Difference hierarchies and duality with an application to formal languages
Célia Borlido, Mai Gehrke, Andreas Krebs, Howard Straubing

To cite this version:
Célia Borlido, Mai Gehrke, Andreas Krebs, Howard Straubing. Difference hierarchies and duality with an application to formal languages. Topology and its Applications, Elsevier, 2019. hal-02413264

HAL Id: hal-02413264
https://hal.archives-ouvertes.fr/hal-02413264
Submitted on 16 Dec 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Difference hierarchies and duality
with an application to formal languages

Célia Borlido\textsuperscript{a,*}, Mai Gehrke\textsuperscript{a}, Andreas Krebs\textsuperscript{b}, Howard Straubing\textsuperscript{c}

\textsuperscript{a}Laboratoire J. A. Dieudonné, CNRS, Université Côte d’Azur, France
\textsuperscript{b}Wilhelm-Schickard-Institut Universität Tübingen, Germany
\textsuperscript{c}Boston College, United States

Abstract

The notion of a difference hierarchy, first introduced by Hausdorff, plays an important role in many areas of mathematics, logic and theoretical computer science such as descriptive set theory, complexity theory, and the theory of regular languages and automata. Lattice theoretically, the difference hierarchy over a distributive lattice stratifies the Boolean algebra generated by it according to the minimum length of difference chains required to describe the Boolean elements. While each Boolean element is given by a finite difference chain, there is no canonical such writing in general. We show that, relative to the filter completion, or equivalently, the lattice of closed upsets of the dual Priestley space, each Boolean element over the lattice has a canonical minimum length decomposition into a Hausdorff difference chain. As a corollary, each Boolean element over a co-Heyting algebra has a canonical difference chain (and an order dual result holds for Heyting algebras). With a further generalization of this result involving a directed family of closure operators on a Boolean algebra, we give an elementary proof of the fact that if a regular language is given by a Boolean combination of universal sentences using arbitrary numerical predicates then it is also given by a Boolean combination of universal sentences using only regular numerical predicates.

\textsuperscript{*}This project has received funding from the European Research Council under the European Union’s Horizon 2020 research and innovation program (grant agreement No. 670624). For further partial funding, see acknowledgements.

\textsuperscript{*}Corresponding author

Email addresses: cborlido@unice.fr (Célia Borlido), mgehrke@unice.fr (Mai Gehrke), mail@krebs-net.de (Andreas Krebs), howard.straubing@bc.edu (Howard Straubing)
Keywords: Difference hierarchies, Stone-Priestley duality, Logic on Words
2000 MSC: 06D50, 68F05

In honour of Aleš Pultr on the occasion of his 80th birthday

1. Introduction

Hausdorff introduced the notion of a difference hierarchy in his work on set theory [17]. Subsequently, the notion has played an important role in descriptive set theory as well as in complexity theory. More recently, it has seen a number of applications in the theory of regular languages and automata [15, 4]. From a lattice theoretic point of view, the difference hierarchy over a distributive lattice $D$ stratifies the universal Boolean envelope of $D$. This is the (unique up to isomorphism) Boolean algebra $B$ containing $D$ as a sublattice and generated (as a Boolean algebra) by $D$. We follow the tradition in lattice theory calling $B$ the Booleanization of $D$. However, we warn any frame theorist reading the paper that this is not the same thing as the Booleanization of a frame [1], which generalizes the construction of the complete Boolean algebra of regular opens of a topological space.

In the difference hierarchy over $D$, the stratification of the Booleanization $B$ of $D$ is made according to the minimum length of difference chains required to describe an element $b \in B$:

$$b = a_1 - (a_2 - (\ldots (a_{n-1} - a_n)\ldots))$$

where $a_1 \geq a_2 \geq \cdots \geq a_{n-1} \geq a_n$ are elements of $D$. One difficulty in the study of difference hierarchies is that in general elements $b \in B$ do not have canonical associated difference chains.

Priestley duality [25] for distributive lattices uses the Stone space of the Booleanization equipped with a partial order to represent the lattice as the closed and open (henceforth called clopen) upsets of the associated Priestley space. Priestley duality provides an elucidating tool for the study of difference hierarchies. For one, the minimum length of difference chains for an element $b \in B$ has a nice description relative to the Priestley dual space $X$ of $D$ as the length of the longest chain of points $x_1 < x_2 < \cdots < x_n$ in $X$ so that $x_i$ belongs to the clopen corresponding to $b$ if and only if $i$ is odd. Further, if we allow difference chains of closed upsets of the Priestley
space, rather than clopen upsets, then every element $b \in B$ has a canonical
difference chain which is of minimum length. If in addition $D$ is a co-Heyting
algebra, then the elements of this canonical difference chain are clopen and
thus every $b \in B$ has a canonical difference chain in $D$. We present this ma-
terial, which is closely related to work by Leo Esakia on skeletal subalgebras
of closure algebras [9], in Section 3. We have chosen to present this first part
in the language of point-set topology to provide a treatment which is more
easily accessible for researchers not used to point-free topology.

The main purpose of Section 5 is to obtain sufficient conditions so that,
given a Boolean algebra $B$, a Boolean subalgebra $B'$, and a sublattice $D$ of $B$,
the intersection of $B'$ and $D$ generates the intersection of $B'$ with the Boolean
algebra generated by $D$. For this purpose we consider a situation where $B$
is equipped with a family of closure operators whose meet-subsemilattices of
closed elements form a directed family the union of which is a sublattice of $B$.
Using Stone-Priestley duality in the point-free form of canonical extensions,
we generalize the results of Section 3 and use them to derive the main theorem
of the section.

In turn, the results of Section 5 are used in Section 8 for an application in
logic on words. More precisely, we give an elementary proof of the equality

$$B\Pi_1[N] \cap \text{Reg} = B\Pi_1[\text{Reg}].$$  \hspace{1cm} (2)

The idea is that this equality may be translated into an instance of the main
theorem of Section 5 where $B = \mathcal{P}(A^*)$ is the Boolean algebra of all formal
languages over the alphabet $A$, $B' = \text{Reg}$ is the Boolean subalgebra of regular
languages, and $D = \Pi_1[N]$ is the sublattice of languages given by universal
sentences using arbitrary numerical predicates. Universally quantifying over
a finite number of variables yields a closure operator, and the directed union
of the closed elements is exactly $D = \Pi_1[N]$. The equation (2) was first
proved by Macial, Péladeau and Thérien [20]. For more details, see [29].

Before each of the main Sections 3, 5, and 8 we include the background
needed. The aim is in this way to make the paper accessible both to lattice
and frame theorists and to researchers working with formal languages and
automata. Thus, in Section 2 we introduce the basics on lattices and duality,
Section 4 is an introduction to canonical extensions, and Section 6 contains
the preliminaries on recognition and logic on words. Section 7 provides a
lattice theoretical perspective on logic on words, as well as its connections
with model theory, which is the point of view adopted in Section 8. Although,
strictly speaking, the results of this section can not be considered as new, to the best of our knowledge, the presentation of the material in this form is an original contribution of the paper. Finally, in Section 9 we discuss possible generalizations of the present work, and in particular, we give some examples of open problems that one could try to handle with the techniques developed here.

2. Preliminaries on lattices and duality

We consider all distributive lattices to be bounded and we view the classes of distributive lattices and of Boolean algebras as categories in which the morphisms are the algebraic homomorphisms, that is, the maps that preserve all the basic operations (including the bounds). For readers needing more detailed preliminaries on lattices and duality, we refer to [7].

Priestley duality. The Priestley dual space of a distributive lattice $D$ consists of the set $S(D)$ of homomorphisms from $D$ to the two-element lattice $2$ (or equivalently, of the prime filters of $D$) ordered point-wise and equipped with the topology generated by the sets $\hat{a} = \{ x \in S(D) \mid x(a) = 1 \}$ and their complements, for $a \in D$. One can show that the resulting ordered topological space, $(X, \leq, \pi)$, is compact and totally order disconnected. That is, if $x \not\approx y$ in $X$ then there is a clopen upset $V$ of $X$ with $x \in V$ and $y \not\in V$. Totally order disconnected compact spaces are called Priestley spaces and the appropriate structure preserving maps are the continuous and order preserving maps. In the other direction, given a Priestley space $(X, \leq, \pi)$ the collection $\text{UpClopen}(X, \leq, \pi)$ of subsets of $X$ that are clopen upsets forms a lattice of sets (this may also be seen as the Priestley morphisms into the Priestley space based on the two-element chain). Morphisms correspond contravariantly and the correspondence is given by pre-composition. These functors account for the dual equivalence of the category of distributive lattices and the category of Priestley spaces. On objects, this means that $D \cong \text{UpClopen}(S(D))$ (via the map $a \mapsto \hat{a}$) for any distributive lattice $D$ and $X \cong S(\text{UpClopen}(X))$ (via the map $x \mapsto \chi_x$ where $\chi_x$ is the characteristic function of the point $x$ restricted to $\text{UpClopen}(X)$) for any Priestley space $X$. In addition, the double dual of a morphism, on either side of the duality, is naturally isomorphic to the original. For more details see [7, Chapter 11].

Booleanization. The Booleanization, $D^\sim$, of a distributive lattice $D$ is a Boolean algebra with a lattice embedding $D \hookrightarrow D^\sim$ so that any lattice
homomorphism $h: D \to B$ into a Boolean algebra, uniquely extends to a homomorphism $h^- : D^- \to B$ making the following diagram commutative:

$$
\begin{array}{c}
D \\
\downarrow h \\
D^- \\
\downarrow h^- \\
B
\end{array}
$$

It is well-known that for any embedding of $D$ into any Boolean algebra, the Boolean algebra generated by the image is isomorphic to $D^-$. The existence of such a map is a consequence of Stone-Priestley duality, but showing that this map is an embedding requires a non-constructive principle. For more on Booleanizations of lattices in a constructive manner, please see [21, 23, 16, 6].

In the setting of Priestley duality, we have seen that $D$ is isomorphic to the lattice of clopen upsets of its dual space $X$. It thus follows that $D^-$ is isomorphic to the Boolean subalgebra of $\mathcal{P}(X)$ generated by $\text{UpClopen}(X)$. One can show that this is the Boolean algebra of all clopen subsets of $X$. That is, $D^- \cong \text{Clopen}(X)$.

**Adjunctions and closure operators.** Let $P$ and $Q$ be posets. We say that maps $f : P \rightleftarrows Q : g$ form an adjoint pair provided

$$
\forall p \in P, q \in Q \quad (f(p) \leq q \iff p \leq g(q)).
$$

Note that in this case, $f$ and $g$ uniquely determine each other since

$$
f(p) = \bigwedge \{ q \in Q \mid p \leq g(q) \} \quad \text{and} \quad g(q) = \bigvee \{ p \in P \mid f(p) \leq q \}.
$$

We call $f$ the lower adjoint of $g$ and $g$ the upper adjoint of $f$. One can show that lower adjoints preserve all existing suprema, while upper adjoints preserve all existing infima. In the case that the posets $P$ and $Q$ are complete lattices this gives a simple criterion for the existence of adjoints.

**Proposition 1.** A map between complete lattices has a lower adjoint if and only if it preserves arbitrary meets, and it has an upper adjoint if and only if it preserves arbitrary joins.

Adjoint pairs are intimately related to closure operators. Recall that, for a poset $P$, a function $c : P \to P$ is a closure operator provided

$$
\forall p_1, p_2 \in P \quad (p_1 \leq c(p_2) \iff c(p_1) \leq c(p_2)).
$$
Proposition 2. If \((f: P \Rightarrow Q : g)\) is an adjoint pair, then \(gf\) is a closure operator on \(P\). Conversely, every closure operator \(c : P \rightarrow P\) may be obtained in this way, e.g. \((c' : P \Rightarrow \text{Im}(c) : \iota)\), where \(c'\) is the co-restriction of \(c\) to its image and \(\iota\) the inclusion map, is such an adjoint pair.

Finally, since upper adjoints preserve all existing infima, one can show that if \(c : P \rightarrow P\) is a closure operator and \(P\) admits a meet-semilattice structure, then \(\text{Im}(c)\) is a meet-subsemilattice of \(P\). For more details, see [7, Chapter 7].

Heyting and co-Heyting algebras. Heyting algebras are the algebras for intuitionistic propositional logic in the same sense that Boolean algebras are the algebras for classical propositional logic. The order dual notion is called a co-Heyting algebra. We will focus on co-Heyting algebras here as this is more convenient for the sequel, but any result about one notion has a corresponding order dual result about the other notion.

Definition 3. A co-Heyting algebra is a distributive lattice equipped with an additional binary operation, \(\slash\), which is the lower adjoint of the operation \(\lor\) in the sense that we have

\[
\forall a, b, c \in D \quad (a/b \leq c \iff a \leq b \lor c).
\]

Notice that this property implies that the co-Heyting operation on a distributive lattice, if it exists, is unique and is given by

\[
a/b = \bigwedge\{c \mid a \leq b \lor c\}.
\]

Thus one may think of a co-Heyting algebra as a special kind of distributive lattice. Indeed, the following proposition, which is the order dual of the corresponding fact for Heyting algebras, see [10, Proposition 3], gives such a characterization.

