \in T^\star$ with

(1) $\eta d(w_j^0) = \eta d(u_j^0)$ for each $0 \leq j \leq n$ and $d \in \mathbb{D}_M$, and

(2) $z = \mathrm{nat}(w_0^0 w_1^0 \dots w_n^0)$.

In (1), consider $d = \mathrm{res}_\varepsilon$ which equals the identity on $T^\star$. Then $\eta(w_j^0) = \eta(u_j^0)$ and therefore (by our assumption on η) $\mathrm{res}_{w_j^0} = \mathrm{res}_{u_j^0}$. Hence we can apply Corollary 9.1.5 and obtain a complete grid $(w_j^i, v_i^j)_{0 \leq i \leq j \leq n}$ in $T^\star$. Now consider $d = \mathrm{res}_{v_0^{j-1} v_1^{j-1} \dots v_{j-1}^{j-1}} \in \mathbb{D}_M$. Note that $w_j^j = w_j^0 \uparrow (v_0^{j-1} v_1^{j-1} \dots v_{j-1}^{j-1}) = d(w_j^0)$. Hence $\eta(w_j^j) = \eta d(w_j^0) = \eta d(u_j^0) = \eta(u_j^j)$. Now we can conclude

$$\begin{aligned}
\eta(w_0^0 v_0^0 w_1^1 v_1^1 \dots w_n^n v_n^n) &= \eta(w_0^0)\eta(v_0^0)\eta(w_1^1)\eta(v_1^1) \dots \eta(w_n^n)\eta(v_n^n) \\
&= \eta(u_0^0)\eta(v_0^0)\eta(u_1^1)\eta(v_1^1) \dots \eta(u_n^n)\eta(v_n^n) \\
&= \eta(u_0^0 v_0^0 u_1^1 v_1^1 \dots u_n^n v_n^n) \in \eta(X).
\end{aligned}$$

Hence $w_0^0 v_0^0 w_1^1 v_1^1 \dots w_n^n v_n^n \in X$. Since $(\mathrm{nat}(w_j^i), \mathrm{nat}(v_i^j))_{0 \leq i \leq j \leq n}$ is a complete grid in M, we obtain $zy = \mathrm{nat}(w_0^0 v_0^0 w_1^1 v_1^1 \dots w_n^n v_n^n) \in \mathrm{nat}(X)$. Hence $y \in z^{-1}\,\mathrm{nat}(X)$ and therefore $x^{-1}\,\mathrm{nat}(X) = z^{-1}\,\mathrm{nat}(X)$ as claimed above. □

9.2 From c-rational to recognizable languages

In this section, we prove closure properties of the set of recognizable languages in a divisibility monoid. These closure properties correspond to c-rational languages that we introduce first:

Let $(M, \cdot, 1)$ be a divisibility monoid. An element $x \in M$ is *connected* if the distributive lattice $\downarrow x$ does not contain any pair of complementary elements. In other words, there are *no* complementary $y, z \in M \setminus \{1\}$ such that $x = y \vee z = y r_y(z)$. A set $L \subseteq M$ is *connected* if all of its elements are connected; a language $X \subseteq T^\star$ is *connected* if $\mathrm{nat}(X) \subseteq M$ is connected.

Let t be a trace over the dependence alphabet (Σ, D). In trace theory, this trace is called "connected" if the letters occurring in it induce a connected subgraph of (Σ, D). One can easily check that this is the case iff t is not the supremum of two complementary traces, i.e. iff t is connected in the sense defined above. For rational trace languages, to be recognizable it suffices that the iteration is applied to connected languages, only. In other words, there is a subset of the trace monoid C (the connected traces) such that the iteration is applied to languages included in C, only. Already for concurrency monoids (cf. [Dro95, Dro96]), it is not sufficient to restrict to connected languages. But there, one still has finitely many pairwise disjoint sets C_q such that the iteration can be restricted to subsets of C_q. For divisibility monoids, we did not find such sets in general (for labeled divisibility monoids, they exist – see below). Therefore, we impose an internal condition on those languages that we want to iterate:

A language $X \subseteq T^\star$ is *residually closed* if for any $u \in X$ and $v \in T^\star$ with $u \mathbin{\mathbf{co}} v$ the following holds:

$$v \in X \iff \mathrm{res}_u(v) \in X.$$

Thus X is residually closed if it is closed under the application of res_u and res_u^{-1} for elements u of X. Note that this need not hold for all $u \in T^\star$. A language $L \subseteq M$ is *residually closed* iff $\{w \in T^\star \mid \mathrm{nat}(w) \in L\}$ is residually closed.

Now we define c-rational languages: The set of *c-rational sets* in a divisibility monoid M is the least class $\mathfrak{C} \subseteq 2^M$ such that

- all finite subsets of M belong to $\mathfrak{C}$,
- $X \cdot Y$ and $X \cup Y$ belong to $\mathfrak{C}$ whenever $X, Y \in \mathfrak{C}$, and
- $\langle X \rangle$ belongs to $\mathfrak{C}$ whenever $X \in \mathfrak{C}$ is connected and residually closed.

Now we are going to show that the set of recognizable languages is closed under multiplication.

Lemma 9.2.1 *Let $(M, \cdot, 1)$ be a divisibility monoid and $X, Y \subseteq T^\star$ be closed. Then* $\mathrm{rk}(XY) \leq 1$.

