Masterarbeit

New Approaches to the Union Closed Sets Conjecture

Nicolas Nagel

April 26, 2022

Contents

1 Introduction

1.1 The Problem

The aim of this thesis is to mainly study so called *union closed set families*, which are families $\mathcal{F}$ consisting of finitely many sets, such that if A and B are two sets from $\mathcal{F}$, then so is $A \cup B$. Such families thus posses an algebraic, order theoretic, but also combinatorial structure to them. The most famous open question about union closed families is the *union closed sets conjecture* and states the following.

Conjecture 1.1 (Union Closed Sets Conjecture)**.** Let $\mathcal{F}$ be a finite, union closed family containing at least one nonempty set. Then, there is an element x from the *ground set* over $\mathcal{F}$ (i.e. a superset of all sets from $\mathcal{F}$), such that

$$\#\{F \in \mathcal{F} : x \in F\} \geq \frac{1}{2} \cdot \#\mathcal{F}.$$

In words, any union closed family containing at least one nonempty set has an element that is contained in at least half of the sets of $\mathcal{F}$. While having a very simple statement, this is one of the major open problems in extremal combinatorics. Everything will be made more precise further down below.
Before starting with the technicalities, it is worthwhile to look at the "philosophical" sides of this conjecture. A survey giving more details on the conjecture and its history can be found in [10]. Since then, [1, 3, 9, 17, 38, 39] have further been published, studying the union closed sets conjecture or related topics. In this context, the collaborative effort in [24] should also be mentioned, which also resulted in a sizeable amount of new conjectures and research.

1.2 Difficulties

It seems surprising that the Union Closed Sets Conjecture 1.1, having such a simple statement, is still open after more then 40 years of extensive research. We give some reasons why it might be so hard.
First of all, union closed families are so broad that they comprise a large number of possible set families which allows for relatively sophisticated combinatorial substructures. Finding a technique that applies to all union closed families thus seems more complicated than initially thought. Adding to that, the property of being union closed only adds a small amount of structure to the set family. In [2] it was determined that

the number of union closed families over the ground set $[n]$ behaves asymptotically as

$$2^{(1+o(1))\binom{n}{n/2}},$$

the exact values for small n are collected in [25]. This alone gives a sizeable amount of possible set families (from a possible number of 2^{2^n}) to study for each $n \in \mathbb{N}$.
Secondly, while we do not assume much structure within our set families, we also do not ask very much of them: we only search for an element contained in at least half of the sets of the family (which seems rather difficult to pinpoint exactly, as will be made clear further down below). This makes an inductive approach, which certainly would be among the first methods that one might try for Conjecture 1.1, rather difficult. Indeed, trying induction on $\#\mathcal{F}$, it is surprisingly difficult to control the behaviour of elements from the ground set $x \in [n]$ with respect to the family $\mathcal{F}$, even after we extend $\mathcal{F}$ by only one or two more sets (whilest still being union closed).
Thirdly, continuing on the previous point, it seems very difficult to generalize Conjecture 1.1. In [24], many possible generalizations were given and shortly after disproven by counterexamples (many conjectures stated there, specializing Conjecture 1.1, were also later proven correct, e.g. in [1]). This suggests that one might better try to construct (or find) a counterexample to Conjecture 1.1. Many conditions on (minimal) counterexamples are already known.

1.3 Summary of this Thesis

Chapters 2 and 3 are a summary of already known results that we might only adapt to our situation. In chapter 4, except for those explicitly named there, we will explore new results and techniques in the study of union closed families and related structures. We start by discussing union closed families in general. In particular, we will investigate equivalent structures that also give us a deeper view into the inner structure of union closed families.
We then devote some time to the conjecture itself and its many faces. This also gives a comparison between what is already known and what this thesis aims to tackle.
We then introduce completely new techniques in the study of the conjecture. Specifically, we introduce weaker, related and generalized versions of the conjecture. Here we also deduce some results that already suggest on how to improve on what is known about the conjecture.
On the way, we pose new and unsolved questions related to the Union Closed Sets Conjecture. At the end, we summarize some open questions suitable for further research.

1.4 Notation

As is common, $\mathbb{N}$, $\mathbb{N}_0, \mathbb{Z}, \mathbb{R}$ denote the natural numbers, nonnegative integers, integers and reals. The power set of a set X will be denoted by $\mathcal{P}(X)$. For $n \in \mathbb{N}$ we write $[n] := \{1, ..., n\}$ (with $[0] = \emptyset$) and also shorten $\mathcal{P}(n) := \mathcal{P}([n])$ (with $\mathcal{P}(0) = \{\emptyset\}$). For a set family $\mathcal{F}$ we use

$$\bigcup \mathcal{F} = \bigcup_{F \in \mathcal{F}} F, \quad \bigcap \mathcal{F} = \bigcap_{F \in \mathcal{F}} F.$$

We will also use the symbol $\dot{\cup}$ to emphasize a disjoint union. For the symmetric difference of two sets we use the symbol Δ, i.e.

$$A \Delta B = (A \setminus B) \cup (B \setminus A) = (A \cup B) \setminus (A \cap B).$$

The cardinality of a set X will be denoted by $\#X$.
For two functions f, g (with positive values) we use the following Landau notation:

- $f = O(g)$ if there is a constant $C > 0$ such that for all sufficiently large n it holds $f(n) \leq C \cdot g(n)$;

- $f = o(g)$ if $\lim\limits_{n \longrightarrow \infty} \frac{f(n)}{g(n)} = 0$;

- $f = \Omega(g)$ if $g = O(f)$.

The natural logarithm will be denoted by ln, all other logarithms just with log and their respective base.

2 Characterizations of Set Families

This chapter introduces some basic notions from the theory of set families and aims to give an overview on how they are connected with each other. For this, it is certainly sensible to introduce a notion of isomorphy for set families.

Definition 2.1. Let X, Y be sets and let $\mathcal{A} \subseteq \mathcal{P}(X), \mathcal{B} \subseteq \mathcal{P}(Y)$ be two set families. We say that $\mathcal{A}$ and $\mathcal{B}$ are **isomorphic** if there is a bijection $f \colon X \to Y$ such that the induced map

$$\tilde{f} \colon \mathcal{P}(X) \to \mathcal{P}(Y), A \mapsto f(A)$$

is a bijection from $\mathcal{A}$ to $\mathcal{B}$.

If no explicit *ground set* X is provided, we may assume that a given set family $\mathcal{A}$ is defined over $\bigcup \mathcal{A}$.

2.1 Union Closed Families

We introduce some common terminology from the literature on set families (see [10]).

Definition 2.2. Let $n \in \mathbb{N}$ and $\mathcal{F} \subseteq \mathcal{P}(n)$. We call the set family $\mathcal{F}$

(i) **nontrivial** if $\bigcup \mathcal{F} = [n]$ and $\bigcap \mathcal{F} = \emptyset$;

(ii) **separating** if for all $x, y \in [n], x \neq y$ there is an $F \in \mathcal{F}$ with $\#(F \cap \{x, y\}) = 1$;

(iii) **union closed** if for $A, B \in \mathcal{F}$ also $A \cup B \in \mathcal{F}$.

An element $i \in [n]$ is called **abundant** (with respect to the family $\mathcal{F}$) if

$$\#\{F \in \mathcal{F} : i \in F\} \geq \frac{1}{2} \cdot \#\mathcal{F}.$$

The element i is called **strictly abundant** if the above inequality holds strictly (i.e. "$>$" instead of just "$\geq$").

We will abbreviate "union closed set family" simply by *union closed family.* In the context of the above, $[n]$ is the *ground set* and we often speak of union closed families *over* $[n]$. If $\mathcal{F} \subseteq \mathcal{P}(n)$ is union closed and nontrivial, then by definition $[n] = \bigcup \mathcal{F} \in \mathcal{F}$

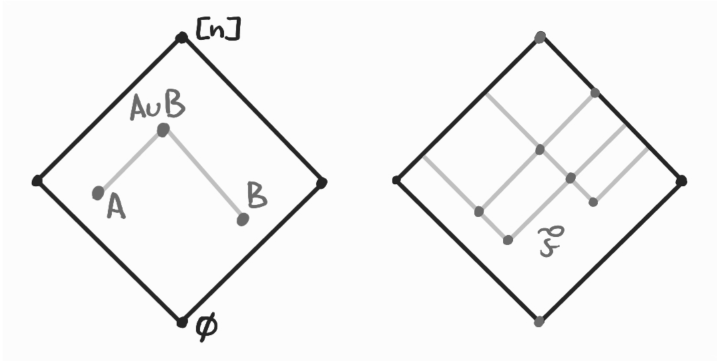

Figure 2.1: A visualization of union closed families

as the (finite) union of sets from $\mathcal{F}$. For what follows, we will investigate several other set structures and how they relate to union closed families. We first observe some immediate consequences.

Observation 2.3.

(i) A set family is separating if and only if the families $\mathcal{F}_x := \{F \in \mathcal{F} : x \in F\}, x \in [n]$ are pairwise distinct. Indeed, the separating property is equivalent to the statement that for all $x, y \in [n], x \neq y$ at least one of the sets $\mathcal{F}_x \setminus \mathcal{F}_y$ or $\mathcal{F}_y \setminus \mathcal{F}_x$ is nonempty. This is precisely the case if the families $\mathcal{F}_x, x \in [n]$ are pairwise distinct (but may be contained in each other).

(ii) For $\mathcal{F}$ a union closed family over $[n]$ and $X \subseteq [n]$ some subset, the derived families

$$\begin{aligned}
\mathcal{F}^1 &:= \{F \cap X : F \in \mathcal{F}\}, \\
\mathcal{F}^2 &:= \{F \setminus X : F \in \mathcal{F}\}, \\
\mathcal{F}^3 &:= \{F \in \mathcal{F} : F \cap X = \emptyset\}, \\
\mathcal{F}^4 &:= \{F \setminus X : F \in \mathcal{F}, X \subseteq F\}, \\
\mathcal{F}^5 &:= \{F \in \mathcal{F} : F \cap X = \emptyset, F \cup X \in \mathcal{F}\}
\end{aligned}$$

are all union closed, in particular for $X = \{x\}$ consisting of only one element. If $\mathcal{F}$ is also nontrivial then so are $\mathcal{F}^1$, $\mathcal{F}^2$ and $\mathcal{F}^4$ (over their respective ground sets $\bigcup \mathcal{F}^1 = X$ and $\bigcup \mathcal{F}^2 = \bigcup \mathcal{F}^4 = [n] \setminus X$), which in general is not the case for $\mathcal{F}^3$ and $\mathcal{F}^5$. If $\mathcal{F}$ is separating then the same holds for $\mathcal{F}^1$ and $\mathcal{F}^2$, but in general not for $\mathcal{F}^3$, $\mathcal{F}^4$ or $\mathcal{F}^5$. We skip the details here but these claims are straight forward to verify.

(iii) Let $\mathcal{F} \subseteq \mathcal{P}(n)$ be a separating, nontrivial, union closed family. Then $\mathcal{F}$ contains a set of size $n-1$. Indeed, assume otherwise and let F be a maximal set (with respect to set inclusion) in $\mathcal{F} \setminus \{[n]\}$. By assumption $\#F \leq n-2$, so there are distinct $x, y \in [n] \setminus F, x \neq y$. Since $\mathcal{F}$ is separating, there is a $G \in \mathcal{F}$ with $x \in G$ and $y \notin G$ (the other way around is analogous). In particular then $F \cup G \in \mathcal{F}$. Since $x \in F \cup G$ but $y \notin F \cup G$ we get $F \subsetneq F \cup G \subsetneq [n]$, contradicting maximality of F.

Definition 2.4. Let $n \in \mathbb{N}$ and $\mathcal{U} \subseteq \mathcal{P}(n)$. The family $\mathcal{U}$ is called an **up-set** if $A \subseteq B \subseteq [n]$ and $A \in \mathcal{U}$ implies $B \in \mathcal{U}$.

That is, up-sets are set families that are closed under taking supersets. They are also refered to as *(order/set) filters* (in $\mathcal{P}(n)$). Every up-set is union closed.

Example 2.5. We give some examples for union closed families.

(i) Trivially, $\mathcal{P}(n)$ and $\{\emptyset, [n]\}$ are union closed families over $[n]$. Of the two, only the first one is separating (for $n \geq 2$). A bit more interesting, for a $k = 1, ..., n$ the family
$$\mathcal{F} := \mathcal{P}(k) \cup \{[l] : l = k+1, ..., n\}$$
is a nontrivial, separating, union closed family over $[n]$. Intuitively, the union closedness of a set family would suggest that $\mathcal{F}$ would have to contain many "large" sets. However, this example shows that one has to be more careful with "largeness".

(ii) Denote by $\mathcal{N} := \{A \subseteq \mathbb{N} : \#A < \infty\}$ the family of finite subsets of $\mathbb{N}$. Define a linear order on $\mathcal{N}$ as follows: For $A, B \in \mathcal{N}, A \neq B$ define $A < B$, if

- $\max A < \max B$ or
- $\max A = \max B$ and $\max(A \Delta B) \in A$.

The order then starts
$$\emptyset < 1 < 12 < 2 < 123 < 23 < 13 < 3 < 1234 < 234 < 134 < 34 < 124 < 24 < ...$$
(here we omit the braces, e.g. 123 stands for the set $\{1, 2, 3\}$). For $m \in \mathbb{N}$ let $\mathcal{H}(m)$ be the first m sets according to this order. Then $\mathcal{H}(m) \subseteq \mathcal{P}(n)$ with $n = \lceil \log_2 m \rceil$ and $\mathcal{H}(m)$ is union closed. Indeed, this follows from $A \leq B \Rightarrow A \cup B \leq B$ for $A, B \in \mathcal{N}$, which can be deduced from the definition of the above order. The union closed family $\mathcal{H}(m)$ is commonly called the *Hungarian family*.
A similar construction results in the so called *Renaud-Fitina families* (for that see [10, 35]).

(iii) There are several other constructions of union closed families with peculiar properties. One of those is given in [33], with a generalization given in [10]. The following example is similar and was first given in [37] (using shorthand as above):
$$\begin{aligned}\{&123, 6789, 4589, 4567, 46789, 45689, 45678, 16789, 24589, 34567, 456789, 146789,\\ &245689, 345678, 1456789, 2456789, 3456789, 1236789, 1234589, 1234567,\\ &12456789, 13456789, 23456789, 12346789, 12345689, 12345678, 123456789\}.\end{aligned}$$
This family consists of 27 sets and is defined over the set [9]. Here, the elements $1, 2, 3$ appear in 13 sets, $4, 6, 8$ appear in 23 and $5, 7, 9$ appear in 20. Notice that the elements 1, 2 and 3 are contained in the unique set of the smallest cardinality, yet they are the least frequent elements among the set family, they are not even contained in at least half of all sets. More visually, the above family can be represented in a *Hasse diagram.*

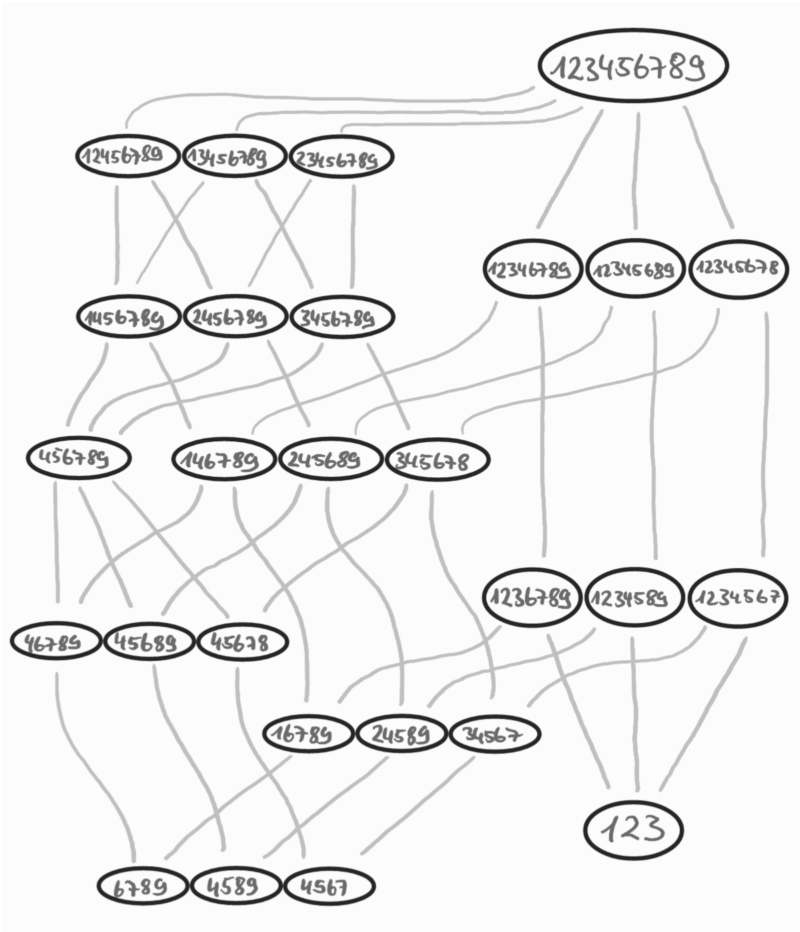

Figure 2.2: The Hasse diagram of the family from Example 2.5 (iii)

2.2 Intersection and Δ-Closed Families

If we want to study union closed families, it is natural to also consider set families which are closed under other set theoretic operations. This motivates the following definition.

Definition 2.6. Let $n \in \mathbb{N}$ and $\mathcal{F} \subseteq \mathcal{P}(n)$. We call the set family $\mathcal{F}$

(i) **intersection closed** if for $A, B \in \mathcal{F}$ also $A \cap B \in \mathcal{F}$;

(ii) **Δ-closed** if for $A, B \in \mathcal{F}$ also $A \Delta B \in \mathcal{F}$ (symmetric difference).

An element $i \in [n]$ is called **rare** (with respect to the family $\mathcal{F}$) if

$$\#\{F \in \mathcal{F} : i \in F\} \leq \frac{1}{2} \cdot \#\mathcal{F}.$$

The element i is called **strictly rare** if the above inequality holds strictly (i.e. "$<$" instead of just "$\leq$").

Again, we will abbreviate "intersection closed set family" to *intersection closed family*, and analogously for Δ-*closed family*. In measure and probability theory in particular, nonempty intersection closed families are also referred to as π-*systems* (see [7]). The following elementary theorem gives a connection between union and intersection closed families.

Theorem 2.7. Let $n \in \mathbb{N}$. Then the following sets are in bijection to each other:

$$\{\mathcal{F} \subseteq \mathcal{P}(n) : \mathcal{F} \text{ union closed}\} \to \{\mathcal{F} \subseteq \mathcal{P}(n) : \mathcal{F} \text{ intersection closed}\}$$
$$\mathcal{F} \mapsto \{[n] \setminus F : F \in \mathcal{F}\}.$$

This bijection preserves the property of being nontrivial.