Proposition 4. A distributive lattice \(D\) is a co-Heyting algebra if and only if the inclusion of \(D\) in its Booleanization has a lower adjoint.

Definition 5. The lower adjoint mentioned in Proposition 4 gives, for each \(b \in D^-\), the least over-approximation of \(b\) in \(D\). We denote it by

\[
D^- \rightarrow D, \quad b \mapsto \lceil b \rceil = \bigwedge\{c \in D \mid b \leq c\}.
\]

and call it the ceiling function (of \(D\)).
It will be useful to understand which Priestley spaces correspond to co-Heyting algebras. The order dual characterization for Heyting algebras is due to Esakia [8] independently of Priestley’s work. We include a proof of the result for co-Heyting algebras to illustrate the correspondence between the algebraic and topological formulations.

Theorem 6 ([8]). Let $D$ be a distributive lattice and $X$ its Priestley dual. Then $D$ admits a co-Heyting structure if and only if for each $V \subseteq X$ clopen, $\uparrow V$ is again clopen. When this is the case, the map $\lceil \rceil : D^\sim \to D$ is naturally isomorphic to the map $V \mapsto \uparrow V$ on clopen subsets of $X$.

Proof. This is a simple consequence of Proposition 4. Note that by total order disconnectedness of $X$, for any closed (and thus compact) $K \subseteq X$, we have

$$\uparrow K = \bigcap\{W \subseteq X \mid K \subseteq W \text{ and } W \text{ clopen upset}\}.$$  \hspace{1cm} (3)

Therefore, there is a least clopen upset (i.e., element of $D$) above $V$ if and only if $\uparrow V$ is clopen. \hfill \Box

Closed upsets in Priestley spaces. Note that closed subspaces of Priestley spaces are again Priestley spaces. In fact, the closed subspaces of a Priestley space $X$ correspond to the lattice quotients of its dual $D$, cf. [7, Section 11.32].

The following well-known fact about upsets of closed sets will be used extensively in the sequel. We include a proof for the sake of completeness. For a subset $S \subseteq P$ of a poset $P$, we use $\min(S)$ to denote the set of minimal elements of $S$.

Proposition 7. Let $X$ be a Priestley space and $K \subseteq X$ a closed subset. Then, $\uparrow K = \uparrow \min(K)$ and this is a closed subset of $X$.

Proof. As seen in (3), $\uparrow K$ is a closed subset of $X$ whenever $K \subseteq X$ is. Now consider $X$ as the dual space of a distributive lattice $D$. Then the points of $X$ are the prime filters of $D$. Let $x$ be any element of $K$ and let $C$ be a maximal chain of prime filters contained in $K$ with $x \in C$. Since $C$ is a chain, it is easy to show that $x_0 = \bigcap_{x \in C} x$ is again a prime filter. Also, if $W = \hat{a}$ is any clopen upset of $X$ with $K \subseteq W$, then $a \in y$ for all $y \in K$ and thus $a \in x_0$. It follows that $x_0 \in W$ for all clopen upsets $W$ of $X$ with $K \subseteq W$ and thus $x_0 \in \uparrow K$. Now by maximality of $C$ it follows that $x_0 \in \min(K)$ and $x_0 \leq x$. Thus $\uparrow K = \uparrow \min(K)$. \hfill \Box
3. The difference hierarchy and closed upsets

Let $D$ be a distributive lattice and $D^-$ its Booleanization. Since $D^-$ is generated by $D$ as a Boolean algebra and because of the disjunctive normal form of Boolean expressions, every element of $D^-$ may be written as a finite join of elements of the form $a - b$ with $a, b \in D$. A fact, that is well known but somewhat harder to see is that every element of $D^-$ is of the form

$$a_1 - (a_2 - (\ldots - (a_{n-1} - a_n))\ldots),$$

for some $a_1, \ldots, a_n \in D$. The usual proof is by algebraic computation and is not particularly enlightening. It is also a consequence of our results here.

We begin with a technical observation.

**Proposition 8.** Let $B$ be a Boolean algebra and let $a_1 \geq \cdots \geq a_{2m}$ be a decreasing sequence of elements of $B$. Then, the following equality holds:

$$a_1 - (a_2 - (\cdots -(a_{2m-1} - a_{2m})\ldots)) = (a_1 - a_2) \lor (a_3 - (\cdots -(a_{2m-1} - a_{2m})\ldots))$$

where the join is disjoint, and by induction we obtain

$$a_1 - (a_2 - (\cdots -(a_{2m-1} - a_{2m})\ldots)) = \bigvee_{n=1}^{m}(a_{2n-1} - a_{2n})$$

where the joinands are pairwise disjoint.

**Proof.** Let us denote $b = a_3 - (a_4 - (\cdots -(a_{2m-1} - a_{2m})\ldots))$. A simple algebraic computation yields $a_1 - (a_2 - b) = (a_1 - a_2) \lor (a_1 \land b)$. Since $a_1, \ldots, a_{2m}$ is a decreasing chain and $b \leq a_3$, it follows that $a_1 \land b = b$, which in turn yields (5). \qed

One problem with difference chain decompositions of Boolean elements over a distributive lattice, which makes them difficult to understand and work with, is that, in general, there is no ‘most efficient’ such decomposition. We give an example of a Boolean element over a distributive lattice that illustrates this problem.

**Example 9.** Consider $X = \mathbb{N} \cup \{x, y\}$ equipped with the topology of the one-point compactification of the discrete topology on $\mathbb{N} \cup \{x\}$. That is, the frame of opens of $X$ is:
The order relation on $X$ is as depicted. That is, the only non-trivial order relation in $X$ is $x \leq y$. It is not hard to verify that $X$ is a Priestley space.

The clopen upsets of $X$ are the finite subsets of $N$ and the cofinite subsets of $X$ containing $y$ and they form the lattice $D$ dual to $X$. Note that $V = \{x\}$ is clopen in $X$ and thus $V \in D^-$. On the other hand, any clopen upset $W$ of $X$ containing $V$ must be cofinite. We can write

$$V = W - W'$$

where $W' = W - \{x\}$ is also a clopen upset of $X$. There is no smallest choice for $W$ as $\uparrow V = \{x, y\}$ is not open and thus not in $D$.

However, if we look for difference chains for $V$ relative to the lattice of closed subsets of $X$, then we have a least choice of difference chain, namely $V = K_1 - K_2$ where $K_1 = \uparrow V$ and $K_2 = K_1 - V$.

We show that there is an algorithm for deriving, for each element of the Booleanization of a distributive lattice, a difference chain of closed upsets (cf. Theorem 14), and that, in the case of a co-Heyting algebra, this provides a difference chain of the form (4) for each element of its Booleanization (cf. Corollary 17). We show that the difference chain thus obtained is of minimum length and is element-wise contained in any other such sequence (cf. Proposition 15). For this reason we will call it the canonical difference chain of closed upsets for the Boolean element in question. Recall that a subset $S \subseteq P$ of a poset is said to be convex provided $x \leq y \leq z$ with $x, z \in S$ implies $y \in S$.

**Definition 10.** If $P$ is a poset, $S \subseteq P$, and $p \in P$, then we say that $p_1 < p_2 < \cdots < p_n$ in $P$ is an alternating sequence of length $n$ for $p$ (with respect to $S$) provided

(a) $p_i \in S$ for each $i \in \{1, \ldots, n\}$ which is odd;
(b) $p_i \not\in S$ for each $i \in \{1, \ldots, n\}$ which is even;

(c) $p_n = p$.

Further, we say that $p \in P$ has degree $n$ (with respect to $S$), written $\deg_S(p) = n$, provided $n$ is the largest natural number $k$ for which there is an alternating sequence of length $k$ for $p$. In particular, if there is no alternating sequence for $p$ with respect to $S$ (i.e. if $p \in P - \uparrow S$) then $\deg_S(p) = 0$.

Notice that an element of finite degree is of odd degree if and only if it belongs to $S$. Also, if $S$ is convex, then every element of $S$ has degree 1, while every element of $\uparrow S - S$ has degree 2. In general, there will be non-empty subsets of posets with respect to which no element has finite degree. However, that is not the case for clopen subsets of Priestley spaces.

**Proposition 11.** Let $X$ be a Priestley space and $V$ a clopen subset of $X$. Then every element of $X$ has finite degree with respect to $V$.

**Proof.** The elements of the Booleanization of a distributive lattice $D$ may be written as a finite disjunctions of differences of elements from $D$. Thus, if $V$ is a clopen subset of a Priestley space $X$, then there is an $m$ so that we may write

$$V = \bigcup_{i=1}^{m} (U_i - W_i),$$

where $U_i, W_i \subseteq X$ are clopen upsets of $X$. In particular, since each $U_i - W_i$ is convex and, by the Pigeonhole Principle, there is no alternating sequence with respect to $V$ of length strictly greater than $2m$.

**Lemma 12.** Let $X$ be a Priestley space and $V$ a clopen subset of $X$. Let

$$K_1 = \uparrow V, \quad K_2 = \uparrow (\uparrow V - V).$$

Then, for each $i \in \{1, 2\}$, $K_i$ is closed and

$$K_i = \{x \in X \mid \deg_V(x) \geq i\} = \uparrow \{x \in X \mid \deg_V(x) = i\}.$$

**Proof.** By Proposition 11, every element of $X$ has a finite degree. Also if $x \leq y$, then it is clear that $\deg_V(x) \leq \deg_V(y)$. Furthermore, by Proposition 7, we have that $\uparrow K = \uparrow \min(K)$ for any closed set $K$. Now since both $V$ and
$\uparrow V - V$ are closed, it suffices to show that the elements of $\min(V)$ have
degree 1, and the elements of $\min(\uparrow V - V)$ have degree 2. It is clear that
deg$_V(x) = 1$ for any $x \in \min(V)$. Now suppose $x \in \min(\uparrow V - V)$. Since
$x \in \uparrow V$, there is $x' \in V$ with $x' \leq x$. Since $x \not\in V$, this is an alternating
sequence of length 2 for $x$. On the other hand, if $x_1 < x_2 < \cdots < x_n = x$ is
an alternating sequence for $x$, then $x_2 \in \uparrow V - V$ and thus $x \not\in \min(\uparrow V - V)$
unless $n = 2$ and $x_2 = x$. Thus deg$_V(x) = 2$ for any $x \in \min(\uparrow V - V)$. \hfill\(\square\)

**Corollary 13.** Let $X$ be a Priestley space, $V$ a clopen subset of $X$, and
$G_1 \supseteq G_2 \supseteq \cdots \supseteq G_{2p}$ a sequence of closed upsets in $X$ satisfying

$$V = G_1 - (G_2 - (\cdots - (G_{2p-1} - G_{2p})\cdots)).$$ \hfill(6)

If $K_1$ and $K_2$ are as defined in Lemma 12 then

$$K_1 \subseteq G_1, \quad K_2 \subseteq G_2, \quad \text{and} \quad G_1 - G_2 \subseteq K_1 - K_2.$$ \hfill(7)

**Proof.** By (6), we have $V \subseteq G_1$. Also, since $G_1$ is an upset we have $K_1 =
\uparrow V \subseteq G_1$. Now, since $G_1 - G_2 \subseteq V$ we have $G_1 - V \subseteq G_2$ and as $G_2$
is an upset, it follows that $\uparrow (G_1 - V) \subseteq G_2$. Also, $K_1 \subseteq G_1$ implies $K_2 =\n\uparrow (K_1 - V) \subseteq \uparrow (G_1 - V)$ and thus, $K_2 \subseteq G_2$. In particular, we have $G_1 - G_2 \subseteq
V - K_2 \subseteq K_1 - K_2$. \hfill\(\square\)

An iteration of Lemma 12 leads to the main result of this section.

**Theorem 14.** Let $X$ be a Priestley space and $V$ a clopen subset of $X$. Define
a sequence of subsets of $\uparrow V$ as follows:

$$K_1 = \uparrow V, \quad K_{2i} = \uparrow (K_{2i-1} - V), \quad \text{and} \quad K_{2i+1} = \uparrow (K_{2i} \cap V),$$

for $i \geq 1$. Then, $K_1 \supseteq K_2 \supseteq \cdots$ is a decreasing sequence of closed upsets of
$X$ and, for every $n \geq 1$, we have

$$K_n = \{x \in X \mid \deg_V(x) \geq n\} = \uparrow \{x \in X \mid \deg_V(x) = n\}.$$ \hfill(7)

In particular,

$$V = \bigcup_{i=1}^{m} (K_{2i-1} - K_{2i}) = K_1 - (K_2 - (\cdots (K_{2m-1} - K_{2m})\cdots)),$$ \hfill(8)

where $2m - 1 = \max\{\deg_V(x) \mid x \in V\}$. 11
Proof. Note that if (7) holds, then (8) holds since $K_{2i-1} - K_{2i}$ will consist precisely of the elements of $V$ of degree $2i - 1$ and since each element of $V$ has an odd degree less than or equal to the maximum degree achieved in $V$.