Proof. Let $u, v \in T^\star$ with $\mathrm{nat}(uv) \in \mathrm{nat}(XY)$. Then there exist $z_0 \in X$ and $z_1 \in Y$ such that $\mathrm{nat}(uv) = \mathrm{nat}(z_0 z_1)$. By Corollary 9.1.6, there exists a complete grid $(u^i_j, v^j_i)_{0 \leq i \leq j \leq 1}$ in $T^\star$ such that

$$\begin{array}{ll} \mathrm{nat}(u) = \mathrm{nat}(u^0_0 u^0_1), & \mathrm{nat}(v) = \mathrm{nat}(v^1_0 v^1_1), \\ \mathrm{nat}(z_0) = \mathrm{nat}(u^0_0 v^0_0), \text{ and} & \mathrm{nat}(z_1) = \mathrm{nat}(u^1_1 v^1_1). \end{array}$$

Since X and Y are closed, this implies $u^0_0 v^0_0 \in X$ and $u^1_1 v^1_1 \in Y$. Hence $\mathrm{rk}(u, v, XY) \leq 1$. □

Corollary 9.2.2 *Let $(M, \cdot, 1)$ be a divisibility monoid with finitely many residuum functions. Let $K, L \subseteq M$ be recognizable. Then $K \cdot L$ is recognizable.*

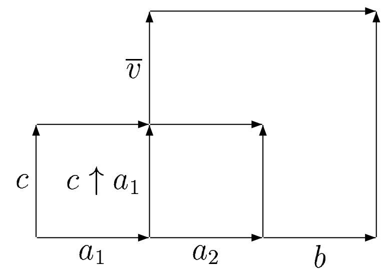

Figure 9.2: The elements from Lemma 9.2.3

Proof. Let $X := \{u \in T^\star \mid \mathrm{nat}(u) \in K\}$ and $Y := \{u \in T^\star \mid \mathrm{nat}(u) \in L\}$. Then $K \cdot L = \mathrm{nat}(XY)$, and X and Y are closed and recognizable in $T^\star$. Hence, by Kleene's Theorem, XY is recognizable in $T^\star$. By Lemma 9.2.1, the rank of XY is finite. Hence Theorem 9.1.8 ensures that $\mathrm{nat}(XY)$ is recognizable in M. □

The rest of this section is devoted to the proof that $\langle L\rangle$ is recognizable for any recognizable language $L \subseteq M$ that is connected and residually closed. But first, we prove some technical lemmas that will be used later on.

Lemma 9.2.3 *Let $(M, \cdot, 1)$ be a divisibility monoid with finitely many residuum functions. Let $a_1, a_2, b, c, \overline{v} \in T^\star$ such that $\mathrm{res}_{a_1 a_2} = \mathrm{res}_b \subseteq \mathrm{id}_{T^\star}$, $c \mathbin{\mathbf{co}} a_1$, $(c \uparrow a_1)\overline{v} \mathbin{\mathbf{co}} a_2 b$ and $a_1 \neq \varepsilon \neq \overline{v}$. Then $\mathrm{nat}((a_1 \uparrow c)\overline{v})$ is not connected.*

Proof. Throughout this proof, we make extensive use of the equations given in Lemma 7.2.7 without mentioning it again.

By $(c \uparrow a_1)\overline{v} \mathbin{\mathbf{co}} a_2 b$, we get in particular $(c \uparrow a_1)\overline{v} \mathbin{\mathbf{co}} a_2$, and $(c \uparrow a_1)\overline{v} \uparrow a_2 \in \mathrm{dom}(\mathrm{res}_b)$. Furthermore, it implies $c{\uparrow}a_1 \mathbin{\mathbf{co}} a_2$ and therefore (together with $c \mathbin{\mathbf{co}} a_1$) $c \mathbin{\mathbf{co}} a_1 a_2$ and $c \uparrow a_1 a_2 = c$. Now we can conclude (with $\tilde{v} := \overline{v} \uparrow (a_2 \uparrow (c \uparrow a_1))$)

$$\begin{aligned} b \mathbin{\mathbf{co}} (c \uparrow a_1)\overline{v} \uparrow a_2 &= ((c \uparrow a_1) \uparrow a_2)\ (\overline{v} \uparrow (a_2 \uparrow (c \uparrow a_1))) \\ &= (c \uparrow a_1 a_2)\tilde{v} \\ &= c\tilde{v}. \end{aligned}$$

Thus we get $c\tilde{v} = (c \uparrow a_1)\overline{v} \uparrow a_2 \in \mathrm{dom}(\mathrm{res}_{a_1 a_2})$ and therefore $c\tilde{v} \uparrow a_1 a_2 = c\tilde{v}$. In particular, we have $a_1 \,\mathbf{co}\, c\tilde{v}$ implying $(a_1 \uparrow c) \,\mathbf{co}\, \tilde{v}$.

Hence we have $(c\uparrow a_1)\overline{v}\uparrow a_2 = c\tilde{v} = (c\tilde{v}\uparrow a_1)\uparrow a_2$. This implies $(c\uparrow a_1)\overline{v} = c\tilde{v}\uparrow a_1$ since res_{a_2} is injective. Note that $c\tilde{v}\uparrow a_1 = (c\uparrow a_1)(\tilde{v}\uparrow(a_1\uparrow c))$. Hence by cancellation we get $\overline{v} = \tilde{v} \uparrow (a_1 \uparrow c)$. This implies $\mathrm{nat}(\tilde{v}) \vee \mathrm{nat}(a_1 \uparrow c) = \mathrm{nat}((a_1 \uparrow c)\overline{v})$.