Proof. Using de Morgan's rule, for every union closed family $\mathcal{F} \subseteq \mathcal{P}(n)$ the derived family $\{[n] \setminus F : F \in \mathcal{F}\}$ is intersection closed. By the same argument, being nontrivial is also preserved. This map is then a bijection, as the inverse is also given by taking complements of all member sets of an intersection closed family (resulting in a union closed family). ■

This shows that union closed families are in a certain sense equivalent (*cryptomorphic*) to intersection closed families. This allows us to translate statements for union closed families to statements about intersection closed families. In many contexts it is more common to work with intersection closed families instead of union closed families (anecdotally, mathematical structures, such as groups, vector spaces, σ-algebras and so on, tend to be better behaved with respect to intersections rather than unions).
In analogy to union closed families, we also define the corresponding concept of up-sets under the above bijection.

Definition 2.8. Let $n \in \mathbb{N}$ and $\mathcal{D} \subseteq \mathcal{P}(n)$. The family $\mathcal{D}$ is called a **down-set** if $A \subseteq B \subseteq [n]$ and $B \in \mathcal{D}$ implies $A \in \mathcal{D}$.

Down-sets are also refered to as *independence systems*, *abstract simplicial complexes* or *(order/set) ideals* (in $\mathcal{P}(n)$).
While Δ-closed families will not be all too important for us, they still have a property closely related to the later chapters of this thesis. For this, we need the following (also elementary) preparations.

Lemma 2.9. Let $\mathcal{F} \subseteq \mathcal{P}(n)$ be a Δ-closed family. Then $(\mathcal{F}, \Delta)$ is a (abelian) group with identity $\emptyset$ and $\#\mathcal{F} = 2^k$ for some $k = 0, 1, ..., n$.

Proof. Interpreting sets $X \subseteq [n]$ as $(0,1)$-strings (in the canonical way) we have the well known group isomorphism $(\mathcal{P}(n), \Delta) \cong ((\mathbb{Z}/2\mathbb{Z})^n, +)$ (componentwise addition modulo 2). Thus, $\mathcal{F}$ is a subgroup of $\mathcal{P}(n)$. The fact that $\#\mathcal{F} = 2^k$ is then an easy consequence of Lagrange's theorem (and the fact that 2 is prime). ■

In a slightly weaker version, the following was already observed in [32].

Theorem 2.10. Let $\mathcal{F} \subseteq \mathcal{P}(n)$ be a nontrivial, Δ-closed family. Then, for all $i \in [n]$ we have

$$\#\{F \in \mathcal{F} : i \in F\} = \frac{1}{2} \cdot \#\mathcal{F}$$

Proof. Fix $i \in [n]$. Because $\mathcal{F}$ is nontrivial, there is an $A \in \mathcal{F}$ with $i \in A$. Define the set $\mathcal{E} := \{F \in \mathcal{F} : i \notin F\}$ which is a proper subset of $\mathcal{F}$ (as $A \notin \mathcal{E}$). By properties of the symmetric difference, the set $\mathcal{E}$ (equipped with the symmetric difference Δ) is a subgroup of $\mathcal{F}$ (if $i \notin X$ and $i \notin Y$, then also $i \notin X\Delta Y$). Especially, one has $\emptyset \in \mathcal{E}$ so that $\mathcal{E}$ is nonempty. Defining the coset

$$A\mathcal{E} := \{A\Delta E : E \in \mathcal{E}\},$$

we claim

$$A\mathcal{E} = \{F \in \mathcal{F} : i \in F\}.$$

Indeed, if $F \in A\mathcal{E}$, then $F = A\Delta E$ for some $E \in \mathcal{E}$. Since $i \in A$ and $i \notin E$, we have $i \in F$. On the other hand, if $i \in F \in \mathcal{F}$, then $F = A\Delta E$ with $E := F\Delta A$. Since $i \in A$ and $i \in F$, we have $i \notin E$, so that $F \in A\mathcal{E}$. This shows the above set equality.
In particular, we get $\mathcal{E}\dot{\cup}A\mathcal{E} = \mathcal{F}$ (as every set in $\mathcal{F}$ either does or does not contain i). Thus, $\{\mathcal{E}, A\mathcal{E}\}$ is a complete set of cosets of $\mathcal{E}$. Since cosets have the same cardinalities, we get

$$\#\mathcal{F} = 2 \cdot \#(A\mathcal{E}) = 2 \cdot \#\{F \in \mathcal{F} : i \in F\},$$

proving the claim. ■

2.3 Simply Rooted Families

We already saw an equivalent characterization of union closed families via intersection closed families. This correspondance came from taking complements of all member sets of a union closed family. We now want to investigate what happens, if we take the complement of a set family $\mathcal{F}$ itself (i.e. $\mathcal{P}(n) \setminus \mathcal{F}$). This will be answered by the following definition and theorem.

Definition 2.11.

(i) Let $i \in A \subseteq B$, then define the **intervals**

$$[A, B] := \{X : A \subseteq X \subseteq B\}$$

and

$$[i, A] := \{X : i \in X \subseteq A\}.$$

(ii) Let $n \in \mathbb{N}$ and let $\mathcal{G} \subseteq \mathcal{P}(n)$ be a set family. We call $\mathcal{G}$ **simply rooted** if for all $\emptyset \neq A \in \mathcal{G}$ there is an $i \in A$ with $[i, A] \subseteq \mathcal{G}$.

The following theorem is well known in the theory of set families and has been used by several authors (e.g. [5, 26]) to investigate the structure of union closed sets.

Theorem 2.12. Let $n \in \mathbb{N}$. Then the following sets are in bijection to each other:

$$\{\mathcal{F} \subseteq \mathcal{P}(n) : \mathcal{F} \text{ union closed}\} \to \{\mathcal{G} \subseteq \mathcal{P}(n) : \mathcal{G} \text{ simply rooted}\}$$
$$\mathcal{F} \mapsto \mathcal{P}(n) \setminus \mathcal{F}.$$

Proof. We have to show that $\mathcal{F} \subseteq \mathcal{P}(n)$ is union closed, if and only if $\mathcal{G} := \mathcal{P}(n) \setminus \mathcal{F}$ is simply rooted. For this, let $\mathcal{F}$ be a union closed family and let $\emptyset \neq G \in \mathcal{G}$. Since $G \notin \mathcal{F}$, we must have

$$G \supsetneq \bigcup \{F \in \mathcal{F} : F \subseteq G\},$$

as the right hand side is the union of sets from $\mathcal{F}$ and thus itself a set in $\mathcal{F}$. Choosing an

$$i \in G \setminus \bigcup \{F \in \mathcal{F} : F \subseteq G\},$$

we obtain $[i, G] \subseteq \mathcal{G}$.
On the other hand, let $\mathcal{G}$ be simply rooted and let $A, B \in \mathcal{F}$. Assume, for contradiction, that $A \cup B \notin \mathcal{F}$. Thus, $A \cup B \in \mathcal{G}$ and there is an $i \in A \cup B$ with $[i, A \cup B] \subseteq \mathcal{G}$. We may assume, by symmetry, that $i \in A$. But then $A \in [i, A] \subseteq [i, A \cup B] \subseteq \mathcal{G}$, contradicting $A \in \mathcal{F} = \mathcal{P}(n) \setminus \mathcal{G}$. This shows that $A \cup B \in \mathcal{F}$, so that $\mathcal{F}$ is union closed. ∎

The usefulness of simply rooted families over $[n]$ lies in the fact that one can more easily control the sets that contain a given element $i \in [n]$. This is due to the fact that element containment is already part of the definition of the simply rootedness, while it is not entirely clear how element containment and union closedness are related (from $i \in A$ and $i \in B$ one can clearly deduce $i \in A \cup B$, but if for example $A = A \cup B$ we do not get a new set that contains i).

2.4 Interior Operators

This and the following section are based on concepts mentioned in [11, 15]. It should be noted that here we diverge a bit from the traditions used in the common literature. For what follows, one usually works with dual concepts that are more natural in the context of intersection closed families. Since we will be interested in union closed families, the concepts will be adapted accordingly. We start with the following definition.

Definition 2.13. Let $n \in \mathbb{N}$ and $\tau : \mathcal{P}(n) \to \mathcal{P}(n)$. We call τ a **pre-interior operator** if

(i) for all $A \subseteq [n]$ we have $\tau(A) \subseteq A$ (exclusivity);

(ii) for all $A \subseteq B \subseteq [n]$ we have $\tau(A) \subseteq \tau(B)$ (monotonicity).

We call τ an **interior operator** if additionally

(iii) for all $A \subseteq [n]$ we have $\tau(\tau(A)) = \tau(A)$ (idempotence).

We have already seen an interior operator in the proof of Theorem 2.12. Interior operators are also known, among other names, as *kernel operators*. Often, one works with *closure operators* instead, which are the dual concept of interior operators and which fulfill (ii) and (iii) from above as well as

(i') for all $A \subseteq [n]$ we have $\tau(A) \supseteq A$ (extensibility).

We start with the following lemma.

Lemma 2.14. Let $n \in \mathbb{N}$ and $\tau : \mathcal{P}(n) \to \mathcal{P}(n)$. Then the following are equivalent:

(i) For all $A \subseteq B \subseteq [n]$ we have $\tau(A) \subseteq \tau(B)$ (monotonicity);

(ii) For all $A, B \subseteq [n]$ we have $\tau(A) \cup \tau(B) \subseteq \tau(A \cup B)$ (superadditivity).

Proof. (i $\Rightarrow$ ii): For $A, B \subseteq [n]$ we have $A \subseteq A \cup B$ and $B \subseteq A \cup B$. Thus, as τ is monotone, we get $\tau(A) \subseteq \tau(A \cup B)$ and $\tau(B) \subseteq A \cup B$. Both together give $\tau(A) \cup \tau(B) \subseteq \tau(A \cup B)$. Since A and B were arbitrary, this proves the superadditivity.
(ii $\Rightarrow$ i): For $A \subseteq B \subseteq [n]$, we have $B = A \cup B$. Thus, by the superadditivity we get $\tau(A) \cup \tau(B) \subseteq \tau(A \cup B) = \tau(B)$, which shows $\tau(A) \subseteq \tau(B)$. Since $A \subseteq B$ were arbitrary, this shows the monotonicity of τ. ∎

The lemma shows that monotone set functions seem to encode union-structures. To make this more precise, for $\tau : \mathcal{P}(n) \to \mathcal{P}(n)$ define the *fixed point set*

$$\operatorname{Fix} \tau := \{A \subseteq [n] : \tau(A) = A\}.$$

We then have the following two characterizations.

Theorem 2.15. Let $\mathcal{F} \subseteq \mathcal{P}(n)$ be a set family. Then the following are equivalent:

(i) $\emptyset \in \mathcal{F}$ and $\mathcal{F}$ is union closed;

(ii) $\mathcal{F} = \operatorname{Fix} \tau$ for some pre-interior operator $\tau : \mathcal{P}(n) \to \mathcal{P}(n)$.

Proof. (i $\Rightarrow$ ii): Let $\emptyset \in \mathcal{F} \subseteq \mathcal{P}(n)$ be a union closed family and define the operator

$$\tau : \mathcal{P}(n) \to \mathcal{P}(n), X \mapsto \bigcup\{F \in \mathcal{F} : F \subseteq X\}.$$

It is straightforward to check that this is a pre-interior operator. For every $F \in \mathcal{F}$, we trivially have $\tau(F) = F$, so that $\mathcal{F} \subseteq \operatorname{Fix} \tau$. On the other hand, if $X \in \mathcal{P}(n) \setminus \mathcal{F}$ then, as in the proof of Theorem 2.12, we have $\tau(X) \subsetneq X$. This shows $\mathcal{P}(n) \setminus \mathcal{F} \subseteq \mathcal{P}(n) \setminus \operatorname{Fix} \tau$, i.e. $\operatorname{Fix} \tau \subseteq \mathcal{F}$. All in all, we get $\mathcal{F} = \operatorname{Fix} \tau$, which shows that (ii) holds.
(ii $\Rightarrow$ i): Let $\tau : \mathcal{P}(n) \to \mathcal{P}(n)$ be a pre-interior operator and set $\mathcal{F} := \operatorname{Fix} \tau$. By Definition 2.13 (i) it is clear that $\tau(\emptyset) = \emptyset$, so that $\emptyset \in \mathcal{F}$. It remains to show that $\mathcal{F}$ is union closed. For this, let $A, B \in \mathcal{F}$ and we need to show $\tau(A \cup B) = A \cup B$. The inclusion "$\subseteq$" holds by monotonicity of τ. For the other inclusion, Lemma 2.14 and $A, B \in \mathcal{F} = \operatorname{Fix} \tau$ yield

$$A \cup B = \tau(A) \cup \tau(B) \subseteq \tau(A \cup B).$$

Thus $\tau(A \cup B) = A \cup B$, as desired. ∎

The above correspondence between union closed families (containing the empty set) and pre-interior operators is not one-to-one. To fix this, we have the following characterization.

Theorem 2.16. Let $n \in \mathbb{N}$. The following sets are in bijection to each other:

$$\{\mathcal{F} \subseteq \mathcal{P}(n) : \emptyset \in \mathcal{F} \text{ union closed}\} \to \{\tau : \mathcal{P}(n) \to \mathcal{P}(n) \text{ interior operator}\}$$
$$\mathcal{F} \mapsto \left(X \mapsto \bigcup\{F \in \mathcal{F} : F \subseteq X\}\right).$$

The inverse of this correspondence is given by $\tau \mapsto \operatorname{Fix} \tau$.

Proof. In the proof of Theorem 2.15, we have seen that $\tau(X) := \bigcup\{F \in \mathcal{F} : F \subseteq X\}$ yields a pre-interior operator, which has fixed point set $\operatorname{Fix} \tau = \mathcal{F}$. It remains to show, that $\tau^2 = \tau$. But this follows immediately since for all $X \subseteq [n]$ we have $\tau(X) \in \mathcal{F}$ (indeed, the maximal $F \in \mathcal{F}$ with $F \subseteq X$). Thus, the above correspondance is well defined.
It remains to show, that the above correspondence is bijective. For this, we have to prove that if $\tau : \mathcal{P}(n) \to \mathcal{P}(n)$ is an interior operator then τ is given by

$$\tau(X) = \bigcup\{F \in \operatorname{Fix} \tau : F \subseteq X\} \tag{2.1}$$

for all $X \subseteq [n]$. Fix an $X \subseteq [n]$. We have (by monotonicity of τ)

$$F = \tau(F) \subseteq \tau(X)$$

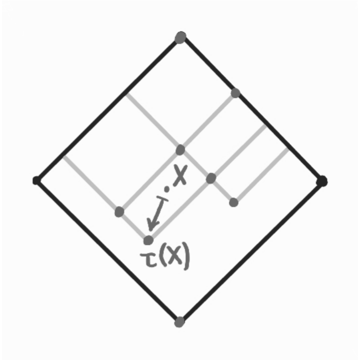

Figure 2.3: A set $X \subseteq [n]$ and its projection $\tau(X) \in \mathcal{F}$

for all $F \in \operatorname{Fix} \tau$ with $F \subseteq X$. Thus $\tau(X) \supseteq \bigcup\{F \in \operatorname{Fix}\tau : F \subseteq X\}$. For the other inclusion, note that $\tau(X) \subseteq X$ (τ exclusive) and $\tau(X) \in \mathcal{F}$ (τ idempotent), so that $\tau(X)$ is part of the above union. Thus, we also get the other inclusion and (2.1) holds. This finishes the proof. ■

For a union closed family $\emptyset \in \mathcal{F} \subseteq \mathcal{P}(n)$ the corresponding interior operator $\tau : \mathcal{P}(n) \to \mathcal{F} \subseteq \mathcal{P}(n)$ may be seen as a "projection" from $\mathcal{P}(n)$ onto $\mathcal{F}$.

2.5 Anticongruence Partitions

We will build upon the insights from the last section. Let $\emptyset \in \mathcal{F} \subseteq \mathcal{P}(n)$ be a union closed family and $\tau : \mathcal{P}(n) \to \mathcal{P}(n)$ the corresponding interior operator from Theorem 2.16. Notice that the image $\tau(\mathcal{P}(n))$ of τ is exactly $\mathcal{F}$ (as τ is idempotent and $\operatorname{Fix}\tau = \mathcal{F}$). For every $F \in \mathcal{F}$ we define its corresponding *cluster* by

$$\mathcal{T}(F) := \tau^{-1}(F) = \{X \subseteq [n] : \tau(X) = F\}. \tag{2.2}$$

It is then clear that $\{\mathcal{T}(F) : F \in \mathcal{F}\}$ forms a partition of $\mathcal{P}(n)$, as every $X \in \mathcal{P}(n)$ is in the cluster $\mathcal{T}(\tau(X))$. We want to investigate what structure these clusters have.

Definition 2.17. Let $n \in \mathbb{N}$ and $\mathbf{P} = \{P_1, ..., P_m\}$ be a partition of $\mathcal{P}(n)$ into nonempty, pairwise disjoint sets $P_i \subseteq \mathcal{P}(n), i = 1, ..., m$. We call $\mathbf{P}$ an **anticongruence partition** if the following holds: For all $A, B, C \subseteq [n]$ with $A, B \in P_i$ there is a $j = j(A, B, C) \in \{1, ..., m\}$ with $A \cap C, B \cap C \in P_j$. The corresponding equivalence relation $\gamma = \gamma_{\mathbf{P}}$ of an anticongruence partition will be called an **anticongruence relation**.

Thus, a partition $\mathbf{P}$ of $\mathcal{P}(n)$ is an anticongruence partition if for two sets A and B from the same partition class, $A \cap C$ and $B \cap C$ are also in the same partition class for arbitrary C (but it may vary in which partition class $A \cap C$ and $B \cap C$ land). In terms of anticongruence relations:

$$A \gamma B \quad \Rightarrow \quad (A \cap C) \gamma (B \cap C).$$

Remark 2.18. In the literature it is more common to work with equivalence relations instead of the corresponding partitions and also to use the dual concepts. There, a *congruence relation* is an equivalence relation θ on $\mathcal{P}(n)$ such that if $A\theta B$ then also $(A \cup C)\theta(B \cup C)$ for all $C \subseteq [n]$. These concepts originate in the theory of dependencies in data banks (see for instance [11, 15]).
Congruence relations are also used in order theory, where they are equivalence relations respecting unions and intersections (or more generally *joins* and *meets*, see Definition 3.9 below). For more on that side see [14]. However, we will use them more akin to the context in [11] from where we get the distinction between compatibility with intersections and unions.