For the first statement and for (7), the proof proceeds by induction on the parameter $i$ used in (8). The case $i = 1$ is exactly Lemma 12. For the inductive step, suppose the statements hold for $n \leq 2i$ and notice that

$$K_{2i+1} = \uparrow(K_{2i} \cap V)$$

and

$$K_{2i+2} = \uparrow(K_{2i+1} - V)$$

are in fact the sets $K_1$ and $K_2$ of Lemma 12 when we apply it to the Priestley space $X' = K_{2i}$ and its clopen subset $V' = K_{2i} \cap V$. Thus, to complete the proof, it suffices to notice that, for every $x \in X'$, we have $\deg_{V'}(x) = \deg_V(x) + 2i$.

Using Corollary 13 we can now prove the minimality of the chain in (8).

**Proposition 15.** Let $X$ be a Priestley space and $V \subseteq X$ be a clopen subset of $X$. Let $G_1 \supseteq G_2 \supseteq \cdots \supseteq G_{2p}$ be a decreasing sequence of closed upsets of $X$ satisfying

$$V = \bigcup_{i=1}^p (G_{2i-1} - G_{2i}) = G_1 - (G_2 - (\cdots (G_{2p-1} - G_{2p})\cdots)).$$

(9)

Then, taking $(K_i)_{i \geq 1}$ as defined in Theorem 14, we have $p \geq m$ and, for every $n \in \{1, \ldots, p\}$, the following inclusions hold:

$$K_{2n-1} \subseteq G_{2n-1}, \ K_{2n} \subseteq G_{2n}, \text{ and } \bigcup_{i=1}^n (G_{2i-1} - G_{2i}) \subseteq \bigcup_{i=1}^n (K_{2i-1} - K_{2i}).$$

(10)

**Proof.** We proceed by induction on $n$. The case $n = 1$ is the content of Corollary 13. Now suppose that (10) holds for a certain $n \in \{1, \ldots, p\}$. As in the proof of Theorem 14, we consider the new Priestley space $X' = K_{2n}$ and its clopen subset $V' = X' \cap V$. Setting $G'_i = X' \cap G_i$ for each $i \in \{1, \ldots, p\}$, we obtain a decreasing sequence of closed upsets of $X'$ that form a difference chain for $V'$. On the other hand, by the induction hypothesis we have

$$\bigcup_{i=1}^n (G'_{2i-1} - G'_{2i}) = \left(\bigcup_{i=1}^n (G_{2i-1} - G_{2i})\right) \cap K_{2n} \subseteq \left(\bigcup_{i=1}^n (K_{2i-1} - K_{2i})\right) \cap K_{2n} = \emptyset$$

so that the first $2n$ sets do not contribute to the writing of $V'$ as a difference. It follows that the sequence $G'_{2n+1} \supseteq G'_{2n+2} \supseteq \cdots \supseteq G'_{2p}$ is a difference chain.
of closed upsets of $X'$ for $V'$. Now applying Corollary 13 to this sequence,
we see that

$$K_{2n+1} = \uparrow(K_{2n} \cap V) = \uparrow V' \subseteq G'_{2n+1} \subseteq G_{2n+1}$$

and

$$K_{2n+2} = \uparrow(K_{2n+1} \setminus V) = \uparrow(\uparrow V' \setminus V') \subseteq G'_{2n+2} \subseteq G_{2n+2}.$$ 

We also obtain that $(G_{2n+1} - G_{2n+2}) \cap K_{2n} = G'_{2n+1} - G'_{2n+2} \subseteq K_{2n+1} - K_{2n+2}$. 

On the other hand, by Theorem 14, we have

$$(G_{2n+1} - G_{2n+2}) - K_{2n} \subseteq V - K_{2n} = \bigcup_{i=1}^{n}(K_{2i} - K_{2i}).$$

We thus conclude that

$$\bigcup_{i=1}^{n+1}(G_{2i-1} - G_{2i}) \subseteq \bigcup_{i=1}^{n+1}(K_{2i-1} - K_{2i})$$

as required for the inductive step.

Note that, for $V = \bigcup_{i=1}^{m}(K_{2i-1} - K_{2i})$ as in Theorem 14, each of the
unions $\bigcup_{i=1}^{j}(K_{2i-1} - K_{2i})$ is an open subset of $X$, while $\bigcup_{i=j}^{m}(K_{2i-1} - K_{2i})$
is closed ($j = 1, \ldots, m$). In the next example, we illustrate Theorem 14 in a

case where $m = 2$, $(K_1 - K_2)$ is not closed, and $(K_3 - K_4)$ is not open.

**Example 16.** Let $X = \mathbb{N} \cup \{x, y\}$ be the same topological space as in Example 9, that is, the one-point compactification by $y$ of the discrete space 
$\mathbb{N} \cup \{x\}$, but now ordered by $1 < 2 < \cdots < y$ and $1 < x < y$. The order on 
$X$ and the dual lattice are depicted below:
Consider the clopen subset $V = \mathbb{N} \cup \{y\}$ of $X$. Then, we have:

$$K_1 = X, \quad K_2 = \{x, y\}, \quad K_3 = \{y\}, \quad \text{and} \quad K_4 = \emptyset.$$ 

So $K_1-K_2 = \mathbb{N}$ is open but not closed and vice versa for $K_3-K_4 = \{y\}$.

As a consequence of Theorem 6, Theorem 14 and Proposition 15 we have the following corollary.

**Corollary 17.** Let $D$ be a co-Heyting algebra and $b \in D^-$. Define the following sequence (recall Definition 5):

$$a_1 = \lfloor b \rfloor, \quad a_{2i} = \lfloor a_{2i-1} - b \rfloor, \quad \text{and} \quad a_{2i+1} = \lfloor a_{2i} \land b \rfloor, \quad \text{for} \ i \geq 1.$$ 

Then, the sequence $\{a_i\}_{i \geq 1}$ is decreasing, and there exists $m \geq 1$ such that $a_{2m+1} = 0$. For the least such $m$ we have

$$b = a_1 - (a_2 - \ldots (a_{2m-1} - a_{2m})\ldots) \quad (11)$$

and, for every other writing

$$b = c_1 - (c_2 - \ldots (c_{2p-1} - c_{2p})\ldots)$$

as a difference chain with $c_1 \geq \cdots \geq c_{2p}$ in $D$, we have $p \geq m$, $c_i \geq a_i$ for $i \in \{1, \ldots, 2p\}$, and for each $n \leq p$ we have $\bigvee_{i=1}^n (c_{2i-1} - c_{2i}) \leq \bigvee_{i=1}^n (a_{2i-1} - a_{2i}).$

We have a order dual algorithm for getting difference chains for Boolean elements over Heyting algebras. To obtain these sequences we use the floor function $\lfloor \cdot \rfloor : D^- \to D$ (its existence is the order dual of Proposition 4).

**Corollary 18.** Let $D$ be a Heyting algebra and $b \in D^-$. Define a sequence of elements in $D$ as follows:

$$a_1 = \lceil -b \rceil, \quad a_{2i} = \lceil a_{2i-1} \lor b \rceil, \quad a_{2i+1} = \lceil a_{2i} \lor -b \rceil, \quad \text{for} \ i \geq 1.$$ 

Then, the sequence $\{a_i\}_{i \geq 0}$ is increasing, and there exists $m \geq 1$ such that $a_{2m+1} = 1$. For the least such $m$ we have

$$b = a_{2m} - (a_{2m-1} - \ldots (a_2 - a_1)\ldots)$$

14
and, for every other writing

\[ b = c_{2p} - (c_{2p-1} - (\ldots (c_2 - c_1)\ldots)) \]

as a difference chain with \( c_1 \leq \cdots \leq c_{2p} \) in \( D \), we have \( p \geq m \), \( c_i \leq a_i \) for \( i \in \{1, \ldots, 2p\} \), and for each \( n \leq p \) we have \( \bigvee_{i=1}^{n}(c_{2i-1} - c_{2i}) \geq \bigvee_{i=1}^{n}(a_{2i-1} - a_{2i}) \).

Proof. Apply Corollary \ref{cor:coHeyting} to the co-Heyting algebra \( D' = \{\neg a \mid a \in D\} \) to get, for any \( b \in B = D^- = (D')^- \), a difference chain of elements in \( D' \):

\[ d_1 = [b], \quad d_{2i} = [d_{2i-1} - b], \quad \text{and} \quad d_{2i+1} = [d_{2i} \land b], \]

where the ceiling function is the one of \( D' \). Since \( a - b = \neg b - (\neg a) \), by Proposition \ref{prop:ceiling}, we may write \( b \in B \) as

\[ b = \neg d_{2m} - (\neg d_{2m-1} - (\ldots (\neg d_2 - \neg d_1)\ldots)), \]

and we have

\[ \neg d_1 = [\neg b], \quad \neg d_{2i} = [\neg d_{2i-1} \lor b], \quad \text{and} \quad \neg d_{2i+1} = [\neg d_{2i} \lor \neg b], \]

since \( \neg [u] = [\neg u] \) for all \( u \in B \), where the ceiling function is the one for \( D' \) while the floor function is the one for \( D \). Each \( a_i \) in the statement of the corollary is the \( \neg d_i \) of the proof.

Since finite distributive lattices are co-Heyting algebras, Corollary \ref{cor:coHeyting} applies. Combined with the fact that every distributive lattice is the direct limit of its finite sublattices and that the Booleanization is the direct limit of the Booleanizations of these finite sublattices, we have a proof of the original observation by Hausdorff.

**Corollary 19.** Every Boolean element over any distributive lattice may be written as a difference chain of elements of the lattice.

The results of this section were proved using Priestley duality. This makes them non-constructive. However, we could have proved them in a point-free setting (with very similar proofs). We trust that anyone interested in the constructive aspect can see for themselves that this is the case. In Section \ref{sec:point-free} we will continue in the point-free setting of so-called canonical extensions as this makes the more involved proofs there simpler.
4. Preliminaries on canonical extensions

Here we provide the required information on canonical extensions. For further details, please see [10] and [12].

Canonical extensions. Let \( D \) be a distributive lattice and \( X \) its dual Priestley space. Then, Priestley duality implies that the Stone map

\[
D \longrightarrow \text{Up}(X, \leq), \quad a \mapsto \hat{a}
\]

is an embedding of \( D \) into the complete lattice \( \text{Up}(X, \leq) \) of upsets of the poset underlying \( X \). An embedding into a complete lattice is called a completion, and canonical extension, first introduced by Jónsson and Tarski [18], comes about from the fact that the above completion can be uniquely characterized in abstract terms among all the completions of \( D \). Indeed, it is the unique completion \( e : D \hookrightarrow C \) (up to isomorphism) satisfying the following two properties:

1. **(dense)** Each element of \( C \) is a join of meets and a meet of joins of elements in the image of \( D \);
2. **(compact)** For \( S, T \subseteq D \) with \( \bigwedge e[S] \leq \bigvee e[T] \) in \( C \), there are finite subsets \( S' \subseteq S \) and \( T' \subseteq T \) with \( \bigwedge e[S'] \leq \bigvee e[T'] \).

Thus, instead of working with the dual space of a distributive lattice \( D \), we will work with its canonical extension, denoted \( D^\delta \). It comes with an embedding \( D \hookrightarrow D^\delta \), which is compact and dense in the above sense. As stated above this implies (modulo a non-constructive axiom) that \( D^\delta \) is isomorphic to \( \text{Up}(X, \leq) \), where \( X \) is the Priestley space of \( D \), and that the embedding \( e \) is naturally isomorphic to the map \( a \mapsto \hat{a} \). In what follows, to lighten the notation, we will assume (WLOG) that the embedding \( e \) is an inclusion so that \( D \) sits as a sublattice in \( D^\delta \).

Filter and ideal elements. Since \( D \) sits in \( D^\delta \) as the clopen upsets sit in \( \text{Up}(X, \leq) \), the join-closure of \( D \) in \( D^\delta \) corresponds to the lattice of open upsets of \( X \). One can show that these are in one-to-one correspondence with the ideals of \( D \). Thus we denote the join-closure of \( D \) in \( D^\delta \) by \( I(D^\delta) \) and call the elements of \( I(D^\delta) \) ideal elements of \( D^\delta \). Similarly the meet-closure of \( D \) in \( D^\delta \) corresponds to the lattice of closed upsets of \( X \) and one can show that these are in one-to-one correspondence with the filters of \( D \).
Accordingly we denote the meet-closure of $D$ in $D^\delta$ by $F(D^\delta)$ and call its elements filter elements of $D^\delta$. Note that, relative to the concepts of filter and ideal elements, the density property of $D^\delta$ states that every element of $D^\delta$ is a join of filter elements and a meet of ideal elements.