Since $a_1 \neq \varepsilon$ and res_c is length preserving, we have $a_1\uparrow c \neq \varepsilon$, i.e. $\mathrm{nat}(a_1\uparrow c) \neq 1$. Similarly, $\overline{v} \neq \varepsilon$ implies $\mathrm{nat}(\tilde{v}) \neq 1$. Finally $(a_1 \uparrow c) \,\mathbf{co}\, \tilde{v}$ proves that $\mathrm{nat}((a_1 \uparrow c)\overline{v})$ is not connected. □

Next, we use Ramsey's Theorem (page 61) to show that in a divisibility monoid with finitely many residuum functions, for any sufficiently long sequence $u_1, u_2, \ldots u_n$ of elements of $T^\star$, there is a nonempty fragment of this sequence such that the residuum function of $u_i u_{i+1} \ldots u_j$ is contained in the identity.

Lemma 9.2.4 *Let M be a divisibility monoid with finitely many residuum functions and let $u_i \in T^\star$ for $1 \leq i \leq R_3(|\mathbb{D}_M|)$. Then there exist $1 \leq i < j < k \leq R_3(|\mathbb{D}_M|)$ such that $\mathrm{res}_{u_i u_{i+1} \ldots u_{j-1}} = \mathrm{res}_{u_j u_{j+1} \ldots u_{k-1}} \subseteq \mathrm{id}_{T^\star}$.*

Proof. For simplicity, let $n = R_3(|\mathbb{D}_M|)$. Consider the mapping d' from the 2-element subsets of $[n]$ into $\mathbb{D}_M$ with $d'(\{i,j\}) = \mathrm{res}_{u_i u_{i+1} \ldots u_{j-1}}$ where $i < j$. By Ramsey's Theorem, there are $1 \leq i < j < k \leq n$ with $d'(\{i,j\}) = d'(\{i,k\}) = d'(\{j,k\}) =: f$. Note that f is an idempotent partial function since $d'(\{i,k\}) = d'(\{j,k\}) \circ d'(\{i,j\})$. In addition, f is injective on its domain by Lemma 7.2.7, implying $f \subseteq \mathrm{id}_{T^\star}$. □

Lemma 9.2.5 *Let M be a divisibility monoid with finitely many residuum functions, $(u^i_j, v^j_i)_{1 \leq i \leq j \leq n}$ a complete grid in $T^\star$ with $n \geq R_3(|\mathbb{D}_M|)$ and $u^i_i \neq \varepsilon \neq v^i_i$. Then there exists $1 \leq i \leq n$ such that $\mathrm{nat}(u^i_i v^i_i)$ is not connected.*

Proof. By Lemma 9.2.4, there are $1 \leq i < j < j \leq n$ with

$$\mathrm{res}_{u^0_i u_{i+1}0 \ldots u^0_{j-1}} = \mathrm{res}_{u^0_j u_{j+1}0 \ldots u^0_{k-1}} = \mathrm{res}_{u^0_i u_{i+1}0 \ldots u^0_{k-1}}.$$

With $a_1 = u^0_i$, $a_2 = u^0_{i+1} u^0_{i+2} \ldots u^0_{j-1}$, $x = u^0_j u_{j+1}0 \ldots u^0_{k-1}$, $y = v^{i-1}_0 v^{i-1}_1 \ldots v^{i-1}_{i-1}$ and $\overline{v} = v^i_i$, the assumptions of Lemma 9.2.3 are satisfied. Hence $\mathrm{nat}(u^i_i v^i_i)$ is not connected. □

Theorem 9.2.6 *Let $(M, \cdot, 1)$ be a divisibility monoid with finitely many residuum functions. Let $X \subseteq T^\star$ be closed, connected, and residually closed. Then the rank $\mathrm{rk}(\langle X \rangle)$ of X is at most $R_3(|\mathbb{D}_M|) + 1$ and therefore finite.*

Proof. Let $u, v \in T^\star$ with $\mathrm{nat}(uv) \in \mathrm{nat}(\langle X\rangle) = \langle \mathrm{nat}(X)\rangle$. Then there exist $n \in \mathbb{N}$ and $x_0, x_1, \ldots, x_n \in X$ such that $\mathrm{nat}(uv) = \mathrm{nat}(x_0 x_1 \ldots x_n)$. By Corollary 9.1.6, there exists a complete grid $(u^i_j, v^j_i)_{0\le i\le j\le n}$ in $T^\star$ such that $\mathrm{nat}(u) = \mathrm{nat}(u^0_0 u^0_1 \ldots u^0_n)$, $\mathrm{nat}(v) = \mathrm{nat}(v^n_0 v^n_1 \ldots v^n_n)$ and $\mathrm{nat}(x_i) = \mathrm{nat}(u^i_i v^i_i)$. Since X is closed and connected, the latter implies that $u^i_i v^i_i \in X$ is connected.

There exist $m \le n$ and $-1 = i_0 < i_1 < i_2 \cdots < i_m < i_{m+1} = n+1$ such that

- $u^{i_k}_{i_k}$ and $v^{i_k}_{i_k}$ are nonempty for $0 < k < m+1$,
- u^i_i or v^i_i is empty for $0 \le i \le n$ with $i \notin \{i_0, i_1, \ldots, i_{m+1}\}$.

