Parameters for minimal unsatisfiability: Smarandache primitive numbers and full clauses*

Oliver Kullmann[1] and Xishun Zhao[2]

[1] Computer Science Department, Swansea University
[2] Institute of Logic and Cognition, Sun Yat-sen University

Abstract. We establish a new bridge between propositional logic and elementary number theory. A *full clause* contains all variables, and we study them in minimally unsatisfiable clause-sets (MU); such clauses are strong structural anchors, when combined with other restrictions. Counting the maximal number of full clauses for a given *deficiency* k, we obtain a close connection to the so-called "Smarandache primitive number" $S_2(k)$, the smallest n such that 2^k divides $n!$.
The deficiency $k \geq 1$ of an MU is the difference between the number of clauses and the number of variables. We also consider the subclass UHIT of MU given by unsatisfiable hitting clause-sets (every two clauses clash). We study the four fundamental quantities $\mathrm{FCH}, \mathrm{FCM}, \mathrm{VDH}, \mathrm{VDM} : \mathbb{N} \to \mathbb{N}$, defined as the maximum number of full clauses in UHIT resp. MU, resp. the maximal minimal number of occurrences of a variable (the variable degree) in UHIT resp. MU, in dependency on the deficiency. We have the relations $\mathrm{FCH}(k) \leq \mathrm{FCM}(k) \leq \mathrm{VDM}(k)$ and $\mathrm{FCH}(k) \leq \mathrm{VDH}(k) \leq \mathrm{VDM}(k)$, together with $\mathrm{VDM}(k) \leq \mathrm{nM}(k) \leq k + 1 + \log_2(k)$, using the "non-Mersenne numbers" $\mathrm{nM}(k)$ as established in [21].
We show the lower bound $S_2(k) \leq \mathrm{FCH}(k)$; indeed we conjecture this to be exact. The proof rests on two methods: Applying *subsumption resolution* and its inverse, and analysing certain recursions, combining an application-specific recursion with a recursion from the field of *meta-Fibonacci sequences*. The S_2-lower bound together with the nM-upper-bound yields a good handle on the four quantities, which we determine for $1 \leq k \leq 13$.

1 Introduction

Long clauses often occur in practical instances; we study the most extreme case, the occurrences of *full clauses*, clauses of maximal possible length, in minimal unsatisfiable clause-sets ($F \in \mathcal{MU}$). The main parameter is the deficiency $\delta(F) = c(F) - n(F) \geq 1$, the number of clauses minus the number of variables. We denote by $\mathrm{FCM}(k)$ the maximal possible number of full clauses in $F \in \mathcal{MU}$ with $\delta(F) = k$ (short: $F \in \mathcal{MU}_{\delta=k}$). From [21, Theorem 15] follows the upper bound $\mathrm{FCM}(k) \leq \mathrm{nM}(k)$ for the *non-Mersenne numbers* $\mathrm{nM}(k) \in \mathbb{N}$, with

* This research was partially supported by NSFC Grand 61272059, NSSFC Grant 13&ZD186 and MOE Grant 11JJD7200020.

$k + \lfloor \log_2(k+1) \rfloor \le \mathrm{nM}(k) \le k + 1 + \lfloor \log_2(k) \rfloor$ ([21, Corollary 10]). Until now no general lower bound on $\mathrm{FCM}(k)$ was known, and we establish $S_2(k) \le \mathrm{FCM}(k)$. Here $S_2(k)$, as introduced in [28], is the smallest $n \in \mathbb{N}_0$ such that 2^k divides $n!$, and various number-theoretical results on S_2 and the generalisation S_p for prime numbers p are known. Actually we show a stronger lower bound, namely we do not consider all $F \in \mathcal{MU}_{\delta=k}$, but only those F which are *hitting* ($F \in \mathcal{UHIT}$), that is, where every two clauses clash, yielding $\mathrm{FCH}(k)$ with $\mathrm{FCH}(k) \le \mathrm{FCM}(k)$, and we show $S_2 \le \mathrm{FCH}$. The elements of $\mathcal{UHIT}$ are known in the DNF language as "orthogonal" or "disjoint" tautological DNF, and when considering arbitrary boolean functions, then also "disjoint sums of products" (DSOP) or "disjoint cube representations" are used; see [25, Section 4.4] or [4, Chapter 7].