Another abstract characterization of $F(D^\delta)$ is that it is the free down-directed meet completion of $D$. As such it is uniquely determined by the following two properties [13, Proposition 2.1]:

(\textbf{filter dense}) Each element of $F(D^\delta)$ is a down-directed meet of elements from $D$;

(\textbf{filter compact}) For $S \subseteq F(D^\delta)$ down-directed and $a \in D$, if $\bigwedge S \leq a$, then there is $s \in S$ with $s \leq a$.

Notice, that in the particular case of a Boolean algebra $B$, the order on the dual space is trivial and thus $B^\delta$ is isomorphic to the full powerset of the dual space $X$ of $B$. Also, the ideal elements of $B^\delta$ correspond to all the opens of $X$ while the filter elements of $B^\delta$ correspond to all the closed subsets of $X$.

Ceiling functions at the level of canonical extensions. We saw in Proposition [4] that if $D$ is a co-Heyting algebra, then the inclusion of $D$ in its Booleanization, $D^\rightarrow$, has a lower adjoint $\lceil \rceil : D^\rightarrow \rightarrow D$. Here we will show, that on the level of canonical extensions any embedding has a lower adjoint with nice properties for filter elements.

Consider a situation where we have a Boolean algebra $B$ and a sublattice $D$ of $B$. Then the embedding of $D$ in $B$ extends to a complete embedding $D^\delta \hookrightarrow B^\delta$ [12, Theorem 3.2] which restricts to embeddings for the filter elements as well as for the ideal elements [12, Theorem 2.19]. Since the embedding is complete, it has both an upper and a lower adjoint, see Proposition [1]. We are interested in the lower adjoint, which we will study via the corresponding closure operator $\left( \right) : B^\delta \rightarrow D^\delta \hookrightarrow B^\delta$, cf. Proposition [2]. Thus we have, for $v \in B^\delta$,

$$\overline{v} = \bigwedge \{ u \in D^\delta \mid v \leq u \} = \bigwedge \{ y \in I(D^\delta) \mid v \leq y \},$$

where the second equality follows by the density property of $D^\delta$. We are particularly interested in the restriction of this closure operator to the filter elements.
Proposition 20. Let $B$ be a Boolean algebra and $D$ a sublattice of $B$, and let $(\overline{\cdot}): B^\delta \to B^\delta$ be the closure operator associated with $D \leq B$ as above. Then the following properties hold:

(a) for each $u \in B^\delta$, $\overline{u}$ is the least element of $D^\delta$ which lies above $u$;

(b) the map $(\overline{\cdot}): B^\delta \to B^\delta$ sends filter elements to filter elements;

(c) the map $(\overline{\cdot}): F(B^\delta) \to F(B^\delta)$ preserves down-directed meets.

Proof. Part (a) follows by the definition of adjoints. For (b), let $v \in F(B^\delta)$ and let $y \in I(D^\delta)$ with $v \leq y$. Since $y = \bigvee \{a \in D \mid a \leq y\}$, by compactness, there is $a_y \in D$ with $v \leq a_y \leq y$. Thus, we have

$$\overline{v} = \bigwedge \{y \in I(D^\delta) \mid v \leq y\} = \bigwedge \{a_y \in D \mid v \leq y\} \in F(D^\delta) \subseteq F(B^\delta).$$

For (c), let $S$ be a down-directed subset of $F(B^\delta)$ with $v = \bigwedge S$. By (b), $\overline{v} \in F(D^\delta)$, and thus $\overline{v} = \bigwedge \{a \in D \mid v \leq a\}$. Let $a \in D$ with $v \leq a$, then $\bigwedge S \leq a$ and by the filter compactness property of $F(B^\delta)$, there is $w_a \in S$ with $w_a \leq a$. Therefore we have

$$\bigwedge \{\overline{w} \mid w \in S\} \leq \bigwedge \{\overline{w_a} \mid v \leq a \in D\} \leq \bigwedge \{a \mid v \leq a \in D\} = \overline{v}.$$ 

On the other hand, since $v \leq w$ for each $w \in S$, by monotonicity of the closure operator, we also have $\overline{v} \leq \bigwedge \{\overline{w} \mid w \in S\}$ and thus the closure operator, restricted to $F(B^\delta)$, preserves down-directed meets.

Remark 21. Notice that if $B$ is the Booleanization of $D$, and $X$ is the Priestley space of $D$, then $B^\delta \cong \mathcal{P}(X)$, $D^\delta \cong \up(X, \leq)$, and the closure operator is the map $S \mapsto \up S$. Furthermore, Proposition 20(b) tells us that if $K \subseteq X$ is closed then so is $\up K$. That is, it is the canonical extension formulation of the second assertion in Proposition 7. We did not prove Proposition 20(c) in topological terms, but we could have. It says that if $\{W_i\}_{i \in I}$ is a down-directed family of closed subsets of a Priestley space, then

$$\bigcap_{i \in I} \up W_i = \up(\bigcap_{i \in I} W_i).$$

A statement that is not true in general for down-directed families of subsets of a poset.
Our main aim in this section is Theorem 27 and, more specifically, the Corollary 30, which will provide a simple proof of an important result in logic on words in Section 8. We will be working in the following general setting. We have a Boolean algebra $B$ and $D$ a sublattice of $B$. Recall that the Boolean subalgebra of $B$ generated by $D$ is (up to isomorphism) the Booleanization $D^{-}$ of $D$. Accordingly, we work with $D^{-}$ as being this generated subalgebra of $B$.

We start by formulating Theorem 14 in terms of canonical extensions and closure operators. As in the previous section, we let $( ) : B^{\delta} \to D^{\delta} \subseteq B^{\delta}$ be the closure operator which is the lower adjoint of the (complete) embedding $D^{\delta} \hookrightarrow B^{\delta}$. Given $b \in B$, we define the sequence $\{k_n\}_{n \geq 1}$ in $D^{\delta}$ as follows:

$$k_1 = \bar{b}, \quad k_{2n} = \frac{k_{2n-1} - b}{b}, \quad \text{and} \quad k_{2n+1} = \frac{k_{2n} \land b}{b}, \quad \text{for } n \geq 1. \quad (12)$$

Notice that $\{k_n\}_{n \geq 1}$ is a decreasing sequence of filter elements of $D^{\delta}$. Moreover, in the case where $B = D^{-}$, by Remark 21, this is exactly the canonical extension incarnation of the sequence $\{K_n\}_{n \in \mathbb{N}}$ for $V = \hat{b}$ in Theorem 14. We thus have the following:

**Theorem 22.** Let $D$ be a distributive lattice, $B = D^{-}$ and, for $b \in D^{-}$, consider the sequence $\{k_n\}_{n \geq 1}$ of $D^{\delta}$ as defined in (12). Then, there exists $m \geq 1$ such that $k_{2m+1} = 0$, and

$$b = k_1 - (k_2 - (\ldots (k_{2m-1} - k_{2m}) \ldots)) = \bigvee_{l=1}^{m} (k_{2l-1} - k_{2l}).$$

We will need the following slight generalization of Theorem 22.

**Corollary 23.** Let $B$ be a Boolean algebra and $D$ a sublattice of $B$. If $b \in D^{-}$ and $\{k_n\}_{n \geq 1}$ is the sequence defined in (12), then there exists $m \geq 1$ such that $k_{2m+1} = 0$ and

$$b = k_1 - (k_2 - (\ldots (k_{2m-1} - k_{2m}) \ldots)) = \bigvee_{l=1}^{m} (k_{2l-1} - k_{2l}).$$

**Proof.** The embedding $e : D \hookrightarrow B$ factors through $D^{-}$ so that $e = e_2 e_1$ where $e_1 : D \hookrightarrow D^{-}$ and $e_2 : D^{-} \hookrightarrow B$. These maps all lift to complete embeddings.
with $e^\delta = e_2^\delta e_1^\delta$ [12, Theorem 3.2 and Theorem 2.33], and these all have lower adjoints, $f$, $f_2$, and $f_1$, respectively. It follows that $f = f_1 f_2$. In a picture:

$$
\begin{array}{c}
D^\delta \\
\overset{e_1^\delta}{\leftarrow} \overset{e_2^\delta}{\rightarrow} (D^-)^\delta \\
\overset{f_1}{\leftarrow} \overset{f_2}{\rightarrow} B^\delta \\
\end{array}
= \begin{array}{c}
D^\delta \\
\overset{e^\delta}{\leftarrow} \overset{f}{\rightarrow} B^\delta \\
\end{array}
$$

Now the closure operator associated with $D^-$ is the map $c_1 = e_1^\delta f_1$, whereas the closure operator associated with $B$ is the map $c = e^\delta f$. The difference between Theorem 22 and Corollary 23 is that in the former $c_1$ is used to produce the sequence $\{k_n\}_{n \geq 1}$, and in the latter $c$ is used. Thus the corollary follows if we can show that for all $x \in (D^-)^\delta$ we have $c(x) = c_1(x)$, or including the action of the inclusion $e_2^\delta$, that $c(e_2^\delta(x)) = e_2^\delta(c_1(x))$. This is verified by the following calculation.

$$
c(e_2^\delta(x)) = (e^\delta fe_2^\delta)(x) = (e_2^\delta e_1^\delta f_1 f_2 e_2^\delta)(x) = e_2^\delta(c_1([f_2 e_2^\delta](x))) = e_2^\delta(c_1(x))
$$

since $f_2 e_2^\delta = \text{id}_{(D^-)^\delta}$.

Motivated by the application to logic on words presented in Section 8, we will now work in the following more general setting:

**Definition 24.** Let $B$ be a Boolean algebra and $I$ a directed partially ordered set. A directed family of closure operators on $B$ indexed by $I$ is a family of closure operators $\{\overline{\cdot}^i : B \to B\}_{i \in I}$ satisfying the following conditions:

(a) The meet-subsemilattices $S_i := \{b \in B \mid \overline{b}^i = b\}$ for $i \in I$ form an $I$-directed family of subsets of $B$. That is, $S_i \subseteq S_j$ whenever $i \leq j$.

(b) $D := \bigcup_{i \in I} S_i$ is a sublattice of $B$. That is, if $a, b \in S_i$ then there is $j$ with $a \vee b \in S_j$.

We start by showing that we have the following relationship between the closure operators $\overline{\cdot}^i$ and the one given by $D$.

**Proposition 25.** Let $B$ be a Boolean algebra and $\{\overline{\cdot}^i : B \to B\}_{i \in I}$ a directed family of closure operators on $B$. Then, for each $x \in B$, we have:

$$
\overline{x} = \bigwedge_{i \in I} x^i
$$

where the meet is taken in $B^\delta$. 

20
Proof. For each \( x \in B \), \( x \leq \overline{x}^i \in S_i \subseteq D \subseteq D^\delta \). Also \( \overline{x} \) is the least element of \( D^\delta \) above \( x \). Thus \( \overline{x} \leq \bigwedge_{i \in I} \overline{x}^i \). On the other hand, by Proposition \ref{prop:20}(b), since \( x \in B \subseteq F(B^\delta) \), we have \( \overline{x} \in F(D^\delta) \). That is, \( \overline{x} = \bigwedge \{ a \in D \mid x \leq a \} \).

Now let \( a \in D \) with \( x \leq a \). Then, since \( D^\delta = \bigcup_{i \in I} S_i \), there is \( j \in I \) with \( a \in S_j \). Now using the fact that \( x \leq a \) and the monotonicity of \((\cdot)^j\) we obtain

\[
\bigwedge_{i \in I} \overline{x}^i \leq \overline{a}^j = a.
\]

We thus have

\[
\bigwedge_{i \in I} \overline{x}^i \leq \bigwedge \{ a \in D \mid x \leq a \} = \overline{x}. \quad \square
\]

Now, for each \( i \in I \), we define the \( \{ c_{n,i} \}_{n \geq 1} \) as follows:

\[
c_{1,i} = b^i, \quad c_{2n,i} = c_{2n-1,i} - b^i, \quad \text{and} \quad c_{2n+1,i} = c_{2n,i} \land b^i \quad (13)
\]

Lemma 26. The following properties hold for the sequences as defined above:

(a) \( i \leq j \) implies \( k_n \leq c_{n,j} \leq c_{n,i} \) for all \( n \in \mathbb{N} \) and \( i, j \in I \);

(b) \( k_n = \bigwedge_{i \in I} c_{n,i} \) for all \( n \in \mathbb{N} \).

Proof. Define \( k_0 = c_{0,i} = 1 \) for all \( i \in I \). Also, define \( b_n = b \) for \( n \) odd and \( b_n = -b \) for \( n \) even then we have, for all \( n \geq 1 \), \( k_{n+1} = k_n \land b_n \) and similarly for the \( c_{n,i} \). Proceeding by induction on \( n \), we suppose \((a)\) holds for \( n \in \mathbb{N} \) and that \( i \leq j \). Note that since \( S_i \subseteq S_j \subseteq D \), we have \( \overline{x} \leq \overline{x}^j \leq \overline{x}^i \) for all \( x \in B \). Also, by the induction hypothesis \( k_n \leq c_{n,j} \leq c_{n,i} \), and thus we have

\[
k_n \land b_n \leq c_{n,j} \land b_n \leq c_{n,j} \land b_n^j \leq c_{n,j} \land b_n^i \leq c_{n,i} \land b_n^i.
\]

That is, \( k_{n+1} \leq c_{n+1,j} \leq c_{n+1,i} \) as required.