Our first aim is to show that the partitions derived from a union closed family $\emptyset \in \mathcal{F} \subseteq \mathcal{P}(n)$ (or more precisely from the corresponding interior operator τ) are precisely the anticongruence partitions of $\mathcal{P}(n)$.

Theorem 2.19. Let $n \in \mathbb{N}$. The following sets are in a bijection to each other:

$$\begin{aligned} \{\tau : \mathcal{P}(n) \to \mathcal{P}(n) \text{ interior operator}\} &\to \{\mathbf{P} \text{ anticongruence partition of } \mathcal{P}(n)\} \\ \tau &\mapsto \{\mathcal{T}(F) : F \in \mathcal{F}\}, \end{aligned}$$

where $\mathcal{F} = \operatorname{Fix} \tau$ and $\mathcal{T}(F)$ as in (2.2). Denoting $\gamma = \gamma_{\mathbf{P}}$ the corresponding equivalence relation, the inverse is given via

$$\mathbf{P} \mapsto \left(X \mapsto \bigcap \{ Y \subseteq [n] : X \gamma Y \} \right). \tag{2.3}$$

Proof. We first show that the above map is well defined. For this, we have to show that for an interior operator $\tau : \mathcal{P}(n) \to \mathcal{P}(n)$ the corrsponding clusters form an anticongruence partition. Indeed, it suffices to show that for all $A, C \subseteq [n]$ and calling $F := \tau(A)$, we have

$$\tau(A \cap C) = \tau(F \cap C)$$

(then, if $A, B \subseteq [n]$ are in the same cluster so that $\tau(A) = \tau(B) = F$, we have $\tau(A \cap C) = \tau(F \cap C) = \tau(B \cap C)$, so that $A \cap C$ and $B \cap C$ are also in the same cluster). Since τ is exclusive, we have $A \supseteq F$ and consequently $A \cap C \supseteq F \cap C$. By monotonicity of τ, this gives the inclusion $\tau(A \cap C) \supseteq \tau(F \cap C)$. For the reverse inclusion, take an arbitrary $H \in \mathcal{F} = \operatorname{Fix} \tau$ with $H \subseteq A \cap C$. Then clearly $H \subseteq C$ and, again by monotonicity of τ, $H = \tau(H) \subseteq \tau(A) = F$. Hence $H \subseteq F \cap C$, which shows

$$\tau(A \cap C) = \bigcup \{ H \in \mathcal{F} : H \subseteq A \cap C \} \subseteq \bigcup \{ H \in \mathcal{F} : H \subseteq F \cap C \} = \tau(F \cap C)$$

and thus proves the other inclusion. This shows the well definedness.
To see that the claimed correspondance is also bijective, we have to verify that (2.3) is indeed an inverse. We are done once we showed that for all interior operators $\tau : \mathcal{P}(n) \to \mathcal{P}(n)$,

$$\tau(X) = \bigcap \{ Y \subseteq [n] : \tau(Y) = \tau(X) \}$$

(this is (2.3) applied to the anticongruence partition derived from the clusters of τ, which has the corresponding anticongruence relation $X \gamma Y$ if and only if $\tau(X) = \tau(Y)$) holds for all $X \subseteq [n]$. Since $Y = \tau(X)$ is part of the intersection (by idempotence of τ),

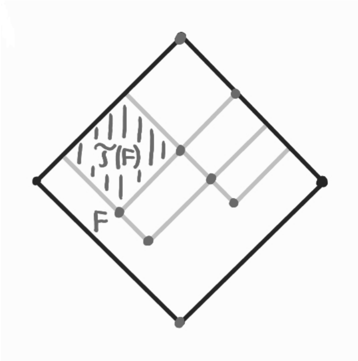

Figure 2.4: The cluster $\mathcal{T}(F)$ for an $F \in \mathcal{F}$

we directly get $\tau(X) \supseteq \bigcap\{Y \subseteq [n] : \tau(Y) = \tau(X)\}$. On the other hand, for all $Y \subseteq [n]$ with $\tau(Y) = \tau(X)$ we have $\tau(X) = \tau(Y) \subseteq Y$, which also shows the other inclusion. All in all, this verifies the claimed inverse map and finishes the proof. ■

From the get-go it is not clear how restrictive the anticongruence condition actually is. However, Theorems 2.16 and 2.19 give an easy way to construct all anticongruence partitions: Start with a union closed family containing the empty set, construct the corresponding interior operator and derive from that the partition via its clusters. In the following we collect some immediate consequences of anticongruence partitions, in particular the structure of its partition classes.

Corollary 2.20. Let $n \in \mathbb{N}$ and let $\mathbf{P} = \{P_1, ..., P_m\}$ be an anticongruence partition of $\mathcal{P}(n)$.

(i) If $A, B \in P_i$ and $A \subseteq B$, then $[A, B] \subseteq P_i$ (see Definition 2.11 (i));

(ii) If $A, B \in P_i$, then $A \cap B \in P_i$.

(iii) For all $i, j \in \{1, ..., m\}$ there is a unique $k = k(i, j) \in \{1, ..., m\}$ such that

$$P_i \Cap P_j \subseteq P_k,$$

where $\mathcal{A} \Cap \mathcal{B} := \{A \cap B : A \in \mathcal{A}, B \in \mathcal{B}\}$ denotes the *elementwise intersection*.

Of course, one can also define the *elementwise union* $\mathcal{A} \Cup \mathcal{B}$ of two set families $\mathcal{A}$ and $\mathcal{B}$ in an analogous manner.

Proof. (i): Let $A \subseteq B$ be sets from the same partition class P_i. For an arbitrary $C \in [A, B]$, Theorem 2.19 yields, that $A = A \cap C$ and $C = B \cap C$ are also in the same partition class. Since $A \in P_i$ thus also $C \in P_i$. This proves $[A, B] \subseteq P_i$.
(ii): If A and B are from the same partition class P_i, setting $C = B$ in Theorem 2.19 yields that $A \cap B$ and $B \cap B = B$ are also from the same partition class. Since $B \in P_i$ thus also $A \cap B \in P_i$.
(iii): Let $A, B \in P_i$ and $C, D \in P_j$. It suffices to show that $A \cap C$ and $B \cap D$ lie in

the same partition class (namely then P_k). Denoting by $\gamma = \gamma_{\mathbf{P}}$ the corresponding anticongruence relation, we have the implications

$$A\gamma B \Rightarrow (A \cap C)\gamma(B \cap C)$$
$$C\gamma D \Rightarrow (B \cap C)\gamma(B \cap D).$$

By transitivity of γ we thus get $(A \cap C)\gamma(B \cap D)$ as desired. ∎

Remark 2.21.

(i) By the previous corollary, we see that the partition classes of an anitcongruence partition are of the following form: Every partition class P_i contains a unique minimal member $X = \bigcap P_i$ and we have

$$P_i = \bigcup_{Y \in P_i} [X, Y].$$

In particular, using Theorems 2.16 and 2.19, for every anticongruence partition $\mathbf{P}$ the family $\{\bigcap P_i : P_i \in \mathbf{P}\}$ is union closed.

(ii) Also note that not all partitions of $\mathcal{P}(n)$ whose partition classes fulfill the two properties from (i) are an anticongruence partition. A simple example of that for $n = 2$ is $\{\{\emptyset\}, \{1\}, \{2, 12\}\}$ (using shorthand notation as in Example 2.5 (ii)). Indeed, the family of minimal members from each partition class has to form a union closed family, which is here not the case. Alternatively, 2 and 12 are in the same partition class, but intersecting with 1 yields $\emptyset$ and 1 respectively, which lie in different partition classes.

(iii) In general, $P_i \Cap P_j \subseteq P_k$ is a proper inclusion. As an example, over $n = 2$ we have the anticongruence partition $\{\{\emptyset, 1\}, \{2\}, \{12\}\}$. For $P_1 = \{\emptyset, 1\}$ and $P_2 = \{2\}$ (again we remind: "2" here actually stands for $\{2\}$, etc.) we have $P_1 \Cap P_2 = \{\emptyset\} \subsetneq P_1$ (i.e. $k = 1$).

The following corollary is particularly noteworthy and will be of importance later on.

Corollary 2.22. Let $n \in \mathbb{N}$, $\emptyset \in \mathcal{F} \subseteq \mathcal{P}(n)$ a union closed family, $\tau : \mathcal{P}(n) \to \mathcal{P}(n)$ the corresponding interior operator and $\mathcal{T}(F), F \in \mathcal{F}$ the corresponding clusters. Furthermore, let $E, F \in \mathcal{F}$ with $E \subseteq F$. Then the map

$$\iota_E^F : \mathcal{T}(F) \to \mathcal{T}(E), X \mapsto X \setminus (F \setminus E)$$

is an *order embedding* (i.e. for all $X, Y \in \mathcal{T}(F)$ it holds $X \subseteq Y$ if and only if $\iota_E^F(X) \subseteq \iota_E^F(Y)$), in particular injective.

Proof. We first prove that the given map is well defined. Clearly, if $X \in \mathcal{T}(F)$ then X and F are in the same cluster (namely $\mathcal{T}(F)$). Thus, by Theorem 2.19 we get that $X \setminus (F \setminus E)$ and $F \setminus (F \setminus E) = E$ are in the same cluster (note that taking "$... \setminus (F \setminus E)$" is the same as taking "$... \cap (([n] \setminus F) \cup E)$"). Hence $X \setminus (F \setminus E) \in \mathcal{T}(E)$, so that the

map is indeed well defined.
Let now $X, Y \in \mathcal{T}(F)$. Assume $X \subseteq Y$. This gives

$$\iota_E^F(X) = X \setminus (F \setminus E) \subseteq Y \setminus (F \setminus E) = \iota_E^F(Y),$$

so that ι_E^F is order preserving. Assume now $\iota_E^F(X) \subseteq \iota_E^F(Y)$. Using $F \subseteq X$ and $F \subseteq Y$ we get

$$X = \iota_E^F(X) \,\dot{\cup}\, (F \setminus E) \subseteq \iota_E^F(Y) \,\dot{\cup}\, (F \setminus E) = Y,$$

so that ι_E^F is an order embedding. It is a general fact from order theory that order embeddings are injective (using antisymmetry, see [14] for more details). In our case, this can be seen a bit more directly using the set operations. Indeed, if $X, Y \in \mathcal{T}(F)$ with $\iota_E^F(X) = \iota_E^F(Y)$, essentially the same argument as was already used above shows

$$X = \iota_E^F(X) \,\dot{\cup}\, (F \setminus E) = \iota_E^F(Y) \,\dot{\cup}\, (F \setminus E) = Y,$$

so that ι_E^F is injective. ■

Clearly, $\iota_E^E = \mathrm{id}_{\mathcal{T}(E)}$ is just the identity map. The above corollary shows that for $E, F \in \mathcal{F}, E \subseteq F$ the cluster $\mathcal{T}(F)$ can be embedded (order theoretically) into the cluster $\mathcal{T}(E)$ (via the map ι_E^F). In particular

$$\bigcup \left\{ \iota_E^F\Big(\mathcal{T}(F)\Big) : F \in \mathcal{F}, E \subseteq F \right\} = \mathcal{T}(E)$$

for all $E \in \mathcal{F}$ (the inclusion $\subseteq$ follows by the above corollary, the choice $F = E$ also shows $\supseteq$).

2.6 The Equivalence Theorem

For the sake of having a better overview, we recollect the major results so far into the following equivalence theorem.

Theorem 2.23. Let $n \in \mathbb{N}$. The following sets are in bijection to each other:

$$\begin{aligned} \mathbf{S}_{UC} &= \{\mathcal{F} \subseteq \mathcal{P}(n) : \emptyset \in \mathcal{F} \text{ union closed}\}, \\ \mathbf{S}_{IC} &= \{\mathcal{F} \subseteq \mathcal{P}(n) : [n] \in \mathcal{F} \text{ intersection closed}\}, \\ \mathbf{S}_{SR} &= \{\mathcal{G} \subseteq \mathcal{P}(n) : \emptyset \notin \mathcal{G} \text{ simply rooted}\}, \\ \mathbf{S}_{IO} &= \{\tau : \mathcal{P}(n) \to \mathcal{P}(n) \text{ interior operator}\}, \\ \mathbf{S}_{AP} &= \{\mathbf{P} \text{ anticongruence partition of } \mathcal{P}(n)\}. \end{aligned}$$

Proof. The bijections are given in the Theorems 2.7, 2.12, 2.16 and 2.19. It should be noted that the there given bijections are compatible with the here given sets (e.g. if $\emptyset \in \mathcal{F} \subseteq \mathcal{P}(n)$ is union closed, then the corresponding intersection closed family contains $[n]$, etc.) ■

Later on, we will mainly be interested in anticongruence partitions.

3 The Union Closed Sets Conjecture

The *Union Closed Sets Conjecture* is the main driving force in the study of set families (from a combinatorial point of view). We will now go into the details of the conjecture itself.

3.1 The Conjecture

We start by given a more precise statement for the Union Closed Sets Conjecture.

Conjecture 3.1 (Union Closed Sets Conjecture)**.** Let $n \in \mathbb{N}$ and let $\mathcal{F} \subseteq \mathcal{P}(n)$ be a nontrivial, union closed family. Then there is an element $i \in [n]$ with

$$\#\{F \in \mathcal{F} : i \in F\} \geq \frac{1}{2} \cdot \#\mathcal{F}.$$

The above may be reformulated as: Every finite, nontrivial union closed family contains an abundant element.

Remark 3.2. The minimal assumption on $\mathcal{F}$ that is certainly necessary (next to the union closedness) is that $\mathcal{F}$ contains at least one nonempty set. However, without losing generality, we can assume $\mathcal{F}$ to be nontrivial, as we may shrink the ground set to $\bigcup \mathcal{F}$ (i.e. disregard all elements that appear in no sets) and also assume $\bigcap \mathcal{F} = \emptyset$, so that no element trivially fulfills Conjecture 3.1. The latter can be done, as deleting trivial elements (in this sense) does not affect the union closedness of the family (see Observation 2.3 (ii)). Furthermore, we may assume $\mathcal{F}$ to be separating (see Definition 2.2 (ii)), as two elements which cannot be separated behave the same with regards to Conjecture 3.1. Thus we may delete indistinguishable elements (which again does not affect the union closedness).
It should also be noted that it is important to only consider finite union closed families (which then by the previous remarks may be assumed to only consist of finite sets). For example, the family

$$\mathcal{F} := \{\{n \in \mathbb{N} : n \geq k\} : k \in \mathbb{N}\}$$

over $\mathbb{N}$ is union closed. However, every $n \in \mathbb{N}$ is only contained in finitely many sets from $\mathcal{F}$, while $\mathcal{F}$ itself is of infinite cardinality. Thus, if one wants to generalize the Union Closed Sets Conjecture 3.1 to infinite families, one seems to need some additional "compactness" or "Noetherianity" assumptions (e.g. consider a similar example to above but over $\mathbb{N} \cup \{\infty\}$) to assure that at least one element is contained in an infinite number of sets from the family.

Using Theorems 2.7 and 2.12 and the bijections given there, an easy inspection leads to the following equivalence.

Observation 3.3. The following statements are equivalent:

(i) Every finite, nontrivial, union closed family contains an abundant element (Union Closed Sets Conjecture 3.1);

(ii) Every finite, nontrivial, intersection closed family contains a rare element (Intersection Closed Sets Conjecture);

(iii) Every finite, nontrivial, simply rooted family contains a rare element.

In Theorem 2.10, the matter for Δ-closed families was already resolved (even in a much stronger version). Unfortunately, this does not seem to help for the case of union closed families. Also, the bijections given in Theorems 2.16 and 2.19 do not seem to easily allow a simple (insightful) reformulation of the Union Closed Sets Conjecture 3.1. However, they help in understanding the structre of union closed families, which will be useful later on. For the rest of this chapter we recollect some known results that, irrespective of the truth of the Union Closed Sets Conjecture 3.1, are certainly interesting on their on rights. On the way, we also prepare some of the basic facts that will be used in the latter sections. We start with a recollection of known cases for Conjecture 3.1.

Remark 3.4.

(i) Every finite, union closed family containing a one-element or two-element set fulfills the Union Closed Sets Conjecture 3.1. Even more, if $\{x\} \in \mathcal{F}$ ($\{x, y\} \in \mathcal{F}$) then x (at least one of x or y) is abundant in $\mathcal{F}$. This is often refered to as a "folklore" fact in combinatorics (see [10] or [36, 37]). The corresponding statement for three-element sets is in general not true (see Example 2.5 (iii)). There is however a characterization of *local configurations* that guarantee the existence of an abundant element among them (see [33]). The latter uses techniques from convex analysis.

(ii) There are many classes of union closed families which are known to fulfill the Union Closed Sets Conjecture 3.1. Many of them are listed in [10] (e.g. if $\mathcal{F}$ is not only union closed but also intersection closed [33] or if for $\mathcal{F} \subseteq \mathcal{P}(n)$ one has $\#\mathcal{F} \geq (1-c)2^{n-1}$ with $c > 0$ an absolute constant [26]). Recently, in [1] it was shown that union closed families generated by translates of subsets of a finite, abelian group contain an abundant element. Even stronger, they showed that the *expected* number of sets containing any given element is $\geq \frac{1}{2} \cdot \#\mathcal{F}$ (see Observation 3.5).

(iii) It is also known (see [27, 42]) that any nontrivial, union closed family $\mathcal{F} \subseteq \mathcal{P}(n)$ contains an element that appears in at least $\Omega(\frac{\#\mathcal{F}}{\log_2(\#\mathcal{F})})$ many sets of $\mathcal{F}$. We will go into similar results later on.

We continue with some more specific ideas in the next sections.

3.2 Averaging

We start with the following observation.

Observation 3.5. Let $\mathcal{F} \subseteq \mathcal{P}(n)$ be a nontrivial, union closed family. If we know that, on average, we have

$$\frac{1}{n}\sum_{i=1}^{n} \#\{F \in \mathcal{F} : i \in F\} \geq \frac{1}{2} \cdot \#\mathcal{F}, \tag{3.1}$$

then $\mathcal{F}$ contains at least one abundant element. Of course, via double counting, the left hand side of the above may be rewritten as

$$\sum_{i=1}^{n} \#\{F \in \mathcal{F} : i \in F\} = \sum_{F \in \mathcal{F}} \#F.$$

Unfortunately (at least if we consider the expected value for the uniform distribution on $[n]$ as above), the inequality (3.1) is generally not true. A simple example of that is $\{\emptyset, \{1\}, [n]\} \subseteq \mathcal{P}(n)$ for $n \geq 3$. By setting the parameters k and n in Example 2.5 (i) right, one can even show that the expected value can be arbitrarily small. However, there is at least the following result from [5].