The central underlying research question is the programme of classification of $\mathcal{MU}$ in the deficiency, that is, the characterisation of the layers $\mathcal{MU}_{\delta=k}$ for $k \in \mathbb{N}$. A special case of the general classification is the classification of $\mathcal{UHIT}_{\delta=k}$. The earliest source [1] showed (in modern notation) $\delta(F) \ge 1$ for $F \in \mathcal{MU}$, and characterised the special case $\mathcal{SMU}_{\delta=1} \subset \mathcal{MU}_{\delta=1}$, where $\mathcal{SMU} \subset \mathcal{MU}$ contains those $F \in \mathcal{MU}$ such that no literals can be added to any clauses without destroying unsatisfiability. Later [5] characterised $\mathcal{MU}_{\delta=1}$ via matrices, while the intuitive characterisation via binary trees was given in [16, Appendix B], where also $\mathcal{SMU}_{\delta=1} = \mathcal{UHIT}_{\delta=1}$ has been shown. In the form of "S-matrices", the class $\mathcal{MU}_{\delta=1}$ had been characterised earlier in [13,11], going back to a conjecture on Qualitative Economics ([7]), and where the connections to this field of matrix analysis, called "Qualitative Matrix Analysis (QMA)", where first revealed in [18] (see [15, Subsection 11.12.1] and [23, Subsection 1.6.4] for overviews). $\mathcal{SMU}_{\delta=2}$ and partially $\mathcal{MU}_{\delta=2}$ were characterised in [14], with further information on $\mathcal{MU}_{\delta=2}$ in [22]. [6] showed that all layers $\mathcal{MU}_{\delta=k}$ are poly-time decidable.

A key element for these investigations into the structure of $\mathcal{MU}$ is the *min-var-degree* $\mu\mathrm{vd}(F) := \min_{v \in \mathrm{var}(F)} |\{C \in F : v \in \mathrm{var}(C)\}|$, the minimal variable-degree of F, and its maximum $\mathrm{VDM}(k)$ over all $F \in \mathcal{MU}_{\delta=k}$. Indeed the key to the characterisation of $\mathcal{MU}_{\delta=1}$ in [5] as well as in [13] was the proof of $\mathrm{VDM}(1) = 2$. The first general upper bound $\forall\, k \in \mathbb{N} : \mathrm{VDM}(k) \le 2k$ was shown in [16, Lemma C.2]. Now in [21], mentioned above, we actually showed the upper bound $\mathrm{VDM}(k) \le \mathrm{nM}(k)$. Using $\mathrm{fc}(F)$ for the number of full clauses in F, obviously $\mathrm{fc}(F) \le \mu\mathrm{vd}(F)$ holds, and $\mathrm{FCM}(k)$ is the maximum of $\mathrm{fc}(F)$ over all $F \in \mathcal{MU}_{\delta=k}$, thus $\mathrm{FCM}(k) \le \mathrm{VDM}(k)$.

For the variation $\mathrm{VDH}(k) \le \mathrm{VDM}(k)$, which only considers hitting clause-sets, we conjecture $\mathrm{VDH}(k) = \mathrm{VDM}(k)$ for all $k \ge 1$. Furthermore we conjecture $\mathrm{FCM}(k) \ge \mathrm{nM}(k) - 1$, and thus the quantities $\mathrm{nM}(k), \mathrm{VDM}(k), \mathrm{VDH}(k), \mathrm{FCM}(k)$ are believed to have at most a distance of 1 to each other. On the other hand we conjecture $\mathrm{FCH}(k) = S_2$, where $S_2(k)$ oscillates between the linear function $k+1$ and the quasi-linear function $k + 1 + \lfloor \log_2(k) \rfloor$. Altogether the "four fundamental quantities" $\mathrm{FCH}, \mathrm{FCM}, \mathrm{VDH}, \mathrm{VDM}$ are fascinating and important structural parameters, whose study continues to reveal new and surprising aspects of $\mathcal{MU}$ and $\mathcal{UHIT}$.