Consider the subgrid $(a^i_j, b^j_i)_{0\le l\le \ell\le m}$ induced by the sequence $(0, i_1, i_2, \ldots, i_m, n)$. Then $a^k_k = u^{i_k}_{i_k} u^{i_k}_{i_k+1} \ldots u^{i_k}_{i_{k+1}-1}$ and $b^k_k = v^{i_{k+1}-1}_{i_k} v^{i_{k+1}-1}_{i_k+1} \ldots v^{i_{k+1}-1}_{i_{k+1}-1}$ are not empty. Hence, by Lemma 9.2.5, $m \le R_3(|\mathbb{D}_M|)$.

Let $0 \le k \le m$ and $i_k < i < i_{k+1}$. Then $u^i_i v^i_i \in X$. Since one of u^i_i and v^i_i is empty, the other belongs to X, i.e. $u^i_i, v^i_i \in X \cup \{\varepsilon\}$.

Now we show $u^i_j, v^j_i \in X \cup \{\varepsilon\}$ by induction on j for $i_k < i \le j < i_{k+1}$:

Assume $i_k + 1 = j$. Then the claim is trivial since $i = j$ follows. Now assume that for any $i_k \le i \le l \le i_{k+1}$ with $l < j$ we have $u^i_l, v^l_i \in X \cup \{\varepsilon\}$. Then $u^i_j \uparrow (v^{j-1}_i v^{j-1}_{i+1} \ldots v^{j-1}_{j-1}) = u^j_j \in X \cup \{\varepsilon\}$. Note that the upper index $j-1$ of the v's is properly between i_k and j. Hence by the induction hypothesis $v^{j-1}_{i'} \in X \cup \{\varepsilon\}$ for $i' = i, i+1, \ldots, j-1$. Since X and therefore $X \cup \{\varepsilon\}$ is residually closed, we get $u^i_j \in X \cup \{\varepsilon\}$. On the other hand, $v^j_i = v^{j-1}_i \uparrow u^i_j$. By the induction hypothesis, $v^{j-1}_i \in X \cup \{\varepsilon\}$. Hence $v^j_i \in X \cup \{\varepsilon\}$ since u^i_j is an element of this residually closed language.

Now consider the subgrid $(a^i_j, b^i_j)_{0\le i\le j\le 2m+1}$ that is generated by the sequence $(i_1, i_1+1, i_2, i_2+1, \ldots, i_m, i_m+1)$. Then

$$
\begin{aligned}
a^{2k}_{2k} &= u^{i_k+1}_{i_k+1} u^{i_k+1}_{i_k+2} u^{i_k+1}_{i_k+3} \ldots u^{i_k+1}_{i_{k+1}-1}, \\
b^{2k}_{2k} &= v^{i_{k+1}-1}_{i_k+1} v^{i_{k+1}-1}_{i_k+2} v^{i_{k+1}-1}_{i_k+3} \ldots v^{i_{k+1}-1}_{i_{k+1}-1}, \\
a^{2k+1}_{2k+1} &= u^{i_k}_{i_k}, \text{ and} \\
b^{2k+1}_{2k+1} &= v^{i_k}_{i_k}.
\end{aligned}
$$

Note that all the factors of a^{2k}_{2k} and of b^{2k}_{2k} belong to $X \cup \{\varepsilon\}$. This implies $a^{2k}_{2k} b^{2k}_{2k} \in \langle X\rangle$. Thus, the complete grid $(a^i_j, b^j_i)_{0\le i\le j\le 2m+1}$ satisfies

- $\mathrm{nat}(u) = \mathrm{nat}(u^0_0 u^0_1 \ldots u^0_n) = \mathrm{nat}(a^0_0 a^0_1 \ldots a^0_{2m+1})$,
- $\mathrm{nat}(v) = \mathrm{nat}(v^n_0 v^n_1 \ldots v^n_n) = \mathrm{nat}(b^{2m+1}_0 b^{2m+1}_1 \ldots b^{2m+1}_{2m+1})$ and
- $a^i_i b^i_i \in \langle X\rangle$.

Therefore $\mathrm{rk}(u, v, \langle X\rangle) \le 2m+1 \le 2R_3(|\mathbb{D}_M|) + 1$. □

Corollary 9.2.7 *Let $(M, \cdot, 1)$ be a divisibility monoid with finitely many residuum functions. Let $L \subseteq M$ be connected, recognizable, and residually closed. Then the iteration $\langle L \rangle$ of L is recognizable.*

Proof. Let $X := \{w \in T^\star \mid \mathrm{nat}(w) \in L\}$. Note that M and X satisfy the assumptions of Theorem 9.2.6. Hence the rank $\mathrm{rk}(X)$ is finite. By the theorem of Kleene, $\langle X \rangle$ is recognizable in $T^\star$. By Theorem 9.1.8, $\langle L \rangle = \mathrm{nat}(\langle X \rangle)$ is recognizable in M. □

Summarizing the results for obtained so far, we can show that any c-rational language is recognizable.