Theorem 3.6. Let $n \in \mathbb{N}$ and let $\mathcal{F} \subseteq \mathcal{P}(n)$ be a union closed family with $\#\mathcal{F} \geq \frac{2}{3} \cdot 2^n$. Then

$$\frac{1}{n}\sum_{i=1}^{n} \#\{F \in \mathcal{F} : i \in F\} \geq \frac{1}{2} \cdot \#\mathcal{F}$$

holds.

This theorem is sharp in the sense that for all $n \in \mathbb{N}$ there is a union closed family $\mathcal{F} \subseteq \mathcal{P}(n)$ of size $\#\mathcal{F} = \lfloor \frac{2}{3} \cdot 2^n \rfloor$ which does not fulfill inequality (3.1). Examples of that are the Hungarian families $\mathcal{H}(\lfloor \frac{2}{3} \cdot 2^n \rfloor)$ from Example 2.5 (ii).

There is at least the following result known about averaging, proved in [34].

Theorem 3.7. Let $n \in \mathbb{N}$ and let $\mathcal{F} \subseteq \mathcal{P}(n)$ be a union closed family consisting of $m = \#\mathcal{F}$ sets. Then

$$\frac{1}{m}\sum_{F \in \mathcal{F}} \#F \geq \frac{\log_2 m}{2}.$$

The bound in the theorem is equivalent to

$$\frac{1}{n}\sum_{x \in [n]} \#\{F \in \mathcal{F} : x \in F\} \geq \frac{\log_2 m}{n} \cdot \frac{m}{2}.$$

In the case of separating families, in [21] the following bound was proven.

Theorem 3.8. Let $n \in \mathbb{N}$ and let $\mathcal{F} \subseteq \mathcal{P}(n)$ be a separating, union closed family consisting of $m = \#\mathcal{F}$ sets. Then

$$\frac{1}{n} \sum_{x \in [n]} \#\{F \in \mathcal{F} : x \in F\} \geq \frac{n+1}{2}.$$

A short proof of this is also given in [10].

3.3 Posets and Lattices

It is natural to not only consider the algebraic and combinatorial structure derived from the union operation of a union closed family, but also the *order theoretic* structure that arises from the canonical order on $\mathcal{F}$ (set inclusion). In view of the Union Closed Sets Conjecture 3.1, which is concerned with element incidence, this seems especially useful.

3.3.1 Terminology from Order Theory

For a more profound discussion, we need some basic notions from order theory.

Definition 3.9. Let X be a nonempty set. Then $(X, \leq)$ is called a **poset** if $\leq$ defines a reflexive, antisymmetric and transitive relation on X. For $x, y \in X$ we say that y **covers** x if $x < y$ and $x \leq z \leq y$ implies $z = x$ or $z = y$ for all $z \in X$.
Furthermore, $(X, \leq)$ is called a **lattice** if every $x, y \in X$ have a unique minimal common upper bound (the **join** $x \vee y$) and a unique maximal common lower bound (the **meet** $x \wedge y$).

Intuitively, y covers x if y is a "successor" of x with respect to the order.

Example 3.10. Every set family $\mathcal{F}$ becomes a poset if we equip it with the order of set inclusion. If $\mathcal{F}$ is also union and intersection closed, then $(\mathcal{F}, \subseteq)$ becomes a lattice with join $A \vee B = A \cup B$ and meet $A \wedge B = A \cap B$. In general, even if $\mathcal{F}$ is neither union nor intersection closed, we may still be able to obtain a meet and join (which then differ from the union and intersection). For this, we have the following observation.

Observation 3.11. Let $n \in \mathbb{N}$ and $\emptyset \in \mathcal{F} \subseteq \mathcal{P}(n)$ be a nontrivial, union closed family containing the empty set. Since $\mathcal{F}$ is union closed, the corresponding poset $(\mathcal{F}, \subseteq)$ has a join, namely set unions. So far, this makes $\mathcal{F}$ into a *join-semilattice*. To see that $\mathcal{F}$ is indeed a lattice, we need to show that any two sets $A, B \in \mathcal{F}$ have a meet. However, this cannot be realized by the intersection (in general), as $\mathcal{F}$ may not be intersection closed. Using the interior operator $\tau : \mathcal{P}(n) \to \mathcal{P}(n), X \mapsto \bigcup\{F \in \mathcal{F} : F \subseteq X\}$ as in Theorem 2.16, we can still verify that

$$A \wedge B := \tau(A \cap B)$$

is a meet operation for $\mathcal{F}$. Indeed, clearly $A \wedge B$ is a set in $\mathcal{F}$ and it is also the largest set among all sets from $\mathcal{F}$, that is a subset of both A and B (by the way τ was defined). This makes $\mathcal{F}$ into a lattice.

Even for finite posets and lattices there exists a deep theory and a rich literature (see [14]). However, here we will only be concerned with the connections to the union closed sets problem. For this, we need some more notions. First note that for finite lattices $(X, \leq)$ (as we are interested in) there is a unique minimal element (the largest common lower bound of X, i.e. the meet of all elements of X, which is well defined for finite X) and a unique maximal element (the lowest common upper bound of X, i.e. the join of all elements of X). These minimal and maximal elements are commonly denoted by 0 and 1 (sometimes also refered to as the *bottom* $\perp$ and *top* $\top$ respectively) in X. Clearly, in the case of $\emptyset \in \mathcal{F} \subseteq \mathcal{P}(n)$ a nontrivial, union closed family the minimal element of $\mathcal{F}$ is $\emptyset$ and the maximal element of $\mathcal{F}$ is $[n]$.

Definition 3.12. Let $(X, \leq)$ be a lattice with join $\vee$ and meet $\wedge$. An element $x \in X$ which is not minimal is called **join-irreducible** if $x = y \vee z$ implies $y = x$ or $z = x$. Analogously, an element $x \in X$ which is not maximal is called **meet-irreducible** if $x = y \wedge z$ implies $x = y$ or $x = z$.

It is common to say that the minimal element of a lattice (if it has one, e.g. in finite lattices) is not join-irreducible and dually for the maximal element that it is no meet-irreducible. For example in the lattice $(\mathbb{N}, |)$ (the natural numbers with the order derived from dividability, the meet is the *greatest common devisor* and the join is the *least common multiple*), the convention is that 1 is not join-irreducible (i.e. prime), see [14] for more. Also note that an element x of a lattice L is join-irreducible if and only if there exists exactly one element of L that is covered by x. Dually, an element x is meet-irreducible if and only if there is a unique element covering x.

Example 3.13. For $n \in \mathbb{N}$ consider the nontrivial, union and intersection closed family $\mathcal{F} = \{\emptyset, [1], ..., [n]\}$ which is linearly (also known as totally) ordered. Then all members of $\mathcal{F}$ are easily seen to be join-irreducible (i.e. union-irreducible) and meet-irreducible (i.e. intersection-irreducible), except for the the minimal set $\emptyset$ and the maximal set $[n]$ of course.

3.3.2 The Lattice Formulation of the Union Closed Sets Conjecture

We now come back to the Union Closed Sets Conjecture 3.1. We start with the following considerations.

Remark 3.14. Let $\emptyset \in \mathcal{F} \subseteq \mathcal{P}(n)$ be a nontrivial, union closed family. For any given $i \in [n]$ we have

$$\{F \in \mathcal{F} : i \in F\} = \mathcal{F} \setminus \{G \in \mathcal{F} : i \notin G\} = \mathcal{F} \setminus \{G \in \mathcal{F} : G \subseteq [n] \setminus \{i\}\}.$$

Let $\tau : \mathcal{P}(n) \to \mathcal{P}(n)$ be the interior operator corresponding to $\mathcal{F}$ (Theorem 2.16). One then has

$$\{G \in \mathcal{F} : G \subseteq [n] \setminus \{i\}\} = \{G \in \mathcal{F} : G \subseteq \tau([n] \setminus \{i\})\}$$

by properties of τ. Thus, $i \in [n]$ is abundant in $\mathcal{F}$ if and only if

$$\#\{G \in \mathcal{F} : G \subseteq \tau([n] \setminus \{i\})\} \leq \frac{1}{2} \cdot \#\mathcal{F}.$$

The point we want to make is that $\tau([n] \setminus \{i\})$ is an explicit member of $\mathcal{F}$. Thus, the Union Closed Sets Conjecture 3.1 holds if and only if every nontrivial, union closed $\mathcal{F}$ has a maximal set $X \in \mathcal{F} \setminus \{[n]\}$ (even meet-irreducible in $\mathcal{F}$) such that

$$\#\{G \in \mathcal{F} : G \subseteq X\} \leq \frac{1}{2} \cdot \#\mathcal{F}. \tag{3.2}$$

If X is maximal in $\mathcal{F} \setminus \{[n]\}$, then for any $i \in [n] \setminus X$ we have $X = \tau([n] \setminus \{i\})$.
Note that in general, a union closed family need not contain a nonempty set $A \in \mathcal{F} \setminus \{\emptyset\}$ such that

$$\#\{F \in \mathcal{F} : A \subseteq F\} \geq \frac{1}{2} \cdot \#\mathcal{F}.$$

A simple example of that is $\{\emptyset, 12, 13, 23, 123\} \subseteq \mathcal{P}(3)$ (shorthand as in Example 2.5 (ii)). Thus, the requirement (3.2) gives a more intrinsic point of view of the structure of union closed families. This will be made more precise in the conjecture below.

The importance of lattices for the Union Closed Sets Conjecture 3.1 lies in the following (see [10]).

Conjecture 3.15. Let $(L, \leq)$ be a finite lattice with at least two elements. Then there is a meet-irreducible element $x \in L$ such that

$$\#\{y \in L : y \leq x\} \leq \frac{1}{2} \cdot \#L.$$

From context it should always be clear what meaning "$\leq$" and "$\geq$" have (i.e. in the lattice L or in $\mathbb{R}$). Also note that we have to require that $\#L \geq 2$, as otherwise L does not contain any meet-irreducible elements.

Theorem 3.16. The Union Closed Sets Conjecture 3.1 and Conjecture 3.15 are equivalent.

For a detailed proof we would need to discuss some more theory on lattices. Instead we will refer to the literature where we would need further results about lattices.

Proof. We use the reformulation of the Union Closed Sets Conjecture 3.1 for intersection closed families discussed in Observation 3.3. The gist of the proof of the above equivalence is then to interpret a lattice as an intersection closed family and vice versa.
<u>($\Rightarrow$)</u>: Assume that any nontrivial intersection closed family contains a rare element. Let $(L, \leq)$ be a lattice with at least two elements. For every $x \in L$ set

$$M(x) = \{y \in L : y \geq x, y \text{ meet-irreducible}\}.$$

One can show (see [14]) that $M(x \vee y) = M(x) \cap M(y)$ and $x = \bigwedge M(x)$. Thus $\mathcal{G} = \{M(x) : x \in L\}$ is an intersection close family. The ground set of $\mathcal{G}$ is the set of meet-irreducible elements in L, and $\mathcal{G}$ is in bijection with L. Applying the assumption on $\mathcal{G}$, we get the claim.
($\Leftarrow$): Assume Conjecture 3.15 holds. Let $\mathcal{F} \subseteq \mathcal{P}(n)$ be a nontrivial union closed family. Interpret $\mathcal{F}$ as a lattice $(\mathcal{F}, \subseteq)$ (see Observation 3.11). If $A \in \mathcal{F}$ is a meet-irreducible set in $\mathcal{F}$ with

$$\#\{F \in \mathcal{F} : F \subseteq A\} \leq \frac{1}{2} \cdot \#\mathcal{F},$$

then every $x \in [n] \setminus A$ is abundant (similar as discussed in Remark 3.14). ■

Remark 3.17.

(i) One usually states Conjecture 3.15 in its dual formulation, that every lattice L as above has a join-irreducible $x \in L$ with

$$\#\{y \in L : y \geq x\} \leq \frac{1}{2} \cdot \#L.$$

We have chosen the other formulation, as it is a bit closer to union closed families (see Remark 3.14), the dual formulation being more natural for intersection closed families.

(ii) Let L be a lattice and for $y \in L$ set

$$J(y) := \{x \in L : x \leq y, x \text{ join-irreducible}\}.$$

Contrary to what one might conjecture at first sight, it is in general not true that $J(y \vee z) = J(y) \cup J(z)$ for all $y, z \in L$. As for an example we can consider the *Fano plane* as a set family

$$\mathcal{F} := \{\emptyset, 1, 2, 3, 4, 5, 6, 7, 123, 145, 167, 246, 257, 347, 356, 1234567\}$$

(shorthand as in Example 2.5 (ii)). Then $(\mathcal{F}, \subseteq)$ becomes a lattice. The meet is here given via set intersection (i.e. $\mathcal{F}$ is an intersection closed family), however the join is not set union (see Remark 4.25). The join-irreducibles of $\mathcal{F}$ are then clearly the singletons $1, ..., 7$. We get

$$J(1 \vee 2) = J(123) = \{1, 2, 3\},$$

but

$$J(1) \cup J(2) = \{1, 2\}.$$

Clearly, one at least has the inclusion $J(y) \cup J(z) \subseteq J(y \vee z)$.

Remark 3.18. The above seems to suggest that every finite lattice L is (isomorphic to) a sublattice of $\mathcal{P}(n)$ for some $n \in \mathbb{N}$ (to be more specific, n the number of join-irreducibles of L). However, note that for $K \subseteq L$ to be a sublattice of L, it is required (see [14]) that the join and meet of K must coincide with the join an meet of L. Therefore, for some subset $K \subseteq L$ there might be a natural way to make K into a lattice using the order on K induced by the order of L, but this does not mean that

K is a sublattice of L. As for an example, simply note that union closed families $\emptyset \in \mathcal{F} \subseteq \mathcal{P}(n)$ are only sublattices, if they are also intersection closed. Otherwise, the meet in $\mathcal{F}$ (as already discussed in Observation 3.11) is not the intersection of sets. In general, $\mathcal{F} \subseteq \mathcal{P}(n)$ is only a join-subsemilattice.

Conjecture 3.15 is known to hold for many particular classes of lattices, see [10] for more details.

3.4 Lower Bounds

The Union Closed Sets Conjecture 3.1 postulates the existence of an element that appears in at least half of the sets of a nontrivial, union closed family. While it is still unknown if there is even an element in a c-fraction of sets for some $c > 0$, there are some known lower bounds that depend on n and m. We collect the most important ones for the sake of having a better view on what is known about union closed families. We start with the bound of Knill proved in [27] and later improved by Wójcik in [42]. The bound of Knill is relatively simple to prove.

Theorem 3.19. Let $n \in \mathbb{N}$ and $\mathcal{F} \subseteq \mathcal{P}(n)$ a nontrivial, union closed family with $m = \#\mathcal{F}$. There exists an element $x \in [n]$ with

$$\#\{F \in \mathcal{F} : x \in F\} \geq \frac{m-1}{\log_2(m+1)}.$$

Proof. Let $S \subseteq [n]$ be a minimum (i.e. of minimal cardinality) set that intersects all $F \in \mathcal{F} \setminus \{\emptyset\}$. We claim that for every $s \in S$ there is a set $F_s \in \mathcal{F}$ with $S \cap F_s = \{s\}$. Indeed, if not then there is an $s \in S$ such that for all $F \in \mathcal{F}$ with $s \in F$ we have $\#(S \cap F) \geq 2$. But then $S \setminus \{s\}$ still intersects every $F \in \mathcal{F} \setminus \{\emptyset\}$, contradicting minimality of S. Since $\mathcal{F}$ is union closed, for every $\emptyset \neq T \subseteq S$ there is an $F \in \mathcal{F}$ with $S \cap F = T$, namely $F = \bigcup_{s \in T} F_s$. Thus $m \geq 2^{\#S} - 1$, i.e. $\#S \leq \log_2(m+1)$. Every element of S appears in a nonempty set of $\mathcal{F}$, thus there must be an $x \in S \subseteq [n]$ with

$$\#\{F \in \mathcal{F} : x \in F\} \geq \frac{m-1}{\log_2(m+1)}.$$

∎

Wójcik strengthened this bound to the following.

Theorem 3.20. Let $n \in \mathbb{N}$ and $\mathcal{F} \subseteq \mathcal{P}(n)$ a nontrivial, union closed family with $m = \#\mathcal{F}$. There exists an element $x \in [n]$ with

$$\#\{F \in \mathcal{F} : x \in F\} \geq \frac{(1+o(1))m}{\log_2(4/3)\log_2 m} \geq \frac{2.4}{\log_2 m} \cdot m$$

for sufficiently large m.

On a different note, averaging also gives a lower bound for the frequency of the most frequent element in a union closed family. The following bound follows directly from Theorem 3.7.

Theorem 3.21. Let $n \in \mathbb{N}$ and $\mathcal{F} \subseteq \mathcal{P}(n)$ a nontrivial, union closed family with $m = \#\mathcal{F}$. There exists an element $x \in [n]$ with

$$\#\{F \in \mathcal{F} : x \in F\} \geq \frac{\log_2 m}{n} \cdot \frac{m}{2}.$$

By applying averaging not to the entire family but just to certain subfamilies, Balla obtained the following bound in [4].

Theorem 3.22. Let $n \in \mathbb{N}$ with $n \geq 16$ and $\mathcal{F} \subseteq \mathcal{P}(n)$ a nontrivial, union closed family with $m = \#\mathcal{F}$. There exists an element $x \in [n]$ with

$$\#\{F \in \mathcal{F} : x \in F\} \geq \sqrt{\frac{\log_2 n}{n}} \cdot \frac{m}{2}.$$

These are, as far as the author is aware, the best known bounds.

4 New Contributions

Using the previously developed theory, we discuss some new approaches for the Union Closed Sets Conjecture 3.1. First we study the structure of union closed families a bit more. We then look at the frequencies of all the elements from the ground set in a given union closed family. The first two sections may be seen as an attempt to prove the Union Closed Sets Conjecture 3.1 via an inductive approach.
In the third section we use some of the previously developed theory to prove a statement that is weaker than the Union Closed Sets Conjecture 3.1 but still gives some insight into the combinatorial structure of union closed families. We also demonstrate how to use this new technique to obtain some statements about the frequencies of elements in a union closed family. These considerations also naturally lead to related statements for other types of set families. In this context we state a conjecture that is again a weakening of the Union Closed Sets Conjecture 3.1 but may be more approachable.
At last, we consider families that are closed under operations similar to the union of two sets. We start with a relatively common technique in combinatorics to consider vector spaces over finite fields. This naturally generalizes to families defined over other set systems. An ambitious question is to determine what fraction of the sets from such families contain a certain element.