It is also possible to go beyond $\mathcal{MU}$: in [23, Section 9] it is shown that when considering the maximum of $\mu\mathrm{vd}(F)$ over all $F \in \mathcal{LEAN}_{\delta=k} \supset \mathcal{MU}_{\delta=k}$, the set of all "lean" clause-sets, that then $\mathrm{nM}(k)$ is the precise maximum for all $k \geq 1$. Lean clause-sets were introduced in [17] as the clause-sets where it is not possible to satisfy some clauses while not touching the other clauses, and indeed were already introduced earlier, as "non-weakly satisfiable formulas (matrices)" in the field of QMA by [12]. Furthermore it is shown in [23, Section 10], that there is a polytime "autarky reduction", removing some clauses which can be satisfied without touching the other clauses, which establishes for arbitrary clause-sets F the upper bound $\mu\mathrm{vd}(F) \leq \mathrm{nM}(\delta(F))$; an interesting open question here is to find the witnessing autarky in polynomial time.

Back to the main result of this paper, the proof of $S_2 \leq \mathrm{FCH}$ is non-trivial. Indeed the proof is relatively easy for a function $S_2'(k)$ defined by an appropriate recursion, motivated by employing full subsumption extension $C \rightsquigarrow C \cup \{v\}, C \cup \{\overline{v}\}$ in an optimal way. Then the main auxiliary result is $S_2' = S_2$. For that we use another function, namely $a_2(k)$ as considered in [24] in a more general form, while a_2 was introduced with a small modification in [3]. These considerations belong to the field of *meta-Fibonacci sequences*, where special nested recursions are studied, initiated by [8, Page 145]. Via a combinatorial argument we derive such a nested recursion from the course-of-value recursion for S_2', which yields $S_2' = 2a_2$. We also show $2a_2 = S_2$ (this equality was conjectured on the OEIS [27]), and we obtain $S_2' = S_2$.

Overview. The main results of this paper are as follows. Theorem 18 proves $S_2 = 2a_2$. Theorem 31 shows a meta-Fibonacci recursion for S_2', where S_2' is introduced by a recursion directly related to our application. Theorem 33 then proves $S_2' = S_2$. After these number-theoretic preparations, we consider subsumption resolution and its inversion (extension); Theorem 38 combines subsumption extension and the recursion machinery, and shows $S_2 \leq \mathrm{FCH}$. In the remainder of the paper, this fundamental result is applied. Theorem 39 proves a tight upper bound on S_2, while Theorem 42 considers the cases where the lower bound via S_2 and the upper bound via nM coincides. Finally in Theorem 47 we determine the four fundamental quantities for $1 \leq k \leq 13$ (see Table 1).

2 Preliminaries

We use $\mathbb{Z}$ for the set of integers, $\mathbb{N}_0 := \{n \in \mathbb{Z} : n \geq 0\}$, $\mathbb{N} := \mathbb{N}_0 \setminus \{0\}$, and finally $\mathbb{R} \supset \mathbb{Z}$ is the set of real numbers. For maps $f, g : X \to \mathbb{R}$ we write $f \leq g$ if $\forall\, x \in X : f(x) \leq g(x)$.

On the set $\mathcal{LIT}$ of literals we have complementation $x \in \mathcal{LIT} \mapsto \overline{x}$, with $\overline{x} \neq x$ and $\overline{\overline{x}} = x$. We assume $\mathbb{Z} \setminus \{0\} \subseteq \mathcal{LIT}$, with $\overline{z} = -z$ for $z \in \mathbb{Z} \setminus \{0\}$. Variables $\mathcal{VA} \subset \mathcal{LIT}$ with $\mathbb{N} \subseteq \mathcal{VA}$ are special literals, and the underlying variable of a literal is given by $\mathrm{var} : \mathcal{LIT} \to \mathcal{VA}$, such that for $v \in \mathcal{VA}$ holds $\mathrm{var}(v) = \mathrm{var}(\overline{v}) = v$, while for $x \in \mathcal{LIT} \setminus \mathcal{VA}$ holds $\overline{x} = \mathrm{var}(x)$. For a set $L \subseteq \mathcal{LIT}$ we define $\overline{L} := \{\overline{x} : x \in L\}$. A clause is a finite set C of literals with

$C \cap \overline{C} = \emptyset$ (C is clash-free). A clause-set is a finite set of clauses, the set of all clause-sets is denoted by $\mathcal{CLS}$.