Theorem 9.2.8 *Let $(M, \cdot, 1)$ be a divisibility monoid with finitely many residuum functions. Let $L \subseteq M$ be c-rational. Then L is recognizable.*

Proof. By Lemma 7.2.3, $\downarrow x$ is finite for any $x \in M$. Hence finite languages in M are recognizable. By Corollary 9.2.2 and 9.2.7, the set of recognizable languages in M is closed under products and the operation $\langle . \rangle$ applied to connected and residually closed languages. □

9.3 From recognizable to c-rational languages

In this section, we will derive conditions on divisibility monoids M which are sufficient to ensure that all recognizable languages in M are c-rational. Let $(M, \cdot, 1)$ be a divisibility monoid. Recall that an equation $\mathrm{nat}(ab) = \mathrm{nat}(cd)$ where a, b, c, d are irreducible generators of M states that the different sequential executions ab and cd give rise to the same effect. If now $a \neq c$, the effect of a in the execution cd has to be resumed by that of d. Therefore, we consider the least equivalence $\equiv$ on the irreducible generators of M identifying a and d that occur in an equation $ab = cd$ with $a \neq c$. To show that any recognizable language is c-rational, we need the property that $\mathrm{nat}(ab) = \mathrm{nat}(cd)$ and $a \equiv c$ imply $a = c$ for any irreducible elements $a, b, c, d \in T$. It is immediate that this is equivalent to the existence of a function $\rho : T \to E$ into some set E satisfying $\rho(s) = \rho(s{\uparrow}t)$ and $\rho(s) \neq \rho(t)$ for any $s, t \in T$ with $s \,\mathbf{co}\, t$. Such a function is called *labeling function*. Since T is finite, we can assume E to be finite, too. A divisibility monoid M together with a labeling function ρ is a *labeled divisibility monoid* (M, ρ). The *label sequence* of a word $u_0 u_1 \ldots u_n \in T^\star$ is the word $\rho(u_0)\rho(u_1)\ldots\rho(u_n) \in E^\star$. We extend the mapping ρ to words over T by $\rho(tw) = \{\rho(t)\} \cup \rho(w)$. Hence ρ is a monoid morphism into the finite monoid $(2^E, \cup, \emptyset)$. By Lemma 8.2.1, $\mathrm{nat}(u) = \mathrm{nat}(v)$ implies $\rho(u) = \rho(v)$ for any $u, v \in T^\star$. Therefore, it is reasonable to define $\rho(\mathrm{nat}(u)) := \rho(u)$, i.e. to extend ρ to a monoid morphism from $(M, \cdot, 1)$ to $(2^E, \cup, \emptyset)$.

A language $L \subseteq M$ is *monoalphabetic* if $\rho(x) = \rho(y)$ for any $x, y \in L$. The class of *mc-rational languages* in the labeled divisibility monoid M is the smallest class $\mathfrak{C} \subseteq 2^M$ satisfying

- any finite subset of M is in $\mathfrak{C}$,
- whenever $L, K \in \mathfrak{C}$ then $L \cup K \in \mathfrak{C}$ and $L \cdot K \in \mathfrak{C}$, and
- whenever $L \in \mathfrak{C}$ is connected and monoalphabetic then $\langle L \rangle \in \mathfrak{C}$.

Note that differently from c-rational languages, here the iteration is restricted to connected and monoalphabetic languages that are not explicitly required to be residually closed. Nonetheless, any mc-rational language is c-rational as Corollary 9.3.2 states.

Lemma 9.3.1 *Let (M, ρ) be a labeled divisibility monoid and $x, y \in M$ with $x \mathbin{\mathbf{co}} y$. Then $\rho(x) \cap \rho(y) = \emptyset$, $\rho(y) = \rho(y \uparrow x)$, and $\rho(x) \cup \rho(y) = \rho(x \vee y)$.*

Proof. By contradiction, assume $\rho(x) \cap \rho(y) \neq \emptyset$. Then there exist monoid elements $x_1, x_2, y_1, y_2 \in M$ and $s, t \in T$ such that $x = x_1 s x_2$, $y = y_1 t y_2$ and $\rho(s) = \rho(t)$. Clearly, $x_1 s \mathbin{\mathbf{co}} y_1 t$. By Lemma 7.2.5(1) and (3), we have $s \mathbin{\mathbf{co}} (y_1 t \uparrow x_1) = (y_1 \uparrow x_1)(t \uparrow (x_1 \uparrow y_1))$. Hence $s' := s \uparrow (y_1 \uparrow x_1) \mathbin{\mathbf{co}} (t \uparrow (x_1 \uparrow y_1)) =: t'$ by Lemma 7.2.5(1). Furthermore, $\rho(s) = \rho(s')$ and $\rho(t) = \rho(t')$. But this contradicts the definition of a labeling function. Hence the first statement is shown.

By a simple induction on the length of x and y we get $\rho(y \uparrow x) = \rho(y)$, i.e. the second statement. Now the last assertion follows since $\rho(x \vee y) = \rho(x(y \uparrow x)) = \rho(x) \cup \rho(y \uparrow x) = \rho(x) \cup \rho(y)$. □

Corollary 9.3.2 *Let $(M, \cdot, 1, \rho)$ be a labeled divisibility monoid. Then any mc-rational language in M is c-rational.*

Proof. Let $x, y \in M$ with $\rho(x) = \rho(y)$ and $x \mathbin{\mathbf{co}} y$. Then Lemma 9.3.1 ensures $x = y = 1$. Hence any monoalphabetic language is residually closed. This implies that mc-rational languages are c-rational. □

Now let $\preceq$ be a linear order on the set E and let $u \in T^\star$. We say that the word *u is in lexicographical normal form* if $\mathrm{nat}(u) = \mathrm{nat}(v)$ implies that the label sequence of v is lexicographically larger than or equal to that of u. Let LNF be the set of all words in lexicographic normal form.