4.1 Appendable Sets

For a given union closed $\mathcal{F} \subseteq \mathcal{P}(n)$ we consider now those sets $A \in \mathcal{P}(n)$ that we can add to $\mathcal{F}$, so that $\mathcal{F} \cup \{A\}$ stays union closed. Such sets may be thought of as *appendable sets* to the family $\mathcal{F}$. We can prove the following theorem about the family of appendable sets.

Lemma 4.1. Let $n \in \mathbb{N}$ and let $\mathcal{F} \subseteq \mathcal{P}(n)$ be a union closed family. Then

$$\mathcal{F}^{\nabla} := \{A \subseteq [n] : \mathcal{F} \cup \{A\} \text{ union closed}\} \subseteq \mathcal{P}(n)$$

is also union closed.

Proof. We have to show, that if $\mathcal{F} \cup \{A\}$ and $\mathcal{F} \cup \{B\}$ are union closed then so is $\mathcal{F} \cup \{A \cup B\}$. If $A \cup B = A$ or $A \cup B = B$, this is clear. We can therefore assume $A \subsetneq A \cup B$ and $B \subsetneq A \cup B$. Fix an $F \in \mathcal{F}$. It suffices to show $F \cup (A \cup B) \in \mathcal{F} \cup \{A \cup B\}$. Observe that since $\mathcal{F} \cup \{A\}$ and $\mathcal{F} \cup \{B\}$ are union closed, we have $F \cup A \in \mathcal{F} \cup \{A\}$ and $F \cup B \in \mathcal{F} \cup \{B\}$. We distinguish the following four cases:
<u>$F \cup A \in \mathcal{F}$ and $F \cup B \in \mathcal{F}$:</u> Since $\mathcal{F}$ is union closed, we get

$$F \cup (A \cup B) = (F \cup A) \cup (F \cup B) \in \mathcal{F} \subseteq \mathcal{F} \cup \{A \cup B\}.$$

$\underline{F \cup A \in \mathcal{F} \text{ and } F \cup B = B}$: Since $\mathcal{F} \cup \{B\}$ is union closed and $B \subsetneq A \cup B$ (especially $F \cup (A \cup B) \neq B$), we get

$$F \cup (A \cup B) = (F \cup A) \cup B \in (\mathcal{F} \cup \{B\}) \setminus \{B\} \subseteq \mathcal{F} \subseteq \mathcal{F} \cup \{A \cup B\}.$$

$\underline{F \cup A = A \text{ and } F \cup B \in \mathcal{F}}$: This is completely analogous to the previous case.
$\underline{F \cup A = A \text{ and } F \cup B = B}$: We have $F \subseteq A$ and $F \subseteq B$, so that

$$F \cup (A \cup B) = A \cup B \in \mathcal{F} \cup \{A \cup B\}.$$

This finishes the case analysis and proves the claim. ■

Remark 4.2. By definition it is clear that $\mathcal{F} \subseteq \mathcal{F}^{\nabla}$ for union closed families $\mathcal{F} \subseteq \mathcal{P}(n)$. Furthermore, if $\mathcal{F} \subsetneq \mathcal{P}(n)$ then $\mathcal{F}^{\nabla} \setminus \mathcal{F}$ is nonempty, as it contains all maximal sets (with respect to set inclusion) in $\mathcal{P}(n) \setminus \mathcal{F} \neq \emptyset$. Thus, if $\mathcal{F}$ is not the entire power set then $\mathcal{F} \subsetneq \mathcal{F}^{\nabla}$ is a proper extension.

It would certainly be interesting (for union closed $\mathcal{F} \subsetneq \mathcal{P}(n)$) to investigate the structure of the family $\mathcal{F}^{\nabla} \setminus \mathcal{F}$ of appendable sets. In particular, the family of minimal such sets

$$\mathcal{F}^{\underline{\nabla}} := \mathrm{Min}(\mathcal{F}^{\nabla} \setminus \mathcal{F})$$

(where for a set family $\mathcal{X}$ the family $\mathrm{Min}\,\mathcal{X}$ consists of the inclusion wise minimal sets in $\mathcal{X}$) should be of special interest. In some sense good control over $\mathcal{F}^{\underline{\nabla}}$, or rather families of the form $\mathcal{F} \cup \{A\}$ for $A \in \mathcal{F}^{\underline{\nabla}}$, could be enough for an inductive proof (over the variable $m := \#\mathcal{F}$) of Conjecture 3.1 above.

Question. What additional structure does $\mathcal{F}^{\nabla}$ posses? What can be said about $\mathcal{F}^{\underline{\nabla}}$?

To use some terminology we have introduced about lattices: If $\mathbf{F}$ is the system of all union closed families over $[n]$ (also including $\emptyset$ and $\{\emptyset\}$), then for a given $\mathcal{F} \in \mathbf{F}$ the set $\{\mathcal{F} \cup \{A\} : A \in \mathcal{F}^{\nabla}\} \setminus \{\mathcal{F}\} \subseteq \mathbf{F}$ consists of those union closed families that *cover* $\mathcal{F}$ (see Definition 3.9). Here we (naturally) equip $\mathbf{F}$ with the order of set inclusion.

4.2 Frequencies in Union Closed Families

There have been many attempts to generalize the Union Closed Sets Conjecture 3.1. As was discussed in the introduction, this seems to be surprisingly hard. For example, we have already seen that there might not be an abundant element among a set of minimum cardinality in $\mathcal{F}$ (see Example 2.5 (iii) and [10, 37]). There is still some hope in strengthening the conjecture to obtain more insights into the structure of union closed families.
As for a new consideration, observe the following.

Observation 4.3. Let $\mathcal{F} = \mathcal{F}_1$ be a nontrivial, union closed family over $[n]$ and assume that the Union Closed Sets Conjecture 3.1 holds. We therefore have an $x_1 \in [n]$ such that

$$\#\{F \in \mathcal{F}_1 : x_1 \in F\} \geq \frac{1}{2} \cdot \#\mathcal{F}_1.$$

The family $\mathcal{F}_2 := \{F \setminus \{x_1\} : F \in \mathcal{F}_2, x_1 \in F\}$ (see also Observation 2.3 (ii)) is then a nontrivial, union closed family over $[n] \setminus \{x_1\}$. We thus again find an $x_2 \in [n] \setminus \{x_1\}$, such that

$$\#\{F \in \mathcal{F}_2 : x_2 \in F\} \geq \frac{1}{2} \cdot \#\mathcal{F}_2 \geq \frac{1}{4} \cdot \#\mathcal{F}_1.$$

We then conclude, that $\{x_1, x_2\}$ is a two element set, that is contained in at least $\frac{1}{4} \cdot \#\mathcal{F}$ of all sets from $\mathcal{F}$. Repeating this procedure k-times, we get a k-element set, that is contained in at least a 2^{-k}-th portion of all sets from $\mathcal{F}$.
In particular, this means that the k-th most frequent element among $\mathcal{F}$ appears in at least $2^{-k} \cdot \#\mathcal{F}$ many sets. We make this more precise and even give a strengthening in the following conjecture.

Conjecture 4.4. Let $\mathcal{F}$ be a nontrivial, union closed family over $[n]$ and assume

$$\#\{F \in \mathcal{F} : 1 \in F\} \geq \#\{F \in \mathcal{F} : 2 \in F\} \geq ... \geq \#\{F \in \mathcal{F} : n \in F\},$$

i.e. the elements are sorted by their frequencies (otherwise relabel them). Then, for all $k \in \{1, ..., n\}$ we have

$$\#\{F \in \mathcal{F} : k \in F\} \geq \frac{1}{2^{k-1} + 1} \cdot \#\mathcal{F}.$$

Remark 4.5. If Conjecture 4.4 indeed holds, then the bound given there is sharp. For $k = 2, ..., n$ this is easily demonstrated by the nontrivial, union closed family $\mathcal{F} := \mathcal{P}(k-1) \cup \{[n]\} \subseteq \mathcal{P}(n)$. The elements $i = 1, ..., k-1$ appear in $2^{k-2} + 1$ many sets, the elements $i = k, ..., n$ appear in one set (namely only $[n]$). Since $\#\mathcal{F} = 2^{k-1} + 1$ we thus get the equality

$$\#\{F \in \mathcal{F} : k \in F\} = 1 = \frac{1}{2^{k-1} + 1} \cdot \#\mathcal{F}.$$

Also, setting $k = 1$ into Conjecture 4.4 yields Conjecture 3.1, which shows that the above conjecture generalizes the Union Closed Sets Conjecture 3.1.

To justify Conjectue 4.4 somewhat, we start with the following lemma.

Lemma 4.6. Let $n \in \mathbb{N}$, $\mathcal{F} \subseteq \mathcal{P}(n)$ a nontrivial union closed family and $x \in [n]$. For every $A \in \mathcal{F}$ with $x \in A$ we have

$$\#\{F \in \mathcal{F} : x \in F\} \geq \frac{1}{2^{\#A-1} + 1} \cdot \#\mathcal{F}.$$

Proof. Let x and A be as in the assumptions and consider the map

$$\varphi\colon \{G \in \mathcal{F} : x \notin G\} \to \{F \in \mathcal{F} : x \in F\}, G \mapsto G \cup A.$$

This map is well defined ($\mathcal{F}$ nontrivial and union closed) but in general not injective. For $G_1, G_2 \in \mathcal{F}$ with $x \notin G_1, G_2$ we have

$$\begin{aligned} & G_1 \cup A = G_2 \cup A \\ \Leftrightarrow\, & G_1 \setminus A = G_2 \setminus A \\ \Leftrightarrow\, & G_1 \setminus (A \setminus \{x\}) = G_2 \setminus (A \setminus \{x\}), \end{aligned}$$

where the first equivalence follows by basic set theory and the second by the fact that $x \notin G_1, G_2$. Thus, G_1 and G_2 map (under the map φ) to the same set if and only if they coincide on $[n] \setminus (A \setminus \{x\})$. In conclusion, for any given set $G_1 \subseteq [n]$ there are (including G_1 itself) at most $2^{\#A-1}$ such sets $G_2 \subseteq [n]$.
It follows that there is a family $\mathcal{G} \subseteq \{G \in \mathcal{F} : x \notin G\}$, such that

$$\#\mathcal{G} \geq \frac{\#\{G \in \mathcal{F} : x \notin G\}}{2^{\#A-1}}$$

and $\varphi|_{\mathcal{G}} : \mathcal{G} \to \{F \in \mathcal{F} : x \in F\}$ is injective. This can be done by defining the equivalence relation $G_1 \sim G_2$ if and only if $\varphi(G_1) = \varphi(G_2)$ on $\{G \in \mathcal{F} : x \notin G\}$ and picking, from every equivalence class, a representative. The family of representatives is then $\mathcal{G}$. The injectivity of $\varphi|_{\mathcal{G}}$ follows by definition and the bound on the cardinality of $\mathcal{G}$ by the above combinatorial argument.
Since $\varphi|_{\mathcal{G}}$ is injective we get

$$\#\{F \in \mathcal{F} : x \in F\} \geq \#\mathcal{G} \geq \frac{\#\{G \in \mathcal{F} : x \notin G\}}{2^{\#A-1}} = \frac{\#\mathcal{F} - \#\{F \in \mathcal{F} : x \in F\}}{2^{\#A-1}}.$$

From this the claim follows. ■

Corollary 4.7. Let $n \in \mathbb{N}$ and $\mathcal{F} \subseteq \mathcal{P}(n)$ a nontrivial, union closed family. Assume that the elements of the ground set $[n]$ are ordered by frequencies (as in Conjecture 4.4). Then $k = n$ and $k = n - 1$ fulfill the bound given in Conjecture 4.4.

Proof. For the case $k = n$ simply apply Lemma 4.6 to $x = n$ and $A = [n] \in \mathcal{F}$. For the case $k = n - 1$ it suffices to consider the case of separating families. Here, since $n - 1$ is contained in a set $F \in \mathcal{F}$ of size at most $n - 1$, applying Lemma 4.6 to this set we get the statement for $k = n - 1$. ■

Conjecture 4.4 may be useful for an inductive proof, as we also have stronger assumptions. Another reason why this approach might be useful is given in the following informal discussion.
Assume again that the elements of $[n]$ are ordered by their frequencies in $\mathcal{F}$. Thus, the element k for $k > 1$ is by definition less frequent than the elements $1, ..., k-1$. Therefore, if we know that k is already relatively frequent among $\mathcal{F}$, then 1 in particular must be frequent as well. This is good news in light of of Conjecture 4.4 and the Union Closed Sets Conjecture 3.1 in general. If on the other hand k is very infrequent, so that maybe even equality occurs in

$$\#\{F \in \mathcal{F} : k \in F\} = \frac{1}{2^{k-1} + 1} \cdot \#\mathcal{F},$$

then this seems to force a certain structure on $\mathcal{F}$ (at least for sufficiently large k) which makes the elements $1, ..., k - 1$ rather frequent (see Remark 4.5). This is again good news in with regard of the Union Closed Sets Conjecture 3.1.

Question. Can the above discussion be made precise?

In a similar spirit, we at least have the following statement.

Remark 4.8. It should be clear that for any given x, we apply the bound guaranteed by the proposition with $A \in \operatorname{Argmin}\{\#A : x \in A \in \mathcal{F}\}$. However, in general this bound does not seem to be so good. For example, if $A = \{x, y\}$, then we have already mentioned (in Remark 3.4 (i)) that x or y must be abundant. The proposition however then only gives a factor $\frac{1}{3}$ instead of the desired $\frac{1}{2}$. In fact, the above proof is a generalization of the (folklore) proof that x is abundant given that $\{x\} \in \mathcal{F}$.
Inspired by the above proof, we do however obtain another interpretation of the bound in Conjecture 4.4: the fact that

$$\#\{F \in \mathcal{F} : k \in F\} \geq \frac{1}{2^{k-1}+1} \cdot \#\mathcal{F}$$

is equivalent to

$$\#\{F \in \mathcal{F} : k \in F\} \geq \frac{1}{2^{k-1}} \cdot \#\{G \in \mathcal{F} : k \notin G\}.$$

We end this chapter with the following question (in a similar spirit to the Erdős-Gallai theorem [19]) that generalizes Conjecture 4.4 even further.

Question. Which sequences $(\#\{F \in \mathcal{F} : x \in F\})_{x \in [n]}$ can occur?

4.3 Weaker Versions

4.3.1 Up-sets and Union Closed Families

Using the notation introduced in Definition 2.11 (i) we may rewrite

$$\{F \in \mathcal{F} : i \in F\} = \mathcal{F} \cap [i, [n]]$$

for $\mathcal{F} \subseteq \mathcal{P}(n)$ and $i \in [n]$. The family $\mathcal{U} = [i, [n]]$ has the following properties:

(i) $\mathcal{U} \subseteq \mathcal{P}(n)$ is an up-set;

(ii) $\#\mathcal{U} = 2^{n-1}$.

Motivated by this, we pose the following weaker (relaxed) version of the Union Closed Sets Conjecture 3.1 in hopes of obtaining new insights.

Conjecture 4.9. Let $n \in \mathbb{N}$ and $\mathcal{F} \subseteq \mathcal{P}(n)$ a nontrivial, union closed family. Then there is an up-set $\mathcal{U} \subseteq \mathcal{P}(n)$ of size $\#\mathcal{U} \leq 2^{n-1}$ with

$$\#(\mathcal{F} \cap \mathcal{U}) \geq \frac{1}{2} \cdot \#\mathcal{F}.$$

We will give a proof of this conjecture further below. The general idea how to get a suitable up-set is the following *greedy* procedure.

Start with $\mathcal{U} = \emptyset$. As long as $\#(\mathcal{F} \cap \mathcal{U}) < \frac{1}{2} \cdot \#\mathcal{F}$, choose that $F \in \mathcal{F} \setminus \mathcal{U}$ that minimizes $\#([F,[n]] \setminus \mathcal{U})$. Enlarge $\mathcal{U} \leftarrow \mathcal{U} \cup [F,[n]]$ and repeat the previous step.

That is, at every step we try to add a set from $\mathcal{F}$ not already in $\mathcal{U}$, that minimizes the number of sets added to $\mathcal{U}$ ($\mathcal{U}$ is at every step an up-set). It remains to justify why $\mathcal{U}$ at the end is not too large (i.e. $\#\mathcal{U} \leq 2^{n-1}$).
With this at our aim, we start with the following. Recall that any union closed family containing the empty set has a corresponding interior operator (Theorem 2.16) and anticongruence partition (Theorem 2.19).

Lemma 4.10. Let $n \in \mathbb{N}$, $\emptyset \in \mathcal{F} \subseteq \mathcal{P}(n)$ union closed, $\tau : \mathcal{P}(n) \to \mathcal{P}(n)$ the corresponding interior operator and $\mathcal{T}(F), F \in \mathcal{F}$ the corresponding clusters. Then the map

$$\mathcal{F} \to \mathbb{N}, F \mapsto \#\mathcal{T}(F)$$

is order reversing, i.e. for $E, F \in \mathcal{F}$ with $E \subseteq F$ we have $\#\mathcal{T}(E) \geq \#\mathcal{T}(F)$.

Proof. This follows directly from the injectivity of the maps ι_E^F from Corollary 2.22. ■

This lemma together with the greedy procedure above gives the following immediate corollary.

Corollary 4.11. Let the situation be as in Lemma 4.10 and let $m := \#\mathcal{F}$. Then the sets in $\mathcal{F}$ may be arranged in such a way $F_1, F_2, ..., F_m$ that

(i) $\#\mathcal{T}(F_1) \leq \#\mathcal{T}(F_2) \leq ... \leq \#\mathcal{T}(F_m)$;

(ii) if $F_i \supseteq F_j$ then $i \leq j$.

Property (ii) can be seen as a (reversed) *linear extension* of the poset $(\mathcal{F}, \subseteq)$ (see [40]). Of course (if $\mathcal{F}$ is nontrivial) we have $F_1 = [n]$ and $F_m = \emptyset$. In particular we get $\#\mathcal{T}(F_1) = 1$ and

$$\#\mathcal{T}(\emptyset) = \max\{\#\mathcal{T}(F) : F \in \mathcal{F}\}.$$

Before coming back to Conjecture 4.9 we need an elementary lemma.

Lemma 4.12. Let $0 \leq n_1 \leq n_2 \leq ... \leq n_m$ be real numbers, set $N := \sum_{i=1}^m n_i$ and let $\vartheta \in [0,1]$. Then

$$\vartheta N \geq \sum_{i=1}^{\lfloor \vartheta m \rfloor} n_i.$$

Here $\lfloor \cdot \rfloor$ is rounding down and $\lceil \cdot \rceil$ is rounding up. Note that this bound also holds for $m = 0$.