For a clause C we define $\operatorname{var}(C) := \{\operatorname{var}(x) : x \in C\} \subset \mathcal{VA}$, and for a clause-set F we define $\operatorname{var}(F) := \bigcup_{C \in F} \operatorname{var}(C) \subset \mathcal{VA}$. We use the measure $n(F) := |\operatorname{var}(F)| \in \mathbb{N}_0$ and $c(F) := |F| \in \mathbb{N}_0$, while the deficiency is $\delta(F) := c(F) - n(F) \in \mathbb{Z}$.

The set of satisfiable clause-sets is denoted by $\mathcal{SAT} \subset \mathcal{CLS}$, which is the set of clause-sets F such that there is a clause C which intersects all clauses of F, i.e., with $\forall\, D \in F : C \cap D \neq \emptyset$; the unsatisfiable clause-sets are $\mathcal{USAT} := \mathcal{CLS} \setminus \mathcal{SAT}$.

The set $\mathcal{MU} \subset \mathcal{USAT}$ of minimally unsatisfiable clause-sets is the set of $F \in \mathcal{USAT}$, such that for $F' \subset F$ holds $F' \in \mathcal{SAT}$. The unsatisfiable hitting clause-sets are given by $\mathcal{UHIT} := \{F \in \mathcal{USAT} \mid \forall\, C, D \in F, C \neq D : C \cap \overline{D} \neq \emptyset\}$. It is easy to see that $\mathcal{UHIT} \subset \mathcal{MU}$ holds, and that for all $F \in \mathcal{UHIT}$ holds $\sum_{C \in F} 2^{-|C|} = 1$. While all definitions are given in this paper, for some more background see [15].

2.1 Full clauses

A **full clause** for $F \in \mathcal{CLS}$ is some $C \in F$ with $\operatorname{var}(C) = \operatorname{var}(F)$ (equivalently, $|C| = n(F)$), and the number of full clauses is counted by $\operatorname{fc} : \mathcal{CLS} \to \mathbb{N}_0$, which can be defined as $\mathbf{fc}(\boldsymbol{F}) := c(F \cap A(\operatorname{var}(F)))$, and where $\boldsymbol{A(V)} \in \mathcal{UHIT}$ for some finite $V \subset \mathcal{VA}$ is the set of all clauses C with $\operatorname{var}(C) = V$. Standardised versions of the $A(V)$ are $\boldsymbol{A_n} := A(\{1, \dots, n\})$ for $n \in \mathbb{N}_0$. The following observation is contained in the proof of [31, Utterly Trivial Observation]:

Lemma 1. *For $F \in \mathcal{UHIT}$, $F \neq \{\bot\}$, the number $\operatorname{fc}(F)$ of full clauses is even.*

Proof. Let $n := n(F)$. We have $\sum_{C \in F} 2^{n-|C|} = 2^n$, and thus $\sum_{C \in F} 2^{n-|C|}$ is even (due to $n > 0$). Since $\sum_{C \in F, |C| \neq n} 2^{n-|C|}$ is even, the assertion follows. □

2.2 The four fundamental quantities

For $F \in \mathcal{CLS}$ we define the var-degree as $\operatorname{vd}_F(v) := c(\{C \in F : v \in \operatorname{var}(C)\}) \in \mathbb{N}_0$ for $v \in \mathcal{VA}$, while in case of $\operatorname{var}(F) \neq \emptyset$ (i.e., $F \notin \{\top, \{\bot\}\}$) we define the min-var-degree $\boldsymbol{\mu}\mathbf{vd}(\boldsymbol{F}) := \min_{v \in \operatorname{var}(F)} \operatorname{vd}_F(v) \in \mathbb{N}$.