Note that so far there may exist different words u and v in LNF with $\mathrm{nat}(u) = \mathrm{nat}(v)$. We show that this is impossible: So assume that $u, v \in T^\star$ have the same label sequence and satisfy $\mathrm{nat}(u) = \mathrm{nat}(v)$. Let s (t) be the first letter of u (of v, respectively) and suppose $s \neq t$. Then $\mathrm{nat}(s)$ and $\mathrm{nat}(t)$ are incomparable since they have the same length. Thus the infimum of them is properly smaller than $\mathrm{nat}(s)$. Since 1 is the only element which is properly smaller than the irreducible element $\mathrm{nat}(s)$, we get $\mathrm{nat}(s) \wedge \mathrm{nat}(t) = 1$. Since $\mathrm{nat}(s)$ and $\mathrm{nat}(t)$ are bounded above by $\mathrm{nat}(u) = \mathrm{nat}(v)$, we have $s \mathbin{\mathbf{co}} t$. Hence $\rho(s) \neq \rho(t)$, contradicting the fact that the label sequences of u and v coincide. Hence $s = t$. Cancelling s and t from the left of u and v, respectively, we can proceed by induction. At the end, we obtain $u = v$. Hence, for any $x \in M$ there exists at most one $u \in \mathrm{LNF}$ such that $x = \mathrm{nat}(u)$. Since, on the other hand, the lexicographical order on $E^\star$ is a well-order, the set of label sequences of words u with $\mathrm{nat}(u) = x$ has a least element. Hence, for any $x \in M$ there exists a unique word $u \in \mathrm{LNF}$ with $x = \mathrm{nat}(u)$. This word is called *the lexicographic normal form of* x. We denote it by $\mathrm{lexNF}(x)$.

Next, we characterize the set of words in lexicographic normal form. This result generalizes the characterization of lexicographic normal forms for trace monoids given in [AK79].

Lemma 9.3.3 *Let (M, ρ) be a labeled divisibility monoid and let $\preceq$ be a linear order on T. Let $u_i \in T$ for $0 \leq i \leq n$. Then $u_0 u_1 \dots u_n \in \mathrm{LNF}$ iff*

$$u_j \in \mathrm{res}_{u_i u_{i+1} \dots u_{j-1}}(T) \Rightarrow \rho(u_i) \prec \rho(u_j) \qquad (\star)$$

for $0 \leq i < j \leq n$.

Proof. For simplicity, let $u := u_0 u_1 \dots u_n$. First let $u \in \mathrm{LNF}$ and assume there are i, j with $0 \leq i < j \leq n$ such that $u_j \in \mathrm{res}_{u_i u_{i+1} \dots u_{j-1}}(T)$ and $\rho(u_i) \succeq \rho(u_j)$. Then there is $t \in T$ with $u_j = \mathrm{res}_{u_i u_{i+1} \dots u_{j-1}}(t)$. Hence

$$\mathrm{nat}(u) = \mathrm{nat}(u_0 \dots u_{i-1} t \, \mathrm{res}_t(u_i \dots u_{j-1}) \, u_{j+1} \dots u_n).$$

Since $\rho(u_i) \succeq \rho(u_j) = \rho(t)$, the label sequence of u is larger than or equals that of $u_0 \dots u_{i-1} t \, \mathrm{res}_t(u_i \dots u_{j-1}) \, u_{j+1} \dots u_n$. Since u is in lexicographical normal form, this implies in particular $\rho(u_i) = \rho(t)$, contradicting $t \in \mathrm{dom}(\mathrm{res}_{u_i})$. Thus, u satisfies the property $(\star)$.

Conversely, let the word u satisfies the property $(\star)$. Let $v \in \mathrm{nat}(u)$ with $u \neq v$. We claim that u is lexicographically smaller than v. Note that any suffix of u satisfies $(\star)$. Hence we may assume that the first letter t of v is different from u_0. Then $\mathrm{nat}(t)$ and $\mathrm{nat}(u_0)$ are bounded above by $\mathrm{nat}(u)$. Since they are different irreducible elements in M, their infimum is trivial. Hence $\mathrm{nat}(t) \mathbin{\mathbf{co}} \mathrm{nat}(u_0)$ implying $\rho(t) \neq \rho(u_0)$. Let j be the least integer such that $\rho(t) \in \rho(u_0 u_1 \dots u_j)$. By Lemma 9.3.1, $\mathrm{nat}(t)$ and $\mathrm{nat}(u_0 u_1 \dots u_j)$ are not complementary. Since they are bounded by $\mathrm{nat}(u)$, the infimum cannot be 1. Hence $\mathrm{nat}(t) \leq \mathrm{nat}(u_0 u_1 \dots u_j)$.