For \((b)\) again the case \( n = 0 \) is clear by definition and we suppose \( k_n = \bigwedge_{i \in I} c_{n,i} \). Then we have

\[
k_{n+1} = k_n \land b_n = (\bigwedge_{i \in I} c_{n,i}) \land b_n = \bigwedge_{i \in I} (c_{n,i} \land b_n).
\]

Now applying Proposition \ref{prop:20}(c) and then Proposition \ref{prop:25} we obtain
Now given \( i, j \in I \), since \( I \) is directed, there is \( k \in I \) with \( i, j \leq k \). By Lemma 26(a) we have \( c_{n,k} \leq c_{n,i} \). Combining this with the fact that \( S_j \subseteq S_k \) we obtain

\[
\frac{c_{n,k} \land b_n}{b_n} \leq \frac{c_{n,i} \land b_n}{b_n} \leq \frac{c_{n,i} \land b_n}{b_n},
\]

and thus

\[
k_{n+1} = \bigwedge_{(i,j) \in I^2} \frac{c_{n,i} \land b_n^j}{\bigwedge_{i \in I} c_{n,i} \land b_n^j} = \bigwedge_{k \in I} \frac{c_{n,k} \land b_n^k}{b_n} = \bigwedge_{k \in I} c_{n+1,k}.
\]

We are now ready to state and prove our main theorem.

**Theorem 27.** Let \( B \) be a Boolean algebra and \( \{ (\cdot)_i : B \to B \}_{i \in I} \) a directed family of closure operators on \( B \). For each \( b \in B \) let \( \{ k_n \}_{n \geq 1} \) be the sequence of filter elements as defined in (12) and \( \{ c_{n,i} \}_{n \geq 1, i \in I} \) be the sequence defined in (13). If \( b \in D^- \subseteq B \), then, there is \( m \in \mathbb{N} \) and \( i \in I \) so that, for each \( j \in I \) with \( i \leq j \) we have

\[
b = k_1 - (k_2 - \ldots - (k_{2m-1} - k_{2m}) \ldots) = \bigvee_{l=1}^m (k_{2l-1} - k_{2l})
\]

\[
= c_{1,j} - (c_{2,j} - \ldots - (c_{2m-1,j} - c_{2m}) \ldots) = \bigvee_{l=1}^m (c_{2l-1,j} - c_{2l,j}).
\]

**Proof.** Note that for \( b \in D^- \) the fact that the first line of the conclusion holds is simply the content of Corollary 23. The fact that the second line holds follows inductively from Lemma 28 below.

**Lemma 28.** Let \( b, b' \in B \) and \( v \in B^\beta \) be such that \( v \land k_{2l+1} = 0 \) and \( b' \leq k_{2l+2} \). Suppose \( b = v \lor (k_{2l+1} - k_{2l+2}) \lor b' \). Then there is an \( i \in I \) so that, for each \( j \in I \) with \( i \leq j \) we have \( b = v \lor (c_{2l+1,j} - c_{2l+2,j}) \lor b' \).

**Proof.** Since both \( v \) and \( k_{2l+1} - k_{2l+2} \) are below \( -k_{2l+2} \) we have \( b \leq -k_{2l+2} \lor b' \), or equivalently, \( b \land k_{2l+2} \leq b' \). Now by Lemma 26(b) we have \( b \land \bigwedge_{i \in I} c_{2l+2,i} \leq b' \) and by compactness there is an \( i_1 \in I \) so that for all \( j \in I \) with \( i_1 \leq j \) we
have \( b \land c_{2l+2,j} \leq b' \), or equivalently, \( b \leq \neg c_{2l+2,j} \lor b' \). Now, for each \( j \in I \) with \( i_1 \leq j \)

\[
b = (-k_{2l+1} \land b) \lor (k_{2l+1} \land b) = v \lor (k_{2l+1} \land b)
\]

\[
\leq v \lor (k_{2l+1} \land (\neg c_{2l+2,j} \lor b')) = v \lor (k_{2l+1} \land \neg c_{2l+2,j}) \lor (k_{2l+1} \land b')
\]

\[
\leq v \lor (k_{2l+1} - c_{2l+2,j}) \lor b'
\]

\[
\leq v \lor (k_{2l+1} - k_{2l+2}) \lor b' = b.
\]

Consequently, for each \( j \in I \) with \( i_1 \leq j \) we have \( b = v \lor (k_{2l+1} - c_{2l+2,j}) \lor b' \).

Now, since \( c_{2l+2,j} = \overline{c_{2l+1,j}} - b' \geq c_{2l+1,j} - b \), and thus, \( b \geq c_{2l+1,j} - c_{2l+2,j} \),
using also the inequality \( c_{2l+1,j} \geq k_{2l+1} \) given by Lemma 26(a), we may deduce

\[
b = v \lor b \lor b' \geq v \lor (c_{2l+1,j} - c_{2l+2,j}) \lor b'
\]

\[
\geq v \lor (k_{2l+1,j} - c_{2l+2,j}) \lor b' = b.
\]

It then follows that for all \( j \in I \) with \( j \geq i_1 \) we have

\[
b = v \lor (c_{2l+1,j} - c_{2l+2,j}) \lor b'.
\]

**Remark 29.** Notice that Corollary 19, stating that Boolean elements over a lattice are difference chains of elements of the lattice, can also be seen as a consequence of Theorem 27. Let \( D \) be any distributive lattice and \( B \) its Booleanization. For each finite sublattice \( D' \) of \( D \), the embedding \( D' \hookrightarrow D \hookrightarrow B \) has an upper adjoint \( g' : B \to D' \) given by \( g'(b) = \wedge\{a \in D' \mid b \leq a\} = \min\{a \in D' \mid b \leq a\} \). Thus, Theorem 27 applies and we get Corollary 19. In fact, in this way, we obtain more information as we see that the minimum length chain in \( D \) is equal to the minimum length chain in \( F(D^p) \), or equivalently, in the lattice of closed upsets of the dual space of \( D \). In turn, this is the same as the maximum length of difference chains in the dual with respect to the clopen corresponding to the given element.

In Section 8 we will give an application of the following consequence of Theorem 27, which needs its full generality.
Corollary 30. Let $B$ be a Boolean algebra and $\{\overline{\cdot}^i : B \to B\}_{i \in I}$ a directed family of closure operators on $B$. Let $B' \leq B$ be a Boolean subalgebra closed under each of the closure operators $\overline{\cdot}^i$ for $i \in I$. Then,

$$ (D \cap B')^+ = D^- \cap B', $$

where we view the Booleanization of any sublattice of $B$ as the Boolean subalgebra of $B$ that it generates.

Proof. Since $D \cap B'$ is contained in both of the Boolean algebras $D^-$ (also viewed as a subalgebra of $B$) and $B'$, the Booleanization of $D \cap B'$ is contained in their intersection.

For the converse, let $b \in D^- \cap B'$. By Theorem 27, there exists an index $j$ so that $b$ can be written as a difference chain

$$ b = c_{1,j} - (c_{2,j} - (\cdots - (c_{2m-1,j} - c_{2m,j}) \cdots)), $$

where $c_{1,j} = b^j$, $c_{2n,j} = c_{2n-1,j} - b^j$ and $c_{2n+1,j} = c_{2n,j} \wedge b^j$, for $n \geq 1$.

But then, by hypothesis that $B'$ is closed under $\overline{\cdot}^j$ and a straightforward induction argument, it follows that $c_{1,j} \geq \cdots \geq c_{2m,j}$ is a chain in $[B']^j \subseteq D \cap B'$. Thus, $b$ belongs to $(D \cap B')^-$.

Remark 31. We remark that the closure of $B'$ under the operators $\overline{\cdot}^i$ for $i \in I$ implies that the closure operator $\overline{\cdot}^i : B' \to B'$ on $B'$, whose image is $S'_i = B' \cap S_i$, is such that $\{S'_i\}_{i \in I}$ is an $I$-directed family of subsets of $B'$. Moreover, $D' = \bigcup_{i \in I} S'_i$ is precisely the distributive lattice $D \cap B'$.

We give an example to show that the conclusion of Corollary 30 is by no means true in general.

Example 32. Let $B = \mathcal{P}(\{a, b, c\})$ be the eight-element Boolean algebra. Further, let $D$ be the sublattice generated by $\{a\}$ and $\{a, b\}$ and let $B'$ be the Boolean subalgebra generated by $\{b\}$. Then $B$ is, up to isomorphism, the Booleanization of $D$, and thus $D^- \cap B' = B'$, whereas $D \cap B' = (D \cap B')^-$ is the two-element Boolean subalgebra of $B$.

In order to formulate the application to the theory of formal languages, we will need some concepts from logic on words.
6. Preliminaries on formal languages and logic on words

Formal languages. An alphabet is a finite set $A$, a word over $A$ is an element of the free $A$-generated monoid $A^*$, and a language is a set of words over some alphabet. For a word $w \in A^*$, we use $|w|$ to denote the length of $w$, that is, if $w = a_1 \ldots a_n$ with each $a_i \in A$, then we have $|w| = n$. Given a homomorphism $f : A^* \to M$ into a finite monoid $M$, we say that a language $L \subseteq A^*$ is recognized by $f$ provided there is a subset $P \subseteq M$ such that $L = f^{-1}(P)$, or equivalently, if $L = f^{-1}(f[L])$. The language $L$ is recognized by a finite monoid $M$ provided there is a homomorphism into $M$ recognizing $L$. Finally, a language is said to be regular if it is recognized by some finite monoid. Notice that the set of all regular languages forms a Boolean algebra. Indeed, if a language is recognized by a given finite monoid then so is its complement, and if $L_1$ and $L_2$ are recognized, respectively, by $M_1$ and $M_2$, then $L_1 \cap L_2$ is recognized by the Cartesian product $M_1 \times M_2$. (The more commonly encountered definitions of ‘regular language’ refer to finite automata or regular expressions; one can show that these are indeed the same as the regular languages defined here.)

The following well-known technical result [28] will be needed in Section 8.

**Lemma 33.** Let $f : A^* \to B^*$ be a homomorphism. Then the forward image under $f$ of a regular language over $A$ is a regular language over $B$.

We are interested in languages defined by first-order formulas of logic on words which we briefly introduce now. For further details please see [29, Chapter II].

Syntax of first-order logic on words. Fix an alphabet $A$. We denote first-order variables by $x, y, z, x_1, x_2, \ldots$. First-order formulas are inductively built as follows. For each letter $a \in A$, we consider a letter predicate, also denoted by $a$, which is unary. Thus, for any variable $x$, $a(x)$ is an (atomic) formula. A $k$-ary numerical predicate is a function $R : \mathbb{N} \to \mathcal{P}(\{1, \ldots, n\}^k)$ satisfying $R(n) \subseteq \{1, \ldots, n\}^k$ for every $n \in \mathbb{N}$. That is, $R$ is an element of the Boolean algebra $\Pi_{n \in \mathbb{N}} \mathcal{P}(\{1, \ldots, n\}^k)$. When we fix a set $\mathcal{R}$ of numerical predicates, we will assume it forms a Boolean subalgebra of $\Pi_{n \in \mathbb{N}} \mathcal{P}(\{1, \ldots, n\}^k)$. Each $k$-ary numerical predicate $R$ and any sequence $x_1, \ldots, x_k$ of first-order variables define an (atomic) formula $R(x_1, \ldots, x_k)$. Finally, Boolean combinations of formulas are formulas, and if $\varphi$ is a formula and $x_1, \ldots, x_k$ are distinct variables, then $\forall x_1, \ldots, x_k \varphi$ is a formula. In order to simplify the notation,
we usually also consider the quantifier $\exists$: the formula $\exists x_1, \ldots, x_k \varphi$ is an abbreviation for $\neg \forall x_1, \ldots, x_k \neg \varphi$. As usual in logic, we say that a variable $x$ occurs freely in a formula provided it is not in the scope of a quantifier that quantifies over $x$, and a formula is said to be a sentence provided it has no free variables. Quantifier-free formulas are those that are Boolean combinations of atomic formulas.