Definition 2. *For $k \in \mathbb{N}$ let*

- $\mathbf{FCH}(\boldsymbol{k}) \in \mathbb{N}$ *be the maximal* $\operatorname{fc}(F)$ *for* $F \in \mathcal{UHIT}_{\delta=k}$*;*
- $\mathbf{FCM}(\boldsymbol{k}) \in \mathbb{N}$ *be the maximal* $\operatorname{fc}(F)$ *for* $F \in \mathcal{MU}_{\delta=k}$*;*
- $\mathbf{VDH}(\boldsymbol{k}) \in \mathbb{N}$ *be the maximal* $\mu\operatorname{vd}(F)$ *for* $F \in \mathcal{UHIT}_{\delta=k}$*;*
- $\mathbf{VDM}(\boldsymbol{k}) \in \mathbb{N}$ *be the maximal* $\mu\operatorname{vd}(F)$ *for* $F \in \mathcal{MU}_{\delta=k}$*.*

For $k = 1$ the case $F = \{\bot\}$ is excluded in the last two definitions.

By [21, Lemma 9, Corollary 10, Theorem 15]:

Theorem 3 ([21]). $\mathrm{VDM}(k) \le \mathrm{nM}(k) = k + \lfloor \log_2(k + 1 + \lfloor \log_2(k+1) \rfloor) \rfloor \le k + 1 + \lfloor \log_2(k) \rfloor$ *for all* $k \in \mathbb{N}$.

Here $\mathrm{nM} : \mathbb{N} \to \mathbb{N}$ is the enumeration of natural numbers excluding the Mersenne numbers $2^n - 1$ for $n \in \mathbb{N}$; the list of initial values is $2, 4, 5, 6, 8, 9, 10, 11, 12, 13, 14, 16, 17$ (`http://oeis.org/A062289`). In [23, Theorem 14.4] it is shown that $\mathrm{VDM}(6) = 8 = \mathrm{nM}(6) - 1$, extending this to an improved upper bound $\mathrm{VDM} \le \mathrm{nM}_1$ ([23, Theorem 14.6], where $\mathrm{nM}_1 : \mathbb{N} \to \mathbb{N}$ can be defined as follows: $\mathrm{nM}_1(k) := \mathrm{nM}(k)$ for $k \in \mathbb{N}$ with $k \neq 2^n - n + 1$ for some $n \ge 3$, while $\mathrm{nM}_1(2^n - n + 1) := \mathrm{nM}(2^n - n + 1) - 1 = 2^n$; see Table 1 for initial values.

Theorem 4 ([23]). *For* $k \in \mathbb{N}$ *holds* $\mathrm{VDM}(k) \le \mathrm{nM}_1(k) \le \mathrm{nM}(k)$.

We conclude these preparations with a special property of $\mathrm{FCH}(k)$ (supporting our Conjecture 48 that $\mathrm{FCH} = S_2$), namely by Lemma 1 we have:

Corollary 5. $\mathrm{FCH}(k)$ *is even for all* $k \in \mathbb{N}$.

3 Some integer sequences

We review the "Smarandache primitive numbers" $S_2(k)$ and the meta-Fibonacci sequences $a_2(k)$. We show in Theorem 18, that $S_2 = 2a_2$ holds.

3.1 Some preparations

We define two general operations $a \mapsto \Delta a$ and $a \mapsto \mathfrak{P}\, a$ for sequences a. First the (standard) Δ-operator:

Definition 6. *For* $a : I \to \mathbb{R}$*, where* $I \subseteq \mathbb{Z}$ *is closed under increment, we define* $\boldsymbol{\Delta a} : I \to \mathbb{R}$ *by* $\Delta a(k) := a(k+1) - a(k)$.

So a is monotonically increasing iff $\Delta a \ge 0$, while a is strictly monotonically increasing iff $\Delta a \ge 1$. Sequences with exactly two different Δ-values, where one of these values is 0, play a special role for us, and we call them "d-Delta", where d is the other value:

Definition 7. *A sequence* $a : \mathbb{N}_0 \to \mathbb{Z}$ *is called* d***-Delta*** *for* $d \in \mathbb{Z} \setminus \{0\}$*, if* $\Delta a(\mathbb{N}_0) = \{\Delta a(n)\}_{n \in \mathbb{N}_0} = \{0, d\}$.