Proof. The statement is clear for $\vartheta < \frac{1}{m}$ (right hand side $= 0$, left hand side ≥ 0), so suppose $\vartheta \geq \frac{1}{m}$. Define the function $f : (0, m] \to \mathbb{R}, f(x) := n_{\lceil x \rceil}$. By monotonicity of the n_i's the function f is monotonically increasing. Thus, since $\frac{m}{\lfloor \vartheta m \rfloor} \geq 1$, we have

$$f(x) \leq f\left(\frac{m}{\lfloor \vartheta m \rfloor} \cdot x\right)$$

for $x \in (0, \lfloor \vartheta m \rfloor]$. Then, using the substitution $y = \frac{m}{\lfloor \vartheta m \rfloor} \cdot x$, we get

$$\begin{aligned}\sum_{i=1}^{\lfloor \vartheta m \rfloor} n_i = \int_0^{\lfloor \vartheta m \rfloor} f(x)dx \leq \int_0^{\lfloor \vartheta m \rfloor} f\left(\frac{m}{\lfloor \vartheta m \rfloor} \cdot x\right) dx = \int_0^m f(y) \cdot \frac{\lfloor \vartheta m \rfloor}{m} dy \\ = \frac{\lfloor \vartheta m \rfloor}{m} \cdot \sum_{i=1}^{m} n_i \leq \vartheta N.\end{aligned}$$

■

At first, the last inequality in the above proof seems a bit wasteful. However, for a given $\vartheta \in [0, 1]$ by simply applying the above lemma for $\vartheta' := \frac{\lfloor \vartheta m \rfloor}{m}$ we preserve the sum on the right hand side, i.e.

$$\sum_{i=1}^{\lfloor \vartheta m \rfloor} n_i = \sum_{i=1}^{\lfloor \vartheta' m \rfloor} n_i,$$

while making the left hand side even smaller, i.e.

$$\vartheta N \geq \vartheta' N.$$

We are now ready to prove Conjecture 4.9 in a slightly more general form.

Theorem 4.13. Let $n, t \in \mathbb{N}$ and $\mathcal{F} \subseteq \mathcal{P}(n)$ a nontrivial, union closed family. Then there is an up-set $\mathcal{U} \subseteq \mathcal{P}(n)$ of size $\#\mathcal{U} \leq \left\lceil \frac{1}{t} \cdot 2^n \right\rceil$ such that

$$\#(\mathcal{F} \cap \mathcal{U}) \geq \frac{1}{t} \cdot \#\mathcal{F}.$$

Proof. We may assume $\emptyset \in \mathcal{F}$. Let $m = \#\mathcal{F}$ and order $\mathcal{F} = \{F_1, F_2, ..., F_m\}$ as in Corollary 4.11. We claim that $\mathcal{U} := \bigcup_{i=1}^{\lceil m/t \rceil} \mathcal{T}(F_i)$ does the job. Indeed, because $F_i \in \mathcal{T}(F_i)$ for all i, we have $\mathcal{F} \cap \mathcal{U} = \{F_1, ..., F_{\lceil m/t \rceil}\}$ and as such

$$\#(\mathcal{F} \cap \mathcal{U}) = \left\lceil \frac{m}{t} \right\rceil \geq \frac{1}{t} \cdot \#\mathcal{F}.$$

It remains to show that $\mathcal{U}$ is an up-set of size at most $\left\lceil \frac{1}{t} \cdot 2^n \right\rceil$. The fact that $\mathcal{U}$ is an up-set stems from the fact that $\{\mathcal{T}(F_1), ..., \mathcal{T}(F_m)\}$ is a partition of $\mathcal{P}(n)$ having the property from Corollary 2.20 (i) and that the F_i's are arranged in a way fulfilling Corollary 4.11 (ii). To bound $\#\mathcal{U}$ we use Corollary 4.11 (i). For this, note that $F_1 = [n]$ (see the remark after Corollary 4.11). Setting $n_i := \#\mathcal{T}(F_i)$ we thus have $1 = n_1 \leq n_2 \leq ... \leq n_m$ and

$$\sum_{i=2}^{m} n_i = 2^n - 1.$$

Applying Lemma 4.12 to $n_2, ..., n_m$ with $\vartheta = \frac{\lceil \frac{m}{t} \rceil - 1}{m-1}$ we get

$$\begin{aligned}
\#\mathcal{U} &= \sum_{i=1}^{\lceil m/t \rceil} n_i = 1 + \sum_{i=1}^{\lceil m/t \rceil - 1} n_{i+1} \leq 1 + \frac{\lceil \frac{m}{t} \rceil - 1}{m-1} \sum_{i=1}^{m-1} n_{i+1} \\
&\leq 1 + \frac{\frac{m+t-1}{t} - 1}{m-1}(2^n - 1) = 1 + \frac{1}{t} \cdot (2^n - 1) \\
&< \frac{1}{t} \cdot 2^n + 1,
\end{aligned}$$

but $\#\mathcal{U}$ is an integer so even $\#\mathcal{U} \leq \left\lceil \frac{1}{t} \cdot 2^n \right\rceil$, finishing the proof. ■

Setting $t = 2$ yields Conjecture 4.9.

Corollary 4.14. Conjecture 4.9 is true.

We can even give a slight strengthening of Theorem 4.13 if we know a bit more about the family $\mathcal{F}$. By Observation 2.3 (iii) the following class of families also includes all separating families.

Theorem 4.15. Let $n \in \mathbb{N}$ and $\mathcal{F} \subseteq \mathcal{P}(n)$ a nontrivial, union closed family of size $m = \#\mathcal{F}$ containing a set of size $n-1$ and let $t \in \{1, ..., m-1\}$. Then there is an up-set $\mathcal{U} \subseteq \mathcal{P}(n)$ of size

$$\#\mathcal{U} \leq \frac{1}{t} \cdot \frac{m-t-1}{m-2} \cdot 2^n + 2\left(1 - \frac{1}{t}\right)\left(1 + \frac{1}{m-2}\right),$$

such that

$$\#(\mathcal{F} \cap \mathcal{U}) \geq \frac{1}{t} \cdot \#\mathcal{F}.$$

Proof. The proof is analogous to the proof of Theorem 4.13 and we will only outline the differences. Let $n_1, ..., n_m$ be as before. Using the existence of a set of size $n-1$ in $\mathcal{F}$ we can further assume $n_1 = n_2 = 1$. This time using $\vartheta = \frac{\lceil m/t \rceil - 2}{m-2}$ we get the desired bound

$$\begin{aligned}
\#\mathcal{U} &= 2 + \sum_{i=1}^{\lceil m/t \rceil - 2} n_{i+2} \leq 2 + \frac{\lceil m/t \rceil - 2}{m-2} \sum_{i=1}^{m-2} n_{i+2} \\
&\leq 2 + \frac{\frac{m+t-1}{t} - 2}{m-2} \cdot (2^n - 2) \\
&= \frac{1}{t} \cdot \frac{m-t-1}{m-2} \cdot 2^n + 2\left(1 - \frac{1}{t}\right)\left(1 + \frac{1}{m-2}\right).
\end{aligned}$$

■

For fixed m and t, the case for separating families of size m (or general families as described by the theorem) thus guarantees an up-set of size at most

$$\frac{1}{t} \cdot \frac{m-t-1}{m-2} \cdot 2^n + O(1),$$

which, for sufficiently large n and $t \geq 2$, is better than

$$\frac{1}{t} \cdot 2^n + O(1)$$

as guaranteed by Theorem 4.13. This might be surprising, as the assumption of having a set of size $n-1$ in our family might seem very minor at first. Again, for $t=2$ the above bound reads

$$\#\mathcal{U} \leq \frac{m-3}{m-2} \cdot 2^{n-1} + 1 + \frac{1}{m-2},$$

giving

$$\#\mathcal{U} \leq \left\lfloor \frac{m-3}{m-2} \cdot 2^{n-1} + \frac{1}{m-2} \right\rfloor + 1.$$

This bound is mainly useful for smaller m (compared to 2^{n-1}).

4.3.2 Bounds on Frequencies for Union Closed Families

While it seems unlikely that Theorem 4.13 alone can lead to a solution of the Union Closed Sets Conjecture 3.1, there might be some hope to gain further insights into union closed families. As mentioned before, it is not even known if every nontrivial union closed family $\mathcal{F}$ contains an element contained in some c-fraction of sets from $\mathcal{F}$ for some constant $c > 0$ ($c = 1/2$ being the Union Closed Sets Conjecture 3.1). Theorem 4.13 however gives some extra structure to the "big" sets of a union closed family $\mathcal{F} \subseteq \mathcal{P}(n)$, namely that half of the sets from $\mathcal{F}$ lie in an up-set of size at most 2^{n-1}. The following theorem aims to show how this information might be useful.

Theorem 4.16. There is a universal constant $C > 0$ with the following property: Let $n, t \in \mathbb{N}$ with $n \geq 2$ and $t \geq \frac{Cn}{\log_2 n}$ and let $\mathcal{U} \subseteq \mathcal{P}(n)$ be an up-set with $\#\mathcal{U} \leq 2^{n-1}$. For any sets $A_1, ..., A_t \in \mathcal{U}$ there are indices $1 \leq i < j \leq t$ with $A_i \cap A_j \neq \emptyset$.

Thus, choosing any t sets from $\mathcal{U}$, as long as t is sufficiently large, two of the chosen sets must intersect. This then also holds for the set family $\mathcal{F} \cap \mathcal{U} \subseteq \mathcal{U}$ from Theorem 4.13.

Proof. We prove the theorem by contraposition. Assume that $\mathcal{U} \subseteq \mathcal{P}(n)$ is an up-set with $\#\mathcal{U} \leq 2^{n-1}$ and that $A_1, ..., A_t \in \mathcal{U}$ are all pairwise disjoint. Denoting $a_i := \#A_i$ for $i = 1, ..., t$, the up-set $\bigcup_{i=1}^{t} [A_i, [n]]$ (see Definition 2.11 (i)) generated by the A_i's is of size (by the principle of inclusion-exclusion)

$$\sum_{\emptyset \neq I \subseteq [t]} (-1)^{\#I-1} 2^{n - \sum_{i \in I} a_i} = 2^n \left(1 - \prod_{i=1}^{t} \left(1 - 2^{-a_i}\right)\right),$$

since $2^{n-\sum_{i\in I} a_i}$ is the size of $\bigcap_{i\in I}[A_i,[n]] = \left[\bigcup_{i\in I} A_i,[n]\right]$. For variables $x_1,...,x_t \geq 0$ with $\sum_{i=1}^t x_i \leq n$, the expression

$$1-\prod_{i=1}^{t}\left(1-2^{-x_i}\right)$$

is minimized when $x_1 = ... = x_t = \frac{n}{t}$. We get

$$2^{n-1} \geq \#\mathcal{U} \geq \#\bigcup_{i=1}^{t}[A_i,[n]] \geq 2^n\left(1-\left(1-2^{-n/t}\right)^t\right) = 2^n - \left(2^{n/t}-1\right)^t$$

so that $\left(2^{n/t}-1\right)^t \geq 2^{n-1}$. Further rearranging and bounding yields

$$n \geq t\cdot\log_2\left(\frac{1}{1-2^{-1/t}}\right) > t\cdot\log_2\left(\frac{t}{\ln 2}\right). \tag{4.1}$$

If now we would have $t \geq \frac{2n}{\log_2 n}$, then

$$t\cdot\log_2\left(\frac{t}{\ln 2}\right) \geq \frac{2n}{\log_2 n}\cdot\log_2\left(\frac{2n}{\ln 2\cdot\log_2 n}\right) = n\cdot 2\log_n\left(\frac{2n}{\ln 2\cdot\log_2 n}\right) \geq n,$$

contradicting (4.1). We thus have to have

$$t < \frac{2n}{\log_2 n},$$

completing the proof (with $C = 2$). ■

Remark 4.17. Using the above proof, we can easily improve the constant to $C = 1.367... < 1.368$. One could also improve C by only considering up-sets of size at most $c\cdot 2^n$ for some $c > 0$ (above $c = 1/2$).
The important part however is that choosing $\Omega(\frac{n}{\log n})$ many sets in $\mathcal{U}$ already gives two intersecting sets. In particular, combining with Theorem 4.13, every nontrivial, union closed family $\mathcal{F} \in \mathcal{P}(n)$ contains a subfamily $\mathcal{E} \subseteq \mathcal{F}$ of size at least $\frac{1}{2}\cdot\#\mathcal{F}$, such that $O(\frac{n}{\log n})$ sets from $\mathcal{E}$ cannot be pairwise disjoint.

We now demonstrate how to to obtain a bound for the frequency of the most frequent element in a nontrivial, union closed family using only the two results above. This should be more understood as a demonstration on how one may use the above results, but further optimizations (obtaining better asymptotics) by a more detailed analysis are conceivable.
We first have to go through some graph theory, see [16]. For a graph $G = (V,E)$ the *clique number* is defined as

$$\omega(G) := \max\left\{\#X : X \subseteq V, \binom{X}{2} \subseteq E\right\},$$

the maximum cardinality of a set of vertices X such that every pair $\{x,y\} \subseteq X$ is an edge in G. We also define the *independence number*

$$\alpha(G) := \max\left\{\#I : I \subseteq V, \binom{I}{2} \cap E = \emptyset\right\},$$

the maximum cardinality of a set of vertices I such that no pair $\{x, y\} \subseteq I$ is an edge in G. The inequality $\alpha(G) < t$ is then equivalent to the statement that any t vertices contain at least one edge among them in G. For a graph $G = (V, E)$ one can define the *complement graph* $\overline{G} := (V, \binom{V}{2} \setminus E)$ consisting of precisely the non-edges of G. By definition, $\alpha(G) = \omega(\overline{G})$. Turán's theorem gives a bound on the number of edges of a graph with a given clique number.

Theorem 4.18. Let $G = (V, E)$ be a graph on $n = \#V$ vertices and $m = \#E$ edges and let $t \in \mathbb{N}$.

(i) If $\omega(G) < t$ then $m \leq \left(1 - \frac{1}{t-1}\right) \cdot \frac{n^2}{2}$.

(ii) If $\alpha(G) < t$ then $m \geq \frac{1}{t-1} \cdot \frac{n^2}{2} - \frac{n}{2}$.

Proof. For (i) see [16]. To prove (ii) we apply (i) to $\overline{G}$. Since $\alpha(G) < t$ we have $\omega(\overline{G}) < t$, so that $\overline{G}$ has at most

$$\left(1 - \frac{1}{t-1}\right) \cdot \frac{n^2}{2}$$

edges. But then G has at least

$$\binom{n}{2} - \left(1 - \frac{1}{t-1}\right) \cdot \frac{n^2}{2} = \frac{1}{t-1} \cdot \frac{n^2}{2} - \frac{n}{2}$$

edges, proving the claim. ∎

The intuition behind Theorem 4.18 (ii) is that if a graph is such that for every t vertices there must be an edge between some of those t vertices, then the graph itself has to contain a sufficiently large number of edges (indeed, for fixed t the graph contains at least a constant fraction of all possible edges). It should be noted that the bound given in Theorem 4.18 (i) is asymptotically tight, for more see [16].
We now come back to bounding the frequency of the most frequent element in a nontrivial, union closed family.

Theorem 4.19. Let $n \in \mathbb{N}$ with $n \geq 2$ and let $\mathcal{F} \subseteq \mathcal{P}(n)$ be a nontrivial, union closed family on $m := \#\mathcal{F}$ sets. Assume

$$m \geq \frac{7n}{\log_2 n}.$$

Then there is an element $x \in [n]$ with

$$\#\{F \in \mathcal{F} : x \in F\} \geq \frac{\sqrt{\log_2 n}}{3n} \cdot \#\mathcal{F}.$$

Proof. By Theorems 4.13 and 4.16 there is a subfamily $\mathcal{E} \subseteq \mathcal{F}$ with $\mu := \#\mathcal{E} \geq \frac{1}{2} \cdot \#\mathcal{F}$ and having the property that, setting $t := \left\lceil \frac{1.368n}{\log_2 n} \right\rceil$, for all $A_1, ..., A_t \in \mathcal{E}$ at least two of the sets have a nonempty intersection. Consider the graph $G := (\mathcal{E}, E(G))$ where

$$E(G) := \left\{ E_1E_2 \in \binom{\mathcal{E}}{2} : E_1 \cap E_2 \neq \emptyset \right\}.$$

By Theorem 4.16, for the independence number of G it holds

$$\alpha(G) < t,$$

so that by Theorem 4.18 (ii) we get

$$\#E(G) \geq \frac{1}{t-1} \cdot \frac{\mu^2}{2} - \frac{\mu}{2}.$$

For every $E_1E_2 \in E(G)$ pick an element $c(E_1E_2) \in E_1 \cap E_2$. This defines an edge coloring of G with n colors. Consequently, there must be a color $x^* \in [n]$ that appears on at least

$$\frac{1}{n} \cdot \left(\frac{1}{t-1} \cdot \frac{\mu^2}{2} - \frac{\mu}{2} \right) \tag{4.2}$$

edges (i.e. for which there are at least this many pairs E_1E_2 with $x^* \in E_1 \cap E_2$). Let G' be the subgraph of G generated by the edges of color x^* (discard isolated vertices that are not incident to an edge of color x^*) and let μ' be the number of vertices in G'. Note that μ' is a lower bound for the number of sets $E \in \mathcal{E}$ with $x^* \in E$. Since G' is a graph on μ' vertices and at least as many edges as described in (4.2), we must have (using $t \leq \frac{1.368n}{\log_2 n} + 1$)

$$\frac{(\mu')^2}{2} \geq \binom{\mu'}{2} \geq \frac{1}{n} \cdot \left(\frac{1}{t-1} \cdot \frac{\mu^2}{2} - \frac{\mu}{2} \right) \geq \frac{\mu^2}{2n \cdot \frac{1.368n}{\log_2 n}} - \frac{\mu}{2n},$$

so that

$$\mu' \geq \sqrt{\frac{\mu^2 \log_2 n}{1.368n^2} - \frac{\mu}{n}}.$$

Notice that the function

$$h_n(s) := \frac{\log_2 n}{1.368n^2} \cdot s^2 - \frac{1}{n} \cdot s$$

is minimized at $s_0 = \frac{1.368n}{2\log_2 n}$. In particular, $h_n(s)$ is increasing for $s \geq s_0$. Since by assumption $\frac{m}{2} \geq s_0$, $h_n(s)$ for $\frac{m}{2} \leq s \leq m$ attains its minimum at $s = \frac{m}{2}$. Thus

$$\sqrt{\frac{\mu^2 \log_2 n}{1.368n^2} - \frac{\mu}{n}} \geq \sqrt{\frac{m^2 \log_2 n}{4 \cdot 1.368n^2} - \frac{m}{2n}} = \sqrt{\frac{1}{4 \cdot 1.368} - \frac{n}{2m \log_2 n}} \cdot \frac{\sqrt{\log_2 n}}{n} \cdot m$$

and using the assumption again we conclude

$$\#\{F \in \mathcal{F} : x^* \in F\} \geq \mu' \geq \sqrt{\frac{1}{4 \cdot 1.368} - \frac{1}{14}} \cdot \frac{\sqrt{\log_2 n}}{n} \cdot m \geq \frac{\sqrt{\log_2 n}}{3n} \cdot m,$$

finishing the proof. ■

The condition on m seems a bit unfortunate, but considering that m can get as large as 2^n we already encapsulate almost all union closed families. We only added it so that the final bound on the frequency looks somewhat cleaner. In [21] (and slightly improved in [31]) it was shown that any nontrivial, separating, union closed family $\mathcal{F} \subseteq \mathcal{P}(n)$ with $\#\mathcal{F} \leq 2n$ contains an abundant element. Thus, for $n \geq 2^{3.5}$ (i.e for $n \geq 12$) the condition on m in Theorem 4.19 is not essential. Indeed, due to [41] it is known that the Union Closed Sets Conjecture 3.1 also holds for all nontrivial, union closed families $\mathcal{F} \subseteq \mathcal{P}(n)$ for $n \leq 12$, so that we may disregard the condition on m entirely.
We again emphasize that this is just a demonstration of how the new insights might be used. Essentially, the above gives the existence of an element contained in an

$$\Omega\left(\frac{\sqrt{\log_2 n}}{n}\right)$$

fraction of the sets from $\mathcal{F}$. Compared to Theorems 3.20 or 3.22, this does not yield any new results.
However, combining known techniques (for example the above cited results) with our new insights might give better asymptotic bounds for the frequency of the most frequent element in a nontrivial, union closed family. For example, we did not even use the fact that $\mathcal{E}$ is also union closed. Still, the Union Closed Sets Conjecture 3.1 in its purest form seems far out of reach.