Semantics of first-order logic on words. Let us fix an alphabet $A$ and a set of numerical predicates $\mathcal{R}$. To each non-empty word $w = a_1 \ldots a_n \in A^*$ with $a_i \in A$, we associate the relational structure $M_w = (\downarrow n, A \cup \mathcal{R})$, where $\downarrow n = \{1, \ldots, n\}$, $a^w = \{i \in \downarrow n \mid a_i = a\}$, for each $a \in A$, and $R^w = R(n)$, for each $R \in \mathcal{R}$. Models of first-order sentences are words, while models of formulas with free variables are the so-called structures. For a set of distinct variables $x = \{x_1, \ldots, x_k\}$, an $x$-structure is an element of $A^* \times (\downarrow |w|)^x$. We identify maps from $x$ to $\downarrow |w|$ with $k$-tuples $i = (i_1, \ldots, i_k) \in (\downarrow |w|)^k$. Given a word $w \in A^*$ and a vector $i = (i_1, \ldots, i_k) \in (\downarrow |w|)^x$, we denote by $w_{x=i}$ the $x$-structure based on $w$ equipped with the map given by $i$. Moreover, if $x = \{x_1, \ldots, x_k\}$ and $y = \{y_1, \ldots, y_\ell\}$ are disjoint sets of variables, $i = (i_1, \ldots, i_k) \in (\downarrow |w|)^x$ and $j = (j_1, \ldots, j_\ell) \in (\downarrow |w|)^y$, then $w_{x=i, y=j}$ denotes the $z$-structure $w_{z=k}$, where $z = x \cup y$ and $k = (i_1, \ldots, i_k, j_1, \ldots, j_\ell)$.

We denoted the set of all $x$-structures by $A^* \otimes x$. Let $\varphi(x)$ be a formula all of whose free variables are in $x$. A model of $\varphi(x)$ is an $x$-structure that satisfies $\varphi(x)$, using the standard interpretation of quantifiers in formulas. We denote by $L_{\varphi(x)}$ the set of all models of the formula. We will say that the formula defines this set of $x$-structures.

Example 34. Suppose $A = \{a, b\}$. The sentence $\varphi = \exists x, y \; (x < y \land a(x) \land b(y))$ is read: “there are positions $x$ and $y$ such that $x$ comes before $y$ and there is an $a$ at position $x$ and a $b$ at position $y$”. Thus, $\varphi$ defines the regular language $L_{\varphi}$ given by the regular expression $A^*aA^*bA^*$.

It is worth remarking how this informal intuitive interpretation matches up with our definitions of the syntax and semantics of first-order formulas. The subformula $x < y$ is the binary numerical predicate, which, formally speaking, maps each $n \in \mathbb{N}$ to the set $\{(i, j) \mid 1 \leq i < j \leq n\} \subseteq (\downarrow n)^2$. Since $\varphi$ has no free variables, $L_{\varphi}$ is a set of $\emptyset$-structures—that is, simply a set of words over $A$.

We give an example involving formulas that contain free variables.
Example 35. Again, let \( A = \{a, b\} \). Let us define a numerical predicate, which we denote informally by \( x = \frac{y}{2} \). Formally, this maps each \( n \in \mathbb{N} \) to the set \( \{(i, 2i) \mid 1 \leq i \leq n/2\} \). The quantifier-free formula \( \psi_1(x, y) \) given by

\[
\left( x = \frac{y}{2} \right) \land b(y)
\]

defines a set of \( \{x, y\}\)-structures. It is convenient to think of each such structure as a word over \( A \) in which each of the two variables labels a position in the word. Thus, for example, \( ab(a, x)ba(b, y)aa \) represents the structure \( ababaaa \) in \( A \times \{x\} \). Let \( \psi_2 \) be the formula \( \exists x \, \psi_1(x, y) \). This formula has a single free variable \( y \), says that \( y \) labels an even-numbered position, which contains the letter \( b \). So, for example, \( ababa(b, y)aa \in L_{\psi_2(y)} \).

Further, let \( \psi_3 \) denote the formula \( \exists x, y \, \exists x \, \psi_1(x, y) \). Its set of models \( L_{\psi_3} \) consists of all words that contain the letter \( b \) in an even-numbered position. This is again a regular language, given by the regular expression \((A^2)^*AbA^*\).

Regular languages of structures. We now give a more formal account of a device we used informally in the examples above. We fix a set of distinct variables \( x = \{x_1, \ldots, x_k\} \). Then, \( 2^x \) is isomorphic to the powerset \( \mathcal{P}(x) \).

There is a natural embedding of the set of all \( x \)-structures into the free monoid \((A \times 2^x)^*\). Indeed, to an \( x \)-structure \( w_{x=i} \), where \( i = (i_1, \ldots, i_k) \), we may assign the word \((a_1, S_1) \cdots (a_n, S_n)\), where \( w = a_1 \cdots a_n \) with each \( a_i \in A \) and, for \( \ell \in \{1, n\} \), \( S_\ell = \{x_j \in x \mid i_j = \ell\} \). It is not hard to see that this mapping defines an injection \( A^* \otimes x \hookrightarrow (A \times 2^x)^* \). Moreover, an element \((a_1, S_1) \cdots (a_n, S_n)\) of \((A \times 2^x)^*\) represents an \( x \)-structure under this embedding precisely when the non-empty sets among \( S_1, \ldots, S_n \) form a partition of \( x \). From hereon, we view \( A^* \otimes x \) as a subset of \((A \times 2^x)^*\) without further mention.

Since we view \( x \)-structures as words over \( A \times 2^x \), we can talk about regular languages of structures. Moreover, it is easy to see that the set \( A^* \otimes x \) of all \( x \)-structures is itself a regular language: To see this, let \( x \in x \) be any variable. Let \( N = \{0, 1, m\} \) be the three-element monoid in which 0 is absorbent, 1 is the identity, and \( m^2 = 0 \), and let \( f : (A \times 2^x)^* \rightarrow N \) be the unique homomorphism satisfying \( f(a, S) = m \) if \( x \in S \) and \( f(a, S) = 1 \) otherwise. This homomorphism recognizes, via \( \{m\} \), the set of words \((a_1, S_1) \cdots (a_n, S_n)\) such that the variable \( x \) occurs exactly once among the \( S_i \). The set of structures \( A^* \otimes x \) is the intersection of these languages over
all $x \in \mathbf{x}$, and thus is a regular language itself, since the family of regular
languages is closed under finite intersection.

As a result, a language $L$ of structures is a regular language if and only
if it consists of all the $\mathbf{x}$-structures in some regular language $L'$ over $A \times 2^x$,
because we then have $L = L' \cap (A^* \otimes \mathbf{x})$. We will make use of this observation
in the next section.

Example 36. We consider again the formulas of Example 35. The set
of structures defined by
$\psi_1(x, y)$ is a non-regular language: Consider the
homomorphism $\alpha : (A \times 2^{\{x,y\}})^* \rightarrow \{c, d\}^*$ defined by mapping every letter of
the form $(e, \emptyset)$ to $c$ and all other letters to $d$. Then $\alpha[L] = \{c^k dc^k dc^\ell\}_{k, \ell \geq 0}$. If
$L$ were regular, then Lemma 33 implies that so is $\alpha[L]$. However, well-known
elementary techniques of the theory of automata show that this is not the
case. (See, e.g. [29, Chapter I], or a standard textbook like Sipser [26].)

On the other hand, the set of structures defined by
$\psi_2(y)$ is regular: It is
just the language given by the regular expression $(A^2)^* A(b, \{y\}) A^*$.

Fragments of first-order logic. Formulas will always be considered up to se-
monic equivalence, even if not explicitly said. We denote by $\text{FO}[\mathcal{N}]$ the set
of all first-order sentences with arbitrary numerical predicates (up to semantic
equivalence). For formulas whose free variables are in $\mathbf{x}$, we will write
$\text{FO}_x[\mathcal{N}]$. And for a set $\mathcal{R}$ of numerical predicates, $\text{FO}[\mathcal{R}]$ denotes the set
of first-order sentences using numerical predicates from $\mathcal{R}$. Notice that, as
a Boolean algebra, $\text{FO}[\mathcal{N}]$ is naturally equipped with a partial order, which
in turn may be characterized in terms of semantic containment: $\varphi \leq \psi$ if
and only if $L_\varphi \subseteq L_\psi$. For this reason, we will identify formulas and the
 corresponding languages of models switching freely between $\varphi$ and $L_\varphi$. In
particular, we see $\text{FO}[\mathcal{N}]$ as a Boolean subalgebra of $\mathcal{P}(A^*)$ and $\text{FO}_x[\mathcal{N}]$ as
a Boolean subalgebra of $\mathcal{P}(A^* \otimes \mathbf{x})$.

Quantifier alternation. We can measure the complexity of first-order formu-
las by the minimum number of alternations of quantifiers that is needed to
express them in prenex-normal formula, that is, in the form
$$\psi = Q_1 x_1 \ldots Q_m x_m \varphi(x_1, \ldots, x_m),$$
(14)
where $\varphi$ is a quantifier-free formula, $Q_1, \ldots, Q_m \in \{\forall, \exists\}$ and $Q_\ell = \forall$ if and
only if $Q_{\ell+1} = \exists$ for each $\ell = 1, \ldots, m - 1$. It is a well-known fact that
for every first-order formula there is a semantically equivalent one in prenex-
normal form. For $m \geq 1$ and a set of numerical predicates $\mathcal{R}$, $\Pi_m[\mathcal{R}]$ consists
of all the sentences of $\text{FO}[\mathcal{R}]$ that are semantically equivalent to a sentence of the form (14) where $Q_1 = \forall$. This is the alternation hierarchy over $\mathcal{R}$. In particular, we have $\Pi_m[\mathcal{R}] \subseteq \Pi_\ell[\mathcal{R}]$ whenever $m \leq \ell$. Similarly, $\Sigma_m[\mathcal{R}]$ denotes the set of all sentences that are semantically equivalent to a sentence of the form (14) with $Q_1 = \exists$, and $\Sigma_m[\mathcal{R}] \subseteq \Sigma_\ell[\mathcal{R}]$ whenever $m \leq \ell$. It is not hard to see that both $\Pi_m[\mathcal{R}]$ and $\Sigma_m[\mathcal{R}]$ are closed under disjunction and conjunction but not under negation in general. In other words, $\Pi_m[\mathcal{R}]$ and $\Sigma_m[\mathcal{R}]$ are lattices, but not Boolean algebras. We denote by $B\Pi_m[\mathcal{R}]$ and by $B\Sigma_m[\mathcal{R}]$ the Boolean algebras generated by $\Pi_m[\mathcal{R}]$ and by $\Sigma_m[\mathcal{R}]$, respectively, that is,

$$B\Pi_m[\mathcal{R}] = (\Pi_m[\mathcal{R}])^- \quad \text{and} \quad B\Sigma_m[\mathcal{R}] = (\Sigma_m[\mathcal{R}])^-.$$

Clearly, we have $B\Pi_m[\mathcal{R}] = B\Sigma_m[\mathcal{R}] \subseteq \Pi_{m+1}[\mathcal{R}], \Sigma_{m+1}[\mathcal{R}]$.

When $\mathcal{R}$ consists of just the order relation $<$, then the languages defined by first-order sentences are all regular. In this case, the alternation hierarchy has been extensively studied. An outstanding open problem is to determine effectively whether a given first-order definable regular language belongs to the $m^{th}$ level of the hierarchy. The only cases for which this is known is when $m \leq 2$ (see [24]).

In this paper we are only concerned with the first level of the hierarchy over different bases of numerical predicates. For notational convenience, we will work with the fragment $\Pi_1[\mathcal{R}]$, although everything we prove for $\Pi_1[\mathcal{R}]$ admits a dual statement for $\Sigma_1[\mathcal{R}]$. Every formula of $\Pi_1[\mathcal{R}]$ is of the form $\psi = \forall x \varphi(x)$, for some quantifier-free formula $\varphi(x)$. Inside $\Pi_1[\mathcal{R}]$, we classify formulas according to the size of $x$: we let $\Pi_1^k[\mathcal{R}]$ consist of all equivalence classes of such formulas for which there is a representative $\psi$ for which $x$ has $k$ variables. We remark that $\Pi_1^k[\mathcal{R}]$ is closed under conjunction, since the formulas $\forall x \varphi(x) \land \forall x \psi(x)$ and $\forall x (\varphi(x) \land \psi(x))$ are semantically equivalent. However, it is easy to see that, in general, $\Pi_1^k[\mathcal{R}]$ fails to be closed under disjunction, as we now show.

Example 37. Let $\varphi(x) = a(x)$ and $\psi(x) = b(x)$. Then, $\forall x \varphi(x)$ defines the language $a^*$, while $\forall x \psi(x)$ defines the language $b^*$ and thus these are both in $\Pi_1^1[\mathcal{N}]$. The disjunction $\forall x \varphi(x) \lor \forall x \psi(x)$ defines the language $a^* \cup b^*$, while $\forall x (\varphi(x) \lor \psi(x))$ defines the language $\{a, b\}^*$. Indeed, one can show that $\forall x \varphi(x) \lor \forall x \psi(x)$ is not in $\Pi_1^1[\mathcal{N}]$ while it is in $\Pi_2^1[\mathcal{N}]$ as witnessed by the sentence $\forall x_1, x_2 (\varphi(x_1) \lor \psi(x_2))$. 

29
We will use \( \text{Reg} \) to denote the set of first-order sentences \( \varphi \) for which \( L_\varphi \) is a regular language. We will also use \( \text{Reg} \) to denote the set of numerical predicates for which the associated language of structures is regular. In [2] it is shown that

\[
\text{FO}[\mathcal{N}] \cap \text{Reg} = \text{FO}[\text{Reg}].
\]

In other words, if a first-order sentence defines a regular language \( L \), then \( L \) can be defined using only regular numerical predicates. This is proved as a consequence of results in circuit complexity.