Since on the other hand the infimum of $\mathrm{nat}(t)$ and $\mathrm{nat}(u_0u_1\dots u_{j-1})$ is trivial, the supremum of these two equals $\mathrm{nat}(u_0u_1\dots u_j)$. Hence $u_j = \mathrm{res}_{u_0u_1\dots u_{j-1}}(t)$. Since u satisfies $(\star)$, this implies $\rho(u_0) \prec \rho(u_j) = \rho(t)$ and hence our claim. Thus $u \in \mathrm{LNF}$. ☐

Using the lemma above, we show that the set of words in lexicographic normal form is recognizable:

Lemma 9.3.4 *Let (M, ρ) be a labeled divisibility monoid. Then* LNF *is recognizable in the free monoid $T^\star$.*

Proof. Recall that $\mathbb{D}_M = \{\mathrm{res}_u \mid u \in T^\star\}$ is a monoid consisting of partial functions from $T^\star$ to $T^\star$. These functions are length preserving. In particular, they map elements of T to elements of T. Hence $\mathbb{D}_M \restriction T := \{\mathrm{res}_u \restriction T \mid u \in T^\star\}$ is a monoid. It is finite since T is finite. Recall furthermore, that the mapping $T^\star \to \mathbb{D}_M$ defined by $u \mapsto \mathrm{res}_u$ is a monoid antihomomorphism. Hence the mapping from $T^\star$ to $\mathbb{D}_M \restriction T$ with $u \mapsto \mathrm{res}_u \restriction T$ is a monoid antihomomorphism, too. This implies that the sets $X_d := \{u \in T^\star \mid \mathrm{res}_u \restriction T = d\}$ for $d \in \mathbb{D}_M \restriction T$ are recognizable in $T^\star$. Hence they are rational by Kleene's Theorem.

For $S \subseteq T$, $d \in \mathbb{D}_M$ or $d \in \mathbb{D}_M \restriction T$, let $d(S) := \{d(s) \mid s \in S \cap \mathrm{dom}(d)\}$ which is a finite set. Now by Lemma 9.3.3 the set of words over T that are not in lexicographical normal form equals the rational language

$$T^\star \setminus \mathrm{LNF} = \bigcup_{\substack{s\in T\\ d\in\mathbb{D}_M\restriction T}} T^\star\{s\}X_d\ (\mathrm{res}_s \circ d)(\{t \in T \mid \rho(t) \preceq \rho(s)\})\ T^\star.$$

Hence LNF is recognizable. ☐

The crucial point in Ochmański's proof of the c-rationality of recognizable languages in trace monoids is that whenever a square of a word is in lexicographic normal form, it is actually connected. This does not hold any more for labeled divisibility monoids. But we can show that whenever a product of $|E| + 2$ words having the same set of labels is in lexicographic normal form, it is connected (cf. Corollary 9.3.6). This enables us to show that recognizable languages are mc-rational.

For a set $A \subseteq E$ and $u \in T^\star$ let $n_A(u)$ denote the number of maximal factors w of u with $\rho(w) \subseteq A$ or $\rho(w) \cap A = \emptyset$. The number $n_A(u)$ is the number of blocks of elements of A and of $E \setminus A$ in the label sequence of u. For example, let $u = u_1u_2\dots u_n \in T^\star$ with $u_i \neq \varepsilon$, $\rho(u_{2i}) \subseteq A$ and $\rho(u_{2i+1}) \subseteq E \setminus A$ for all suitable i. Then $n = n_A(u) = n_{E\setminus A}(u)$. Furthermore, we put $n_A(x) := n_A(\mathrm{lexNF}(x))$ for $x \in M$.

Lemma 9.3.5 *Let* $(M, \cdot, 1, \rho)$ *be a labeled divisibility monoid,* $x, y \in M$ *and* $x \mathbin{\mathbf{co}} y$. *Then* $n_{\rho(x)}(x \vee y) \leq |E| + 1$.

Proof. If $n_{\rho(x)}(x \vee y) = 1$, the statement is trivial. So let $n_{\rho(x)}(x \vee y) \geq 2$. Since $n_A(x \vee y) = n_{E \setminus A}(x \vee y)$ for any $A \subseteq E$, we may assume that the label sequence of $\mathrm{lexNF}(x \vee y)$ starts with a letter from $A := \rho(x)$.

Hence there exist words $u, v \in T^+$ and $u' \in T^\star$ with $\rho(u) \subseteq A$, $\rho(v) \subseteq E \setminus A$ and $\mathrm{lexNF}(x \vee y) = uvu'$. Now let a be the first letter of u and b the first one of v.

First we show $\rho(\mathrm{nat}(ub) \wedge y) = \rho(b)$: Let $h := \mathrm{nat}(ub) \wedge y$. Then there exist uniquely determined $k, l \in M$ with $\mathrm{nat}(ub) = h \cdot k$ and $y = h \cdot l$. By Lemma 7.1.1, $k \mathbin{\mathbf{co}} l$. Hence $\rho(k) \cap \rho(l) = \emptyset$ by Lemma 9.3.1. We write $\#_e h$ for the number of occurrences of the letter $e \in E$ in the label sequence of any representative of $h \in M$, which is well-defined by Lemma 8.2.1 and the requirements on ρ. So we get on the other hand, $\#_e \mathrm{nat}(ub) = \#_e h + \#_e k$ and $\#_e y = \#_e h + \#_e l$ for any $e \in E$. Hence $\#_e h = \min(\#_e \mathrm{nat}(ub), \#_e y)$. Note that $\rho(b) \in \rho(y)$. Hence $\rho(b) \in \rho(h)$. Now let $e \in \rho(y) \setminus \rho(b)$. Then $e \notin \rho(x)$ and therefore $\#_e h = 0$. Thus $\rho(\mathrm{nat}(ub) \wedge y) = \rho(b)$.

Hence there exists a word $w \in T^\star$ with $x \vee y = \mathrm{nat}(w)$ such that the label sequence of w starts with $\rho(b)$. Since the label sequence of the lexicographical normal form of $x \vee y$ starts with $\rho(a)$, we get $\rho(a) \prec \rho(b)$.