4.3.3 Bounds on Frequencies for Intersecting Families

The previous section also had us consider other types of set families than union closed ones. In particular, we can ask about the frequencies in families of the following type.

Definition 4.20. Let $n \in \mathbb{N}$ and $\mathcal{F} \subseteq \mathcal{P}(n)$. Then $\mathcal{F}$ is called an **intersecting family** if for all $A, B \in \mathcal{F}$, $A \cap B \neq \emptyset$.

Intersecting families are very well known in combinatorics, especially the Erdős-Ko-Rado theorem [20] about the size of a uniform intersecting family. It does not seem like intersecting families were investigated with respect to the frequency of elements yet.

Question. For $\mathcal{F} \subseteq \mathcal{P}(n)$ an intersecting family, what can be guaranteed about the frequency of the most frequent element in $\mathcal{F}$?

Remark 4.21. The Fano plane

$$\{123, 145, 167, 246, 257, 347, 356\}$$

shows that we might have no element that appears in at least a half of the sets of an intersecting family.
More generally, the Fano plane is a special example of a *projective plane* (in the combinatorial sense). A projective plane of *order* $n \in \mathbb{N}$ is a set family $\mathcal{F} \subseteq \mathcal{P}(n^2+n+1)$, each set having cardinality $n+1$ and every pair $\{x, y\} \subseteq [n^2+n+1]$ being contained

in exactly one set of $\mathcal{F}$. Furthermore, two distinct sets in $\mathcal{F}$ intersect in exactly one element (so in particular projective planes are intersecting) and every $x \in [n^2+n+1]$ is contained in exactly $n+1$ sets of $\mathcal{F}$. For such an $\mathcal{F}$ we thus have

$$\#\{F \in \mathcal{F} : x \in F\} = \frac{n+1}{n^2+n+1} \cdot \#\mathcal{F}.$$

Since $\frac{n+1}{n^2+n+1}$ can get arbitrarily small, for general intersecting families we cannot expect to find an element of the ground set contained in a constant fraction of the sets from the family.
Projective planes themselves are special cases of a much larger class of structures called *(combinatorial) designs*. These objects are very well studied, see for example [6].

We will see projective planes later on below, including how one can construct such families. It should be noted that there are two different types of projective planes: the combinatorial structures as described above and the algebraic/geometric structures as described in Remark 4.26. There, the somewhat strange choice of parameters (n^2+n+1 and $n+1$) should be made clearer. Projective planes in the combinatorial sense are a generalization of projective planes over finite fields in the algebraic sense.
We return to the question, how frequent an element among an intersecting family can be. The proof of the following theorem is a direct adaptation of the proof of Theorem 4.19.

Theorem 4.22. Let $n \in \mathbb{N}$ and let $\mathcal{F} \subseteq \mathcal{P}(n)$ be an intersecting family of size $m = \#\mathcal{F}$. Then there is an $x \in [n]$ with

$$\#\{F \in \mathcal{F} : x \in F\} \geq \sqrt{\frac{m-1}{mn}} \cdot m.$$

Proof. Let $\mathcal{F}$ be as in the statement. Consider the graph $G = (\mathcal{F}, \binom{\mathcal{F}}{2}) \cong K_m$. For every $F_1F_2 \in \binom{\mathcal{F}}{2}$ pick a $c(F_1F_2) \in F_1 \cap F_2$. This defines an edge coloring of G with n colors. Therefore, there must be a color $x^* \in [n]$ that appears on at least

$$\frac{1}{n}\binom{m}{2}$$

edges of G. Let G' be the graph induced by the edges of color x^* and let m' be the number of vertices incident to an edge of color x^*. Then

$$\binom{m'}{2} \geq \frac{1}{n}\binom{m}{2},$$

which gives us

$$(m')^2 - m' - \frac{m^2-m}{n} \geq 0.$$

We thus obtain

$$\#\{F \in \mathcal{F} : x^* \in F\} \geq m' \geq \frac{1}{2} + \sqrt{\frac{1}{4} + \frac{m^2-m}{n}} \geq \sqrt{\frac{m^2-m}{n}} = \sqrt{\frac{m-1}{mn}} \cdot m.$$

■

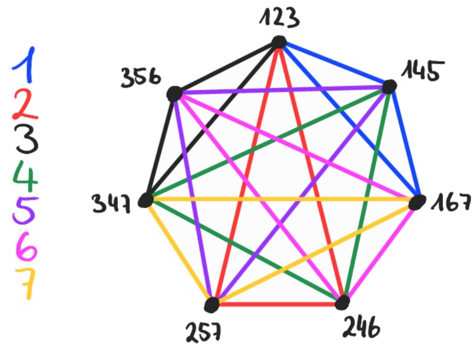

Figure 4.1: The coloring from the proof of Theorem 4.22 for the Fano plane

If we want to make the factor in the above bound independent of m, for $m \geq 2$ (the only case of interest) we can estimate

$$\sqrt{\frac{m-1}{mn}} \geq (2n)^{-1/2}.$$

As the projective planes showed, if we want a bound that only depends on n, this is, up to a constant factor, the best possible that we can guarantee. The projective planes also show that

$$\#\{F \in \mathcal{F} : x \in F\} = \frac{1}{2} + \sqrt{\frac{1}{4} + \frac{m^2 - m}{n}}$$

can hold for all $x \in [n]$.
As the following construction shows, there is no reasonable bound that only depends on m. Consider the complete graph $K_m = ([m], \binom{[m]}{2})$. Define a family

$$\mathcal{F} = \{F_1, ..., F_m\} \subseteq \mathcal{P}\left(\binom{[m]}{2}\right)$$

with $F_i := \{X \in \binom{[m]}{2} : i \in X\}$ for $i \in [m]$. Then for $i \neq j$ it holds $F_i \cap F_j = \{\{i, j\}\}$, so $\mathcal{F}$ is intersecting. Also, every $\{i, j\} \in \binom{[m]}{2}$ is contained in exactly two sets, namely in F_i and F_j.
Clearly, every intersecting family on $m \geq 2$ sets contains an element that is contained in at least 2 sets from the family. Thus, for every intersecting family $\mathcal{F}$ of size $m = \#\mathcal{F}$, if we want a factor that only depends on m, we can generally only guarantee the existence of an element x from the (here unspecified) ground set with

$$\#\{F \in \mathcal{F} : x \in F\} \geq \frac{2}{m} \cdot \#\mathcal{F} = 2.$$

If $\mathcal{F} \subseteq \mathcal{P}(n)$ is a nontrivial family and $x \in [n]$, then $\{F \in \mathcal{F} : x \in F\}$ is clearly intersecting (namely in x). Thus, the following is again a weakening of the Union Closed Sets Conjecture 3.1.

Conjecture 4.23. Let $n \in \mathbb{N}$ and $\mathcal{F} \subseteq \mathcal{P}(n)$ be a nontrivial, union closed family. Then there exists an intersecting subfamily $\mathcal{E} \subseteq \mathcal{F}$ with $\#\mathcal{E} \geq \frac{1}{2} \cdot \#\mathcal{F}$.

If the Union Closed Sets Conjecture 3.1 holds and x is an abundant element in $\mathcal{F}$, then $\mathcal{E} := \mathcal{F} \cap [x, [n]]$ is an intersecting family with $\#\mathcal{E} \geq \frac{1}{2} \cdot \#\mathcal{F}$. Conjecture 4.23 generalizes Conjecture 4.9. If we would be able to prove the above conjecture though we would get (combining it with Theorem 4.22) for all nontrivial, union closed $\mathcal{F} \subseteq \mathcal{P}(n)$ the existence of an element $x \in [n]$ with

$$\#\{F \in \mathcal{F} : x \in F\} \geq \frac{1}{2\sqrt{2n}} \cdot \#\mathcal{F}$$

(using the small remark after the proof of Theorem 4.22). This is again worse than Theorem 3.22, but again, we did not use the fact that $\mathcal{E}$ is intersecting and union closed.

Question. For an intersecting, union closed family $\mathcal{F} \subseteq \mathcal{P}(n)$, what can be guaranteed about the frequency of the most frequent element $x \in [n]$ in $\mathcal{F}$?

4.4 Generalizing Union Closedness

One can ask how important the condition of being union closed actually is. To this end, we investigate families closed under other set operations.

4.4.1 A q-Analogue

In combinatorics it can be expected that many statements about sets can be carried over to statements about subspaces of finite dimensional vector spaces over finite fields. Then, we are not so interested in the cardinality of the subspaces but rather their dimension. An example of that is [13], where this was done for the Kruskal-Katona theorem. We give a full account of the general strategy for such considerations. If the ground field is the finite field $\mathrm{GF}(q)$ with $q = p^k$ elements (p prime), then these statements are referred to as *q-analogues* of the original problem. As far as the author is aware of, there are no such investigations with regards to the Union Closed Sets Conjecture 3.1.
We first introduce some notation. Let $q = p^k$ be a prime power and $\mathrm{GF}(q)$ be the field with q elements. Then $\mathrm{GF}(q)^n$ is canonically a $\mathrm{GF}(q)$-vector space. Denote by $\mathcal{L}_q(n)$ (often called the *linear lattice*) the family of all subspaces of $\mathrm{GF}(q)^n$. If $V, W \in \mathcal{L}_q(n)$, we can define their *sum* $V + W$ as the $\mathrm{GF}(q)$-linear span of $V \cup W$ (equivalently, the intersection of all subspaces containing $V \cup W$). In what follows, we will study *subspace families* $\mathcal{Q} \subseteq \mathcal{L}_q(n)$.

Definition 4.24. Let $n \in \mathbb{N}$ and $q = p^k$ be a prime power. We call the subspace family $\mathcal{Q} \subseteq \mathcal{L}_q(n)$ **sum closed** if for $V, W \in \mathcal{Q}$ also $V + W \in \mathcal{Q}$.

We again abbreviate “sum closed subspace family” simply to *sum closed family.*

Remark 4.25.

(i) One has to be a bit cautious when one wants to translate the terminology from set families to the setting of subspace families. Since all subspaces contain the zero of $\mathrm{GF}(q)^n$, we call a subspace family $\mathcal{Q} \subseteq \mathcal{L}_q(n)$ *nontrivial* if $\bigcup \mathcal{Q} = \mathrm{GF}(q)^n$ and $\bigcap \mathcal{Q} = \{0\}$. For sum closed families $\mathcal{Q}$ we have $\bigcup \mathcal{Q} = \mathrm{GF}(q)^n$ if and only if $\mathrm{GF}(q)^n \in \mathcal{Q}$.

(ii) A priori, it does not make sense to talk of separability for subspace families. This is due to the fact that if $x \in V \in \mathcal{L}_q(n)$, then also $\lambda x \in V$ for all $\lambda \in \mathrm{GF}(q)$. This suggest that we have to work in a projective setting. To make this more precise, let $V \in \mathcal{L}_q(n)$ be a subspace. On $V \setminus \{0\}$ define the equivalence relation $\sim$ by

$$x \sim y \quad :\Leftrightarrow \quad \exists \lambda \in \mathrm{GF}(q) \setminus \{0\} : x = \lambda y.$$

The *projective space* is then the quotient $\mathbb{P}V := (V \setminus \{0\})/\sim$. Alternatively, we could also write $\mathbb{P}V = (V \setminus \{0\})/\,\mathrm{GF}(q)^\times$ where $\mathrm{GF}(q)^\times = \mathrm{GF}(q) \setminus \{0\}$ acts on V via scalar multiplication (this defines a group action of $\mathrm{GF}(q)^\times$ on $V \setminus \{0\}$). Here we use the convention that $\mathbb{P}\{0\} = \emptyset$. The elements of $\mathbb{P}V$ may be seen as "lines" in V through the origin 0. For $\mathcal{Q} \subseteq \mathcal{L}_q(n)$ we then set $\mathbb{P}\mathcal{Q} := \{\mathbb{P}V : V \in \mathcal{Q}\}$. In $\mathbb{P}\mathcal{Q}$, which can be seen as a set family over $\mathbb{P}\,\mathrm{GF}(q)^n$, it now makes sense to speak of separation of elements. We then call a subspace family $\mathcal{Q} \subseteq \mathcal{L}_q(n)$ *separating*, if $\mathbb{P}\mathcal{Q} \subseteq \mathcal{P}(\mathbb{P}\,\mathrm{GF}(q)^n)$ is separating as a set family.

(iii) Using the language introduced in (ii), we can also define $\mathbb{P}V + \mathbb{P}W := \mathbb{P}(V + W)$ in $\mathbb{P}\,\mathrm{GF}(q)^n$. This is well defined, as the map $\mathbb{P}\colon \mathcal{L}_q(n) \to \mathbb{P}\mathcal{L}_q(n)$ is a bijection. We can therefore speak of *sum closed* families $\mathbb{P}\mathcal{Q} \subseteq \mathcal{P}(\mathbb{P}\,\mathrm{GF}(q)^n)$. Note that in general $\mathbb{P}V \cup \mathbb{P}W \subsetneq \mathbb{P}V + \mathbb{P}W$ (often $\mathbb{P}V \cup \mathbb{P}W$ is not even of the form $\mathbb{P}U$ for some $U \in \mathcal{L}_q(n)$). This allows us to switch between $\mathcal{L}_q(n)$ and $\mathbb{P}\mathcal{L}_q(n)$.

Remark 4.26. Continuing on 4.21, we now have a way to construct projective planes in the combinatorial sense whenever $n = p^k$ is a prime power. For this take the the projective versions of all the two-dimensional subspaces of $\mathrm{GF}(q)^3$. This gives a set family over

$$\frac{n^3 - 1}{n - 1} = n^2 + n + 1$$

elements and every set from the family contains

$$\frac{n^2 - 1}{n - 1} = n + 1$$

elements.

We emphazise again that generally

$$\mathbb{P}\mathcal{L}_q(n) \subsetneq \mathcal{P}(\mathbb{P}\,\mathrm{GF}(q)^n)$$

and that the structure on $\mathbb{P}\mathcal{L}_q(n)$ comes via subspace sums rather than set unions.
The above remark suggest a similarity between union closed families and sum closed families $\mathcal{Q}$ (or rather their projective versions $\mathbb{P}\mathcal{Q}$). In particular, one might try to

formulate a conjecture similar to the Union Closed Sets Conjecture 3.1 but for sum closed families. However, it is not so clear how the factor of "$\frac{1}{2}$" needs to change in this q-analogue situation. For this, we have the following remark which benchmarks the factor against the entire space $\mathcal{L}_q(n)$.

Remark 4.27. An adapted version of the Union Closed Sets Conjecture 3.1 certainly has to hold in the case of $\mathcal{Q} = \mathcal{L}_q(n)$. A basic counting argument (via ordered bases) shows that there are

$$\binom{n}{d}_q := \frac{(q^n-1)...(q^{n-d+1}-1)}{(q^d-1)...(q-1)}$$

d-dimensional subspace in $\mathcal{L}_q(n)$ (see [12]). These numbers are often referred to as q-*binomial coefficients* or *Gaussian coefficients* (see [28, 29]) and one recovers the common binomial coefficient by taking $q \longrightarrow 1$ (since $\lim\limits_{q \longrightarrow 1} \frac{q^n-1}{q-1} = n$). One thus sees

$$\#\mathcal{L}_q(n) = \sum_{d=0}^{n} \binom{n}{d}_q .$$

Among those, there are

$$\sum_{d=1}^{n} \binom{n-1}{d-1}_q$$

subspaces that contain a fixed element $x \neq 0$. Thus, every nonzero x appears in a

$$\frac{\#\{V \in \mathcal{L}_q(n) : x \in V\}}{\#\mathcal{L}_q(n)} = \frac{\sum_{d=1}^{n} \binom{n-1}{d-1}_q}{\sum_{d=0}^{n} \binom{n}{d}_q} =: \kappa_{q,n}$$

fraction of all subspaces from $\mathcal{L}_q(n)$. This shows that if a q-analogue of the Union Closed Sets Conjecture 3.1 holds, then the factor must be at most $\kappa_{q,n}$. Note that this factor already can be very small for large q and n ($\kappa_{q,n}$ is roughly of the order $q^{-\frac{n}{2}+\frac{1}{4}}$ since the dominating term in $\binom{n}{d}_q$ is $q^{d(n-d)}$).

Given the above remark, we cautiously state the following conjecture.