It is conjectured that this equality holds at each level of the alternation hierarchy, in other words that, for \( m \geq 0 \), we have

\[
\text{BΠ}_m[\mathcal{N}] \cap \text{Reg} = \text{BΠ}_m[\text{Reg}].
\]

(15)

For \( m > 1 \), this question is open. In Section 8, we will use the results of Section 5 to provide a proof of the case \( m = 1 \). This was first proved in [20], and a different proof appears in [30]. The proof in [20] relies on some hard results in semigroup theory and ideas from circuit complexity, and the one in [30] on Ramsey theory coupled with the algebra of finite semigroups. In contrast, our argument is entirely different and quite elementary.

Example 38. In an earlier example, we saw that the sentence

\[
\exists x, y \ (x = \frac{y}{2} \land b(y))
\]

defines the regular language \( (A^2)^*AbA^* \) over \( A = \{a, b\} \). On the other hand \( x = \frac{y}{2} \) is a non-regular numerical predicate. The dual equality of (15) for existential fragments and \( m = 0 \) implies that the same language can be defined by a sentence of \( \Sigma_1[\text{Reg}] \). Such a sentence is given by

\[
\exists y \ ((y \equiv 0 \pmod{2}) \land b(y)),
\]

in which the numerical predicate is regular.

7. A lattices-and-duality perspective on logic on words

Universal quantifiers as adjoints. Again, we fix a finite alphabet \( A \) and a set of variables \( x \). We consider the projection map given by
\[ \pi : A^* \otimes x \rightarrow A^*, \quad w_{x=1} \mapsto w. \]

This gives rise, via the duality between sets and complete and atomic Boolean algebras, to the complete embedding of Boolean algebras

\[ \pi^{-1} = ( ) \otimes x : \mathcal{P}(A^*) \hookrightarrow \mathcal{P}(A^* \otimes x), \quad L \mapsto \pi^{-1}(L) = L \otimes x \]

This embedding, being a complete homomorphism between complete lattices, has an upper adjoint which we may call \( \forall \) (and a lower adjoint \( \exists \)). These are given by

\[ \forall : \mathcal{P}(A^* \otimes x) \rightarrow \mathcal{P}(A^*) \]
\[ K \mapsto \forall K = \max \{ L \in \mathcal{P}(A^*) \mid L \otimes x \subseteq K \} \]
\[ = \{ w \in A^* \mid \forall i \in \{1, \ldots, |w|\}^x, \ w_{x=1} \in K \} \]
\[ = (\pi[K^c])^c \]

and similarly

\[ \exists : \mathcal{P}(A^* \otimes x) \rightarrow \mathcal{P}(A^*) \]
\[ K \mapsto \exists K = \min \{ L \in \mathcal{P}(A^*) \mid K \subseteq L \otimes x \} \]
\[ = \{ w \in A^* \mid \exists i \in \{1, \ldots, |w|\}^x, \ w_{x=1} \in K \} \]
\[ = \pi[K] \]

As is well-known in categorical logic, \( \forall \) and \( \exists \) are the semantic incarnations of the classical universal and existential quantifiers, respectively. Explicitly, for the universal quantifier, when \( K = L_{\varphi(x)} \) is definable by a formula \( \varphi(x) \), we have

\[ \forall L_{\varphi(x)} = \{ w \in A^* \mid \forall i \in \{1, \ldots, |w|\}^x, \ w_{x=1} \models \varphi(x) \} = L_{\forall x \varphi(x)}. \quad (16) \]

**Recognition, model theoretic types, and duality.** Let \( \text{LOG}_x \) denote the Boolean algebra of formulas with free variables in \( x \) relative to some logic (up to semantic equivalence), e.g. \( \text{LOG}_x = \text{FO}_x[A] \). Further, let \( X \) be the dual space
of this Boolean algebra and let \( \mathbf{x}\text{-Str} \) denote the set (or class) of intended models for this logic, e.g. \( \mathbf{x}\text{-Str} = A^* \otimes \mathbf{x} \). Then we have a mapping

\[
\text{typ} : \mathbf{x}\text{-Str} \rightarrow X \\
(M, i) \mapsto \{ \varphi(\mathbf{x}) | M \models \varphi(i) \}
\] (17)

which sends a model to the ultrafilter of formulas that it satisfies. In model theory, the image of this map is known as the type of the given model. In language theory, this kind of map is used to study the logic itself since it is encoded via recognition. That is, this map topologically recognizes the logic in the sense that the Boolean algebra \( \text{LOG}_x \) is isomorphic to the Boolean subalgebra of those \( L \subseteq \mathcal{P}(\mathbf{x}\text{-Str}) \) such that

\[
L = \text{typ}^{-1}(V)
\]

for some clopen \( V \subseteq X \).

In classical model theory, Gödel’s Completeness Theorem tells us that every type \( x \in X \) is realized (by some model). However, for this to be the case here, we would need to consider not just the finite models of our logic but the so-called pseudo-finite models. Then the map in (17) becomes surjective and topological methods may be applied. This is closely related to recognition by profinite monoids as studied in language theory, where these ideas are combined with those of monoid recognition as described above. See [14, 19] for a study of the connections between model theoretic type theory and recognition in language theory and [11] for a study of the connections between Stone duality and recognition in language theory.

Here, we will be able to work just with the finite models. This is because the logic fragment we want to consider here consists of the quantifier-free formulas in \( \text{FO}_x[N] \) and, as we will see in Corollary 40, these form a complete and atomic Boolean algebra and thus fall within the discrete duality between sets and complete and atomic Boolean algebras.

**Quantifier-free formulas.** Consider a set of distinct variables \( \mathbf{x} = \{x_1, \ldots, x_k\} \).

We will give an algebraic characterization of the languages of the form \( L_{\varphi(\mathbf{x})} \) for \( \varphi(\mathbf{x}) \) a quantifier-free formula whose free variables are in \( \mathbf{x} \).

We first provide a characterization of these languages via discrete duality, bringing out the connection between recognition and the notion of types from model theory as described above. For this purpose, we say that \( L \subseteq X \) is set theoretically recognized by \( f : X \rightarrow Y \) provided there is a subset \( P \subseteq Y \) with \( L = f^{-1}(P) \).
We will need the following notation: For \( w = a_1 \ldots a_n \) with each \( a_i \in A \) and \( i = (i_1, \ldots, i_k) \), we write \( w(i) \) for the tuple \((a_{i_1}, \ldots, a_{i_k})\). For a vector of letters \( a = (a_1, \ldots, a_k) \in A^x \), we denote by \( a(x) \) the conjunction \( a_1(x_1) \land \cdots \land a_k(x_k) \).

**Lemma 39.** Let \( K \subseteq A^* \otimes x \). Then, \( K \) is given by a quantifier-free first-order formula over \( x \) if and only if is it set theoretically recognized by the map

\[
c_A : A^* \otimes x \to \mathbb{N}^{k+1} \times A^k, \quad w_{x=1} \mapsto (|w|, i, w(i)).
\]

**Proof.** First suppose that \( K = c_A^{-1}(P) \) for some \( P \subseteq \mathbb{N}^{k+1} \times A^k \). For each \( a \in A^k \) and \( n \in \mathbb{N} \), let

\[
R^a(n) = \{ i \in (\downarrow n)^k \mid (n, i, a) \in P \}.
\]

Then \( R^a \) is a \((k\text{-ary})\) numerical predicate for each \( a \in A^k \) and it is not difficult to see that \( c_A^{-1}(P) = \varphi(x) \) for

\[
\varphi(x) = \bigvee_{a \in A^k} (a(x) \land R^a(x)).
\]

On the other hand, for \( a \in A \) and \( i \in \{1, \ldots, k\} \)

\[
L_{a(x_i)} = c_A^{-1}(\mathbb{N}^{k+1} \times \{ a \in A^k \mid a_i = a \})
\]

and for \( R \subseteq \mathbb{N}^{m+1} \) an \( m\text{-ary} \) numerical predicate, and (not necessarily distinct) variables \( y_1, \ldots, y_m \in x \), we have

\[
L_{R(y_1, \ldots, y_m)} = c_A^{-1}(R' \times A^k)
\]

where \((i_1, \ldots, i_k) \in R'(n)\) if and only if \((j_1, \ldots, j_m) \in R(n)\) where \( j_s = i_t \) if and only if \( y_s = x_t \).

Now, we obtain an algebraic characterization of the quantifier free formulas. That is, a characterization in the form of recognition by a monoid rather than just by a set. For this purpose, let \( \varepsilon \notin A \) be a new symbol and denote \( A_\varepsilon = A \cup \{ \varepsilon \} \). We consider the homomorphism \( \Theta_x : (A \times 2^x)^* \to (A_\varepsilon \times 2^x)^* \) given by

\[
\Theta_x(a, S) = \begin{cases} 
(a, S), & \text{if } S \neq \emptyset; \\
(\varepsilon, S), & \text{if } S = \emptyset.
\end{cases}
\]
Notice that, given $x$-structures $v_{x=i}$ and $w_{x=j}$, we have

$$\Theta_x(v_{x=i}) = \Theta_x(w_{x=j}) \iff |v| = |w|, \ i = j, \text{ and } v(i) = w(j). \quad (18)$$

Using this observation, it is straightforward to show:

**Corollary 40.** Let $L \subseteq A^* \otimes x$ be a language. Then, the following are equivalent:

(a) $L$ is definable by a quantifier-free formula;

(b) $L = \Theta_x^{-1}(\Theta_x[L])$;

(c) there is a subset $P \subseteq A^*_x \otimes x$ such that $L = \Theta_x^{-1}(P)$.

In particular, the set of all quantifier-free formulas of $\text{FO}_x[N]$ forms a complete and atomic Boolean algebra.

**8. An application to Logic on Words**

In this section we combine Corollary 30 and Remark 31 to prove the equality

$$\mathcal{B}\Pi_1[N] \cap \text{Reg} = \mathcal{B} \Pi_1[\text{Reg}]. \quad (19)$$

The idea is the following. Combining the fact that universal quantification may be seen as an adjoint and our algebraic characterization of quantifier-free formulas we obtain a directed family of adjunctions on $\mathcal{P}(A^*)$ with joint image equal to $\Pi_1[N]$ allowing us to fit into the setting of Theorem 27. Finally we show that these adjunctions restrict correctly to the regular fragment thus allowing us to apply Corollary 30 and Remark 31 thereby concluding that (19) holds.

Let $x$ be a set of $k$ variables. Universal quantification (as an adjoint) and recognition of quantifier-free formulas are based on the following two maps, respectively

$$A^* \xleftarrow{\pi} A^* \otimes x \xrightarrow{\Theta_x} A^*_x \otimes x.$$
Dually this gives rise to

\[
\begin{align*}
\mathcal{P}(A^*) &\xrightarrow{\pi^{-1}} \mathcal{P}(A^* \otimes x) & \mathcal{P}(A^* \otimes x) &\xleftarrow{\Theta_x^{-1}} \mathcal{P}(A^*_\epsilon \otimes x) \\
\exists & & (\Theta_x[(\ )^c])^c &\forall \Theta x[ ( )]
\end{align*}
\]

In particular, we have a (correct) composition of adjunctions as follows

\[
\begin{align*}
\mathcal{P}(A^*) &\xrightarrow{\pi^{-1}} \mathcal{P}(A^* \otimes x) & \mathcal{P}(A^* \otimes x) &\xleftarrow{\Theta_x^{-1}} \mathcal{P}(A^*_\epsilon \otimes x) \\
\exists & & (\Theta_x[(\ )^c])^c &\forall \Theta x[ ( )]
\end{align*}
\]

That is,

\[f_k = \Theta_x[\pi^{-1}(\ )]: \mathcal{P}(A^*) \xhookrightarrow{\pi^{-1}} \mathcal{P}(A^*_\epsilon \otimes x): \forall(\Theta_x^{-1}(\ ) = g_k\]

is an adjunction and, combining quantification as adjunction with the description of quantifier-free formulas in Corollary \ref{corollary:quantifier_free_forms}, we have \(L \subseteq A^*\) is in \(\Pi^1_k[\mathcal{N}]\) if and only if \(L = \forall(\Theta_x^{-1}(P)) = g_k(P)\) for some \(P \subseteq A^*_\epsilon \otimes x\). That is, \((f_k, g_k)\) is an adjunction with associated closure operator

\[\lceil L \rceil_k := g_k f_k(L) = \forall\Theta_x^{-1}(\Theta_x[L \otimes x]),\]

and \(\text{Im}(\lceil \quad \rceil_k) = \text{Im}(g_k) = \Pi^1_k[\mathcal{N}]\). Notice that, since \(\lceil \quad \rceil_k\) is a closure operator on \(\mathcal{P}(A^*)\), we have

\[L \subseteq [K]_k \iff [L]_k \subseteq [K]_k,\]

for every \(K, L \subseteq A^*\). Therefore, \([L]_k\) may be seen as the best over-approximation of \(L\) by a language definable in \(\Pi^1_k[\mathcal{N}]\). Accordingly, we are in the situation of Theorem \ref{theorem:approximation} with

\[B = \mathcal{P}(A^*)\text{ and } D = \bigcup_{k \in \mathbb{N}} \text{Im}(g_k) = \bigcup_{k \in \mathbb{N}} \Pi^1_k[\mathcal{N}] = \Pi_1[\mathcal{N}].\]

We now aim to apply Corollary \ref{corollary:regular_languages} with \(B' = \text{Reg}\), the Boolean algebra of all regular languages over \(A\). This is possible given the following fact.
Lemma 41. For each $k \in \mathbb{N}$, if $L \subseteq A^*$ is regular, then so is $\lceil L \rceil_k$.