So let $\mathrm{lexNF}(x \vee y) = u_0 v_0 u_1 v_1 \ldots u_n v_n$ with $u_i \neq \varepsilon$ for all $i \leq n$, $v_i \neq \varepsilon$ for $i < n$, $\rho(u_i) \subseteq A$ and $\rho(v_i) \subseteq E \setminus A$. Let a_i (b_i) be the first letter of u_i for $i \leq n$ (v_i for $i < n$, resp.). Using Lemma 7.1.1, we can apply the above result inductively and obtain $\rho(a_i) \prec \rho(b_i) \prec \rho(a_{i+1})$ for each $i < n$. Hence $2n + 1 \leq |E|$. □

Corollary 9.3.6 *Let* $(M, \cdot, 1, \rho)$ *be a labeled divisibility monoid and let* $X \subseteq T^\star$ *be a monoalphabetic language. Then* $\mathrm{nat}(w)$ *is connected for any word* $w \in X^{|E|+2} \cap \mathrm{LNF}$.

Proof. Let $n = |E|+1$. Then there exist $x_i \in \mathrm{nat}(X)$ with $\mathrm{nat}(w) = x_0 x_1 \ldots x_n$. Now let $x, y \in M$ with $x \mathbin{\mathbf{co}} y$ and $x \vee y = \mathrm{nat}(w)$. Then $\rho(x) \cap \rho(y) = \emptyset$ by Lemma 9.3.1. If $\rho(x_i) \cap \rho(x) \neq \emptyset$ and $\rho(x_i) \cap (E \setminus \rho(x)) \neq \emptyset$ for all $0 \leq i \leq n$, we would obtain $n_{\rho(x)}(\mathrm{nat}(w)) > n = |E| + 1$, contradicting Lemma 9.3.5. Hence there exists $i \in \{0, 1, \ldots, n\}$ such that $\rho(x_i) \subseteq \rho(x)$ or $\rho(x_i) \subseteq E \setminus \rho(x)$.

First consider the case $\rho(x_i) \subseteq \rho(x)$. Since X is monoalphabetic, this implies $\rho(x_j) = \rho(x_i) \subseteq \rho(x)$ for all $0 \leq j \leq n$. Now $\rho(y) = \emptyset$ follows from the inclusions $\rho(y) \subseteq \rho(w) \subseteq \rho(x)$ and from $\rho(x) \cap \rho(y) = \emptyset$. Hence $y = 1$.

Now consider the case $\rho(x_i) \subseteq E \setminus \rho(x)$. From Lemma 9.3.1, we obtain $\rho(x) \cup \rho(y) = \rho(\mathrm{nat}(w)) \supseteq \rho(x_i)$ and this implies $\rho(x_i) \subseteq \rho(y)$. Now we can argue as above (with x and y interchanged) and obtain $x = 1$. □

Corollary 9.3.7 *Let $(M, \cdot, 1, \rho)$ be a labeled divisibility monoid. Let $L \subseteq M$ be recognizable. Then L is mc-rational.*

Proof. Let $X := \{u \in T^\star \mid \mathrm{nat}(u) \in L\} \cap \mathrm{LNF}$. Since any $x \in M$ has a unique lexicographical normal form, we have $\mathrm{nat}(X) = L$. Then X is recognizable in $T^\star$ and therefore rational. By Lemma 7.1.2, it can be constructed from finite languages in $T^\star$ by the operation $\cdot$, $\cup$ and $\langle . \rangle$ applied to monoalphabetic languages, only. Since $X \subseteq \mathrm{LNF}$, any intermediate language in the construction of X is contained in LNF, too. Let Y be such an intermediate language and suppose that the iteration $\langle . \rangle$ is applied to Y. Hence Y is monoalphabetic. Then $\langle Y \rangle$ is another intermediate language and therefore contained in LNF. Hence by Corollary 9.3.6, $\mathrm{nat}(Y)^{|E|+2}$ is connected. Note that $\langle Y \rangle = \left(\bigcup_{0 \le i \le |E|+1} Y^i\right) \left\langle Y^{|E|+2} \right\rangle$. Therefore, we can construct $\mathrm{nat}(X) = L$ as required. □

We can summarize our results on recognizable, c-rational and mc-rational languages as follows.

Theorem 9.3.8 *Let $(M, \cdot, 1, \rho)$ be a labeled divisibility monoid with finitely many residuum functions. Let $L \subseteq M$. Then the following are equivalent:*

1. *L is recognizable*
2. *L is c-rational*
3. *L is mc-rational.*

Proof. The implications $2 \to 1 \to 3 \to 2$ are Theorem 9.2.8, Corollary 9.3.7 and 9.3.2, respectively. □

Let T be a finite set and $\rho : T \to T$ the identity on T. Then ρ is a labeling function on the free monoid $T^\star$. Since free monoids are divisibility monoids, the theorem above generalizes Kleene's Theorem - but our proof used Kleene's Theorem as well. The situation regarding trace monoids is different: Any trace monoid can be considered as a labeled divisibility monoid with finitely many residuum functions. Hence Theorem 9.3.8 generalizes Ochmański's Theorem (and extends it slightly by the consideration on mc-rational languages). In [DK00], we showed that Theorem 9.3.8 also generalizes the main result from [Dro95] on concurrency monoids. Even more, we could show that Droste's result holds for all stably concurrent automata and not only, as shown in [Dro95], for a certain subclass.