Conjecture 4.28. Let $n \in \mathbb{N}$, $q = p^k$ a prime power, $\mathcal{Q} \subseteq \mathcal{L}_q(n)$ a sum closed family and

$$\kappa_{q,n} = \frac{\sum_{d=1}^{n} \binom{n-1}{d-1}_q}{\sum_{d=0}^{n} \binom{n}{d}_q}.$$

Then there is an $x \in \mathrm{GF}(q)^n \setminus \{0\}$ such that

$$\#\{V \in \mathcal{Q} : x \in V\} \geq \kappa_{q,n} \cdot \#\mathcal{Q}.$$

We will refer to it as the q-analogue of the Union Closed Sets Conjecture 3.1 or perhaps just as the Sum Closed Sets Conjecture. It should also be noted that, in a specific algebraic sense, the Union Closed Sets Conjecture 3.1 may be recovered from its q-analogue by taking $q = 1$ and working with a suitable theory of the *field with one element* (the set of "vector spaces of dimension $\leq n$" over such a theoretical object

would then correspond to $\mathcal{P}(n)$). For more on that see [12].
One now has the opportunity to take everything known about union closed families and to try to prove analogous statements for sum closed families. We will demonstrate this with the following theorem. It can be seen as an adapted version of the "folklore" proof that if a union closed family $\mathcal{F}$ contains a singleton $\{x\} \in \mathcal{F}$, then x is abundant in $\mathcal{F}$ (see [10]).

Theorem 4.29. Let $n \in \mathbb{N}$, $q = p^k$ a prime power and $\mathcal{Q} \subseteq \mathcal{L}_q(n)$ a nontrivial, sum closed family. Assume $U \in \mathcal{Q}$ is a one-dimensional subspace of $\mathrm{GF}(q)^n$ and $x \in U \setminus \{0\}$, then

$$\#\{V \in \mathcal{Q} : x \in V\} \geq \frac{1}{1+q^{n-1}} \cdot \#\mathcal{Q}.$$

Proof. Let $0 \neq x \in U \in \mathcal{Q}$ as in the assumptions. Consider the function

$$\varphi\colon \{V \in \mathcal{L}_q(n) : x \notin V\} \to \{W \in \mathcal{L}_q(n) : x \in W\}, V \mapsto V + U.$$

How large can $\#\varphi^{-1}(W)$ get? Let $W \in \mathcal{L}_q(n)$ with $x \in W$ and $\dim W = d$. If we want to have $V + U = W$, then V has to be a subspace of W of dimension $\dim V = d - 1$ with $x \notin V$. Every such V contains

$$(q^{d-1} - 1)(q^{d-1} - q)...(q^{d-1} - q^{d-2})$$

ordered bases $(v_1, ..., v_{d-1})$. Indeed, there are $q^{d-1} - 1$ possibilities to choose the first vector v_1 (every vector in V except the zero vector), there are $q^{d-1} - q$ possibilities to choose the second vector v_2 (every vector in V except for the ones in $\mathrm{span}\{v_1\}$), there are $q^{d-1} - q^2$ possibilities to choose the third vector v_3 (every vector in V except for the ones in $\mathrm{span}\{v_1, v_2\}$), and so on.
Also note that there are

$$(q^d - q)(q^d - q^2)...(q^d - q^{d-1})$$

ordered bases of W of the form $(x, w_1, ..., w_{d-1})$. Observe that for every such ordered basis, the space $V := \mathrm{span}\{w_1, ..., w_{d-1}\}$ fulfills $x \notin V$ and $U + V = W$. Also, if V is any such space and $(v_1, ..., v_{d-1})$ is an ordered basis of V, then $(x, v_1, ..., v_{d-1})$ is an ordered basis of W of the above mentioned type.
By symmetry, for any given V as above, the number of ordered bases $(x, w_1, ..., w_{d-1})$ of W so that $V = \mathrm{span}\{w_1, ..., w_{d-1}\}$ is independent of the choice of V. We conclude

$$\#\{V \in \mathcal{L}_q(n) : x \notin V, U + V = W\} = \frac{(q^d - q)(q^d - q^2)...(q^d - q^{d-1})}{(q^{d-1} - 1)(q^{d-1} - q)...(q^{d-1} - q^{d-2})} = q^{d-1}.$$

This also gives

$$\#\varphi^{-1}(W) = q^{\dim W - 1} \leq q^{n-1}$$

for all $W \in \mathcal{L}_q(n)$ with $x \in W$. Since φ restricts to a map

$$\varphi\colon \{V \in \mathcal{Q} : x \notin V\} \to \{W \in \mathcal{Q} : x \in W\},$$

we thus have

$$q^{n-1} \cdot \#\{W \in \mathcal{Q} : x \in W\} \geq \#\{V \in \mathcal{Q} : x \notin V\},$$

proving the theorem. ∎

Note that the factor $\frac{1}{1+q^{n-1}} \approx q^{-n+1}$ is in general much smaller than the desired factor $\kappa_{q,n} \approx q^{-\frac{n}{2}+\frac{1}{4}}$. How can this be?
Inspecting the proof, this is due to the fact that $\#\varphi^{-1}(W)$ depends on $\dim W$, in particular it can take different values for different W's. Compare this with Lemma 4.6, where we did not have this issue in the proof there and indeed the bound there can be sharp. Due to the fact that $\#\varphi^{-1}(W)$ depends on W, the bound in Theorem 4.29 can be assumed to almost never be sharp.
Thus, there is still hope to raise the factor all the way to $\kappa_{q,n}$. The following discussion offers a heuristic justification, that a modification to the above proof might yield this better constant already. Consider the weighted average

$$\frac{\sum_{d=1}^{n} q^{d-1}\binom{n-1}{d-1}_q}{\sum_{d=1}^{n}\binom{n-1}{d-1}_q}.$$

The dominating term of the summand corresponding to d in the numerator is

$$q^{(d-1)(n-d+1)},$$

so that the dominating term in the numerator corresponds to $d-1=\frac{n}{2}$ (as to maximize $(d-1)(n-d+1)$, also disregarding rounding). Similarly, the dominating term of the summand corresponding to d in the denominator is

$$q^{(d-1)(n-d)},$$

so that the dominating term in the denominator corresponds to $d=\frac{n+1}{2}$ (as to maximize $(d-1)(n-d)$, again disregarding rounding). Thus, the weighted average above is of order roughly

$$q^{\left(\frac{n}{2}\right)^2-\left(\frac{n-1}{2}\right)^2} = q^{\frac{n}{2}-\frac{1}{4}}.$$

Compared to $\kappa_{n,q} \approx q^{-\frac{n}{2}+\frac{1}{4}}$, this gives roughly the same asymptotic.

Question. Can the above, informal discussion be applied to prove Theorem 4.29 with the constant $\frac{1}{1+q^{n-1}}$ replaced by $\kappa_{q,n}$?

4.4.2 A Vaster Setting

The considerations in the previous section can further be generalized. For example, what if instead of linear subspaces, we consider the subrings or ideals of a (finite) ring? Or for that sake any structure, so that the system of substructures is intersection closed? A very general question would then be the following.

Remark 4.30. Let $n \in \mathbb{N}$ and let $\mathfrak{N} \subseteq \mathcal{P}(n)$ be a nontrivial, intersection closed family with $[n] \in \mathfrak{N}$. As we have already done with union closed family, $\mathfrak{N}$ is a lattice. The meet operation is just set intersection, the join is given by

$$A \vee B = \bigcap\{N \in \mathfrak{N} : A \subseteq N, B \subseteq N\}.$$

One can now study the structure of nontrivial, $\vee$-closed families $\mathcal{F} \subseteq \mathfrak{N}$. We get the theory for union closed families for $\mathfrak{N} = \mathcal{P}(n)$, and the case for sum closed families if $\mathfrak{N}$ consists of the (projective versions of) subspaces of $\mathrm{GF}(q)^n$.

Question. For a given $\mathfrak{N} \subseteq \mathcal{P}(n)$, what is the largest constant $\kappa = \kappa(\mathfrak{N})$, such that for any nontrivial, $\vee$-closed family $\mathcal{F} \subseteq \mathfrak{N}$ there is an element $x \in [n]$ with

$$\#\{F \in \mathcal{F} : x \in F\} \geq \kappa \cdot \#\mathcal{F}?$$

The vast scope that the above question captures makes it almost impossible to answer in a satisfactory way. However single cases may be answered quite easily.

Example 4.31. If $\mathfrak{N} = \{[k] : k \in \{0, 1, ..., n\}\}$, then any $\mathcal{F} \subseteq \mathfrak{N}$ is $\vee$-closed (since $\vee$ is basically just the maximum here). For any nontrivial $\mathcal{F} \subseteq \mathfrak{N}$ it holds

$$\#\{F \in \mathcal{F} : 1 \in F\} \in \{\#\mathcal{F}, \#\mathcal{F} - 1\},$$

depending on whether $\emptyset \in \mathcal{F}$ or not. In any case

$$\#\{F \in \mathcal{F} : 1 \in F\} \geq \frac{1}{2} \cdot \#\mathcal{F}.$$

This shows that $\kappa(\mathfrak{N}) \geq \frac{1}{2}$ and $\mathcal{F} = \{\emptyset, [n]\}$ even gives equality.

The above example is rather simple. For the Union Closed Sets Conjecture 3.1 we would like to know about $\mathfrak{N} = \mathcal{P}(n)$. One way one could get new insights is by considering the following.

Question. Does there exist a sequence $\{\emptyset, [n]\} = \mathfrak{N}_2 \subsetneq \mathfrak{N}_3 \subsetneq ... \subsetneq \mathfrak{N}_{2^n} = \mathcal{P}(n)$ of intersection closed families $\mathfrak{N}_i \subseteq \mathcal{P}(n)$, such that for all $i = 2, ..., 2^n$ it holds

$$\kappa(\mathfrak{N}_i) \geq \frac{1}{2}?$$

The idea of the above is in a way another inductive approach to the Union Closed Sets Conjecture 3.1, but instead of building up the union closed family $\mathcal{F}$, we build up the "universe" around it. We leave this question for future research.

5 Final Remarks

The Union Closed Sets Conjecture 3.1 is still an open and very active part of research in combinatorics. The author hopes that this thesis sheds some new light on union closed families and their famous conjecture. To summarize, the main idea of this thesis was to investigate the inner structure of a union closed family $\mathcal{F}$ to obtain new insights into the behaviour of $\mathcal{F}$. By doing so, we obtain a subfamily $\mathcal{E} \subseteq \mathcal{F}$ containing at least half of the sets from $\mathcal{F}$ and possessing some additional structure. This additional structure allowed us to conclude some statements about the frequency of elements in $\mathcal{E}$ which, up to a factor of $\frac{1}{2}$, translates to a statement about frequencies in $\mathcal{F}$. While we were not able to do so in this thesis, we hope that this approach can at least yield better asymptotics on the frequency of elements in $\mathcal{F}$.
We also stated some new conjectures that could give some new approaches to the Union Closed Sets Conjecture 3.1. To finish, we summarize the most important open questions discussed in this thesis:

(1) Conjecture 4.4: Does the k-th most frequent element in a nontrivial, union closed family $\mathcal{F}$ appear in at least $\frac{1}{2^{k-1}+1} \cdot \#\mathcal{F}$ many sets from $\mathcal{F}$?

(2) Conjecture 4.23: Does every nontrivial, union closed family $\mathcal{F}$ contain an intersecting subfamily $\mathcal{E} \subseteq \mathcal{F}$ of cardinality $\#\mathcal{E} \geq \frac{1}{2} \cdot \#\mathcal{F}$? What can we say about the frequency of the most frequent element in a nontrivial, union closed, intersecting family $\mathcal{F}$?

(3) Remark 4.30 and the discussion after: What can we say about nontrivial, $\vee$-closed families in an intersection closed family $\mathfrak{N}$? Can we build up a chain $\{\emptyset, [n]\} = \mathfrak{N}_2 \subsetneq ... \subsetneq \mathfrak{N}_{2^n} = \mathcal{P}(n)$ of intersection closed families, such that every nontrivial, $\vee$-closed family $\mathcal{F} \subseteq \mathfrak{N}_i$ has an element appearing in at least half of all sets from $\mathcal{F}$?

Acknowledgement

I would like to thank my advisers Prof. Dr. Christoph Helmberg and Dr. Martin Winter for assisting in the making of this thesis. Especially the discussions with Dr. Martin Winter helped in developing the results presented here. I would also like to thank the entire working group of the faculty on algorithmic and discrete mathematics at the TU Chemnitz, that always had an open ear for my endeavours in their regular seminar. Together with Dr. Martin Winter, I would like to thank Dr. Thomas Jahn and Kurt Klement Gottwald from the TU Chemnitz for advices outside of a purely mathematical setting that nevertheless helped in the creation of this thesis.
I would also like to thank Prof. Dr. Tibor Szabó from the FU Berlin, who also suggested new and interesting directions to investigate.

Bibliography

[1] J. Aaronson, D. Ellis, I. Leader, A note on transitive union-closed families, *arXiv:2010.08795* (2020)

[2] V. B. Alekseev, On the number of intersection semilattices [in Russian], *Diskretnaya Mat.* **1**, pp. 129-136 (1989)

[3] B. Amaral, L. Dalton, D. Polakowski, A. Raymond, B. Thomas, The Linear Relaxation of an Integer Program for the Union-Closed Conjecture, *arXiv:2004.05210* (2020)

[4] I. Balla, Minimum density of union-closed families, *arXiv:1106.0369* (2011)

[5] I. Balla, B. Bollobás, T. Eccles, Union-closed families of sets, *Journal of Combinatorial Theory, Series A* **120**, Issue 3, pp. 531-544 (2013)

[6] T. Beth, D. Jungnickel, H. Lenz, Design Theory, *Cambridge University Press*, Volume 1, Second Edition (1999)

[7] P. Billingsley, Probability and Measure, *New York: Wiley*, Second Edition (1986)

[8] P. Billingsley, Convergence of Probability Measures, *New York: Wiley*, Second Edition (1999)

[9] G. Brinkmann, R. Deklerck, Generation of Union-Closed Sets and Moore Families, *Journal of Integer Sequences* **21**, Article 18.1.7 (2018)

[10] H. Bruhn, O. Schaudt, The Journey of the Union-Closed Sets Conjecture, *Graphs and Combinatorics* **31**, pp. 2043-2074 (2015)

[11] N. Caspard, B. Monjardet, The lattices of closure systems, closure operators,and implicational systems on a finite set:a survey, *Discrete Applied Mathematics* **127**, Issue 2, pp. 241-269 (2003)

[12] H. Cohn, Projective Geometry over F_1 and the Gaussian Binomial Coefficients, Vol. 111, No. 6, pp. 487-495 (2005)

[13] A. Chowdhury, B. Patkós, Shadows and intersections in vector spaces, *Journal of Combinatorial Theory, Series A* **117**, pp. 1095-1106 (2010)

[14] B. A. Davey, H. A. Priestley, Introduction to Lattices and Order, *Cambridge University Press*, Second Edition (2002)

[15] A. Day, The Lattice Theory of Functional Dependencies and Normal Decompositions, *International Journal of Algebra and Computation* **2**, Number 4, pp. 409-431 (1992)

[16] R. Diestel, Graph Theory, *Springer-Verlag, Graduate Texts in Mathematics*, Volume 173, Fifth Edition (2016/17)

[17] D. Ellis, M.-R. Ivan, I. Leader, Small sets in union-closed families, *arXiv:2201.11484* (2022)

[18] K. Engel, H. D. Gronau, Sperner Theory in Partially Ordered Sets, *Teubner-Texte zur Mathematik* **78** (1985)

[19] P. Erdős, T. Gallai, Gráfok előírt fokszámú pontokkal, *Matematikai Lapok* **11**, pp. 264-274 (1960)

[20] P. Erdős, C. Ko, R. Rado, Intersection Theorems for Systems of Finite Sets, *The Quarterly Journal of Mathematics*, Volume 12, Issue 1, pp. 313-320 (1961)

[21] V. Falgas-Ravry, Minimal Weight in Union-Closed Families, *The Electronic Journal of Combinatorics*, Vol 18, Issue 1, P95 (2011)

[22] P. Fránkl, On the Trace of Finite Sets, *Journal of Combinatorial Theory, Series A* **34**, Issue 1, pp. 41-45 (1983)

[23] P. Fránkl, A new short proof for the Kruskal-Katona theorem, *Discrete Mathematics* **48**, Issues 2–3, pp. 327-329 (1984)

[24] W. T. Gowers, et al., *gowers.wordpress.com/2016/01/21/frankls-union-closed-conjecture-a-possible-polymath-project/* (2016)

[25] M. Harris, Number of ACI algebras (or semilattices) on n generators, *OEIS A102897* (2005)

[26] I. Karpas, Two Results on Union-Closed Families, *arXiv:1708.01434* (2017)

[27] E. Knill, Graph Generated Union-closed Families of Sets, *arXiv:math/9409215* (1994)

[28] J. Konvalina, Generalized Binomial Coefficients and the Subset-Subspace Problem, *Advances in Applied Mathematics*, Volume 21, Issue 2, pp. 228-240 (1998)

[29] J. Konvalina, A unified interpretation of the Binomial Coefficients, the Stirling numbers, and the Gaussian coefficients, *The American Mathematical Monthly*, Vol. 107, No. 10, pp. 901-910 (2000)

[30] U. Leck, I. T. Roberts, J. Simpson, Minimizing the weight of the union-closure of families of two-sets, *Australasian Journal of Combinatorics* **52**, pp. 67-73 (2012)

[31] J. Massberg, The Union-Closed Sets Conjecture for Small Families, *Graphs and Combinatorics* **32**, pp. 2047-2051 (2016)

[32] Y. Peng, P. Sissokho, C. Zhao, An extremal problem for set families generated with the union and symmetric difference operations, *Journal of Combinatorics* **3**, Number 4, pp. 651-668 (2012)

[33] B. Poonen, Union-Closed Families, *Journal of Combinatorial Theory, Series A* **59**, Issue 2, pp. 253-268 (1992)

[34] D. Reimer, An Average Set Size Theorem, *Combinatorics, Probability and Computing* **12**, Issue 1, pp. 89-93 (2003)

[35] J. C. Renaud, L. F. Fitina, On union-closed sets and Conway's sequence, *Bulletin of the Australian Mathematical Society* **47**, pp. 321-332 (1993)

[36] D. G. Sarvate, J. C. Renaud, On the union-closed sets conjecture, *Ars Combinatoria* **27**, pp. 149–154 (1989)

[37] D. G. Sarvate, J. C. Renaud, Improved bounds for the union-closed sets conjecture, *Ars Combinatoria* **29**, pp. 181-185 (1990)

[38] L. Studer, An asymptotic version of the union-closed sets conjecture, *arXiv:2005.00361* (2020)

[39] C. Tian, Union-closed Sets Conjecture Holds for Height $H(\mathcal{F}) \leq 3$ and $H(\mathcal{F}) \geq n-1$, *arXiv:2112.06659* (2021)

[40] W. T. Trotter, Combinatorics and Partially Ordered Sets: Dimension Theory, *Johns Hopkins Series in the Mathematical Sciences* (1992)

[41] B. Vučković, M. Živković, The 12-Element Case of Frankl's Conjecture, *preprint* (2017)

[42] P. Wójcik, Union-closed families of sets, *Discrete Mathematics* **199**, Issues 1-3, pp. 173-182 (1999)