Proof. Fix $k \in \mathbb{N}$ and suppose $L \subseteq A^*$ is regular. We proceed through the four maps whose composition defines $\lceil \cdot \rceil_k$.

Claim 1. $L \otimes x$ is regular.

Note that if $\mu : A^* \rightarrow M$ is a finite monoid recognizing $L$, then $\pi^* : (A \times 2^x)^* \rightarrow A^*$ composed with $\mu$, where $\pi^*$ is the homomorphism extending the projection of $A \times 2^x$ onto $A$, recognizes $L \otimes x$ once we restrict to structures. As we remarked earlier, the set of all $x$-structures is itself regular so it follows from closure under intersection that $L \otimes x$ is itself regular.

Claim 2. $\Theta_x[L \otimes x]$ is regular.

This follows from the previous claim and Lemma 33.

Claim 3. $\Theta^{-1}_x(\Theta_x[L \otimes x])$ is regular.

This is immediate as the inverse image with respect to a homomorphism between free monoids of a regular language is always regular: If $\Theta_x[L \otimes x]$ is recognized by $f' : (A_x \times 2^x)^* \rightarrow M'$ then the composition $f' \circ \Theta_x : (A \times 2^x)^* \rightarrow M'$ recognizes $\Theta^{-1}_x(\Theta_x[L \otimes x])$.

Claim 4. $\forall(\Theta^{-1}_x(\Theta_x[L \otimes x]))$ is regular.

As observed in Section 7, the upper adjoint $\forall$ is given by $K \mapsto (\pi[K^c])^c$ where $\pi : A^* \otimes x \rightarrow A^*$ is the restriction of $\pi^* : (A \times 2^x)^* \rightarrow A^*$ to structures. It follows that $\pi[K^c] = \pi^*[K^c \cap (A^* \otimes x)]$. Now, since $K = \Theta^{-1}_x(\Theta_x[L \otimes x])$ is regular, $K^c$ is also regular and $K^c \cap (A^* \otimes x)$ is regular. Further, it follows by Lemma 33 that $\pi^*[\Theta^{-1}_x(\Theta_x[L \otimes x])]^c \cap A \otimes x$ is regular. Finally, we conclude that its complement $\forall(\Theta^{-1}_x(\Theta_x[L \otimes x])) = (\pi^*[\Theta^{-1}_x(\Theta_x[L \otimes x])]^c \cap A \otimes x)^c$ is regular as required.

As a consequence, Corollary 30 applies and we obtain:

Corollary 42. Considering each of the following Booleanizations as subalgebras of $\mathcal{P}(A^*)$, we have

$$(\Pi_1[\mathcal{N}] \cap \text{Reg})^- = (\Pi_1[\mathcal{N}])^- \cap \text{Reg}.$$

Finally, applying Remark 31 in this particular case, we see that

$$\Pi_1[\mathcal{N}] \cap \text{Reg} = \bigcup_{k \in \mathbb{N}} g_k f_k [\text{Reg}].$$

The languages in $g_k f_k [\text{Reg}]$ are exactly the languages $\lceil L \rceil_k$ for $L$ regular. By the proof of Lemma 41 we have that $\lceil L \rceil_k = \forall(\Theta^{-1}_x(P))$ where $P = \ldots$
$Θ_x[L \otimes x] \subseteq (A_x \times 2^x)^*$, and, by Claim 2 in particular, we have that $P$ is regular. That is, $[L]_k = L_{\forall x \varphi(x)}$ where the atomic formula $\varphi(x)$ is regular, or equivalently, $[L]_k \in \Pi_1[\text{Reg}]$.

On the other hand if $\varphi(x)$ is an atomic formula that is regular, then the arguments in Claims 3 and 4 show that $L_{\forall x \varphi(x)}$ is regular. Thus

$$g_k f_k[\text{Reg}] = \Pi_1^k[\text{Reg}]$$

and we have

$$(\Pi_1[\text{Reg}])^- = \left( \bigcup_{k \in \mathbb{N}} \Pi_1^k[\text{Reg}] \right)^- = (\Pi_1[\mathcal{N}] \cap \text{Reg})^- = (\Pi_1[\mathcal{N}])^- \cap \text{Reg}.$$

We thus obtain the desired result:

**Theorem 43.** The following equality holds:

$$B\Pi_1[\text{Reg}] = B\Pi_1[\mathcal{N}] \cap \text{Reg}.$$

### 9. Final remarks

Can the techniques of this paper be pushed further? A few possible directions suggest themselves. As we mentioned earlier, Theorem 43 is not new, but the techniques of our proof are completely different from what was used before.

In [30], a more general result is proved, concerning *modular quantifiers*: If $x$ is a set of $k$ variables and $0 \leq j < m$, $t > 0$, then we allow quantified formulas of the form

$$\exists^{(j,m,t)} x \varphi(x).$$

Such formulas are interpreted as follows: If $w \in A^*$, then $w \models \exists^{(j,m,t)} x \varphi(x)$ if the number of $k$-tuples $i$ such that

$$(w, i) \models \varphi(x)$$

is both congruent to $j$ modulo $m$ and at least $t$. Observe that with $m = t = 1$, this is just ordinary existential quantification. If $\varphi(x)$ is quantifier-free, then $\exists^{(j,m,t)} x \varphi(x)$ is called a *generalized* $\Sigma_1$-sentence. The main result of [30] is that the analogue of Theorem 43 holds for Boolean combinations of such
generalized $\Sigma_1$-sentences, as well as for Boolean combinations of ordinary $\Sigma_1$-sentences. It would be interesting to know whether the approach of the present paper can be used to give a different proof of this result.

What about rising higher in the alternation hierarchy? The identity

$$\text{FO}[\mathcal{N}] \cap \text{Reg} = \text{FO}[\text{Reg}],$$

proved in [2], is equivalent to well-known lower bounds results in circuit complexity, and the only proof known depends on the circuit complexity results. Yet this equality appears to be saying something fundamental, and rather simple, about automata and logic: If you can define a regular language with a sentence that uses unrestricted numerical predicates, then you can define it using just regular numerical predicates. Can the methods used here provide a different proof of this fact? One might first try to show that

$$\text{BII}_2[\mathcal{N}] \cap \text{Reg} \subseteq \text{FO}[\text{Reg}].$$

It should be cautioned that things are not so neat when we get to $\Pi_2$ formulas, and that this would require the development of new techniques.

Combining the two directions of generalization - from ordinary to modular quantifiers, and from $\Pi_1$ formulas to formulas with more levels of quantifier alternation - one might conjecture

$$\text{FOMOD}[\mathcal{N}] \cap \text{Reg} = \text{FOMOD}[\text{Reg}].$$

Here, $\text{FOMOD}[\mathcal{R}]$ represents all the formulas one can build, starting from numerical predicates in $\mathcal{R}$, using both modular quantification and Boolean operations. This conjecture is equivalent to the long-open question in circuit complexity of whether the complexity class $\text{ACC}^0$ is strictly contained in $\text{NC}^1$ (see [2] and [29]).

Acknowledgments

In addition to the funding received from the ERC, during the revision process, the first-named author was partially supported by the Center for Mathematics of the University of Coimbra (UID/MAT/00324/2019) and by the Center for Mathematics of the University of Porto (UID/MAT/00144/2019), which are funded by the Portuguese Government through FCT/MCTES and cofunded by the European Regional Development Fund through the Partnership Agreement PT2020.
Finally, we would like to thank the anonymous referee for the careful reading of the paper and for the useful suggestions that helped to improve the presentation of our work. We also want to thank Charles Paperman for fruitful discussions with the two last-named authors.

References

[1] B. Banaschewski and A. Pultr. Booleanization. Cahiers de Topologie et Géométrie Différentielle Catégoriques 37 (1), 41-60, 1996.

[2] D. Barrington, K. Compton, H. Straubing, and D. Thérien. Regular languages in $NC^1$. Journal of Computer and System Sciences 44(3):478-499, 1992.

[3] C. Borlido and M. Gehrke. A note on powers of Boolean spaces with internal semigroups. ArXiv e-prints, page arXiv:1811.12339, November 2018.

[4] O. Carton, D. Perrin, and J.-É. Pin. A survey on difference hierarchies of regular languages. Logical Methods in Computer Science 14(1), 2017.

[5] E. Casanovas and R. Farré. Omitting Types in Incomplete Theories. The Journal of Symbolic Logic 61(1), 236–245, 1996.

[6] C. C. Chen. Free Boolean extensions of distributive lattices. Nanta Math., 1:1–14, 1966/1967.

[7] B. A. Davey and H. A. Priestley. Introduction to Lattices and Order, 2nd edition. Cambridge University Press, 2002.

[8] L. L. Esakia. Topological Kripke models. Dokl. Akad. Nauk SSSR, 214:298–301, 1974.

[9] L. L. Esakia. Heyting algebras: Duality theory. “Metsniereba”, Tbilisi, 1985. (in Russian).

[10] M. Gehrke. Canonical extensions, Esakia spaces, and universal models. In Leo Esakia on duality in modal and intuitionistic logics, volume 4 of Outst. Contrib. Log., pages 9–41. Springer, Dordrecht, 2014.
[11] M. Gehrke, S. Grigorieff, and J.-É. Pin. Duality and equational theory of regular languages. In Automata, languages and programming. Part II, volume 5126 of Lecture Notes in Comput. Sci., pages 246–257. Springer, Berlin, 2008.

[12] M. Gehrke and B. Jónsson. Bounded distributive lattices expansions. Mathematica Scandinavica, 94(2):13–45, 2004.

[13] M. Gehrke and H. A. Priestley. Canonical extensions and completions of posets and lattices. Rep. Math. Logic, 43:133–152, 2008.

[14] S. van Gool and B. Steinberg. Pro-aperiodic monoids and model theory. Israel Journal of Mathematics, to appear.

[15] C. Glaßer, H. Schmitz, and V. Selivanov. Efficient algorithms for membership in Boolean hierarchies of regular languages. Theoret. Comput. Sci., 646:86–108, 2016.

[16] G. Grätzer and E. T. Schmidt. On the generalized Boolean algebra generated by a distributive lattice. Indag. Math., 20:547–553, 1959.

[17] F. Hausdorff. Set theory. Chelsea Publishing Company, New York, 1957. Translated by John R. Aumann, et al.

[18] B. Jónsson and A. Tarski. Boolean algebras with operators I. Amer. J. Math., 73:891–939, 1951.

[19] L. Libkin. Elements of finite model theory. Springer, Berlin, 2004.

[20] A. Maciel, P. Péladeau, and D. Thérien. Programs over semigroups of dot-depth one. Theoret. Comput. Sci., 245(1):135–148, 2000. Semigroups and algebraic engineering (Fukushima, 1997).

[21] H. M. MacNeille. Extension of a distributive lattice to a Boolean ring. Bull. Amer. Math. Soc., 45(6):452–455, 1939.

[22] A. Nerode. Some Stone spaces and recursion theory. Duke Math. J., 26:397–406, 1959.

[23] W. Peremans. Embedding of a distributive lattice into a Boolean algebra. Indag. Math., 60:73–81, 1957.
[24] T. Place and M. Zeitoun. Going higher in first-order quantifier alternation hierarchies on words. *Journal of the ACM*, 2018.

[25] H. A. Priestley. Representation of distributive lattices by means of ordered stone spaces. *Bull. London Math. Soc.*, 2:186–190, 1970.

[26] M. Sipser. *Introduction to the theory of computation, 2nd edition* Thomson, Boston, MA, 2006.

[27] M. H. Stone. Applications of the theory of Boolean rings to general topology. *Trans. Amer. Math. Soc.*, 41(3):375–481, 1937.

[28] H. Straubing. Recognizable sets and power sets of finite semigroups. *Semigroup Forum*, 18(4):331–340, 1979.

[29] H. Straubing. *Finite automata, formal logic, and circuit complexity*. Progress in Theoretical Computer Science. Birkhäuser Boston, Inc., Boston, MA, 1994.

[30] H. Straubing. Languages defined with modular counting quantifiers. *Information and Computation* 166(2):112–132, 2001.