While the Δ-operator determines the change to the next value, the *plateau-operator* determines subsequences of unchanging values:

Definition 8. *For a sequence* $a : \mathbb{N} \to \mathbb{R}$ *which is non-stationary (for all* i *there is* $j > i$ *with* $a_j \neq a_i$*) we define* $\mathfrak{P}\, a : \mathbb{N} \to \mathbb{N}$ *(the "plateau operator") by letting* $\mathfrak{P}\, a(n)$ *for* $n \in \mathbb{N}$ *be the size of the* n*-th (maximal) plateau of equal values (maximal intervals of* $\mathbb{N}$ *where* a *is constant).*

So $\mathfrak{P}\, a(1)$ is the size of the first plateau, $\mathfrak{P}\, a(2)$ the size of the second plateau, and so on; $\forall\, i \in \mathbb{N} : a(i) \neq a(i+1)$ iff $\mathfrak{P}\, a$ is the constant 1-function. For a d-Delta sequence a from $\mathfrak{P}\, a$ and the initial value a_1 we can reconstruct a.

3.2 Smarandache primitive numbers

The “Smarandache Primitive Numbers” were introduced in [28, Unsolved Problem 47]:

Definition 9. *For $k \in \mathbb{N}_0$ let $S_2(k)$ be the smallest $n \in \mathbb{N}_0$ such that 2^k divides $n!$. Using $\operatorname{ord}_2(n)$, $n \in \mathbb{N}$, for the maximal $m \in \mathbb{N}_0$ such that 2^m divides n, we get that $S_2(k)$ for $k \in \mathbb{N}_0$ is the smallest $n \in \mathbb{N}_0$ such that $k \leq \sum_{i=1}^{n} \operatorname{ord}_2(i)$.*

So $S_2(0) = 0$, and $\Delta S_2(\mathbb{N}_0) = \{0, 2\}$. The following is well-known and easy to show (see Subsection III.1 in [9] for basic properties of $S_2(k)$):

Lemma 10. *The sequence $S_2(1), S_2(2), S_2(3), \ldots$ is obtained from the sequence $1, 2, 3, \ldots$ of natural numbers, when each element $n \in \mathbb{N}$ is repeated $\operatorname{ord}_2(n)$ many times.*

Example 11. The numbers $S_2(k)$ for $k \in \{1, \ldots, 25\}$ are $2, 4, 4, 6, 8, 8, 8, 10, 12, 12, 14, 16, 16, 16, 16, 18, 20, 20, 22, 24, 24, 24, 26, 28, 28$. The corresponding OEIS-entry (which has 1 as first element (index 0), instead of 0 as we have it, and which we regard as appropriate) is `http://oeis.org/A007843`.

Lemma 12 ([30]). *For $k \in \mathbb{N}$ holds $k + 1 \leq S_2(k) = k + O(\log k)$.*

We give an independent proof for the lower bound in Lemma 40, while we sharpen the upper bound in Theorem 39. For more number-theoretic properties of S_2 see [29]. To understand the plateaus of S_2, we need the *ruler function*:

Definition 13. *Let $\mathrm{ru}_n := \operatorname{ord}_2(2n) \in \mathbb{N}$ for $n \in \mathbb{N}$.*

The plateaus of S_2 are given by the ruler function: in Lemma 10 we determined the number of repetitions of values $v \in \mathbb{N}$ as $\operatorname{ord}_2(v)$, while for the plateaus we skip zero-repetitions, which happen at each odd number, and thus for the associated index n we have $n = \frac{v}{2}$ for even v, and the number of repetitions is $\operatorname{ord}_2(v) = \operatorname{ord}_2(2n)$; we obtain

Lemma 14. $\mathfrak{P}(S_2(k))_{k \in \mathbb{N}} = (\mathrm{ru}_n)_{n \in \mathbb{N}}$.

3.3 Meta-Fibonacci sequences

Started by [8, Page 145], various nested recursions for integer sequences have been studied. Often the focus in this field of “meta-Fibonacci sequences” is on “chaotic behaviour”, but we consider here only a well-behaved case (but in detail):

Definition 15. *In [24] the sequence $a_2 : \mathbb{N}_0 \to \mathbb{N}_0$[3], has been defined recursively via*

$$a_2(k) = a_2(k - a_2(k-1)) + a_2(k - 1 - a_2(k-2)),$$

while $a_2(k) := k$ for $k \in \{0, 1\}$.

[3] hiding two parameters $d \in \mathbb{N}$, $s \in \mathbb{Z}$ used in [24], which are $d = 2$, $s = 0$ in our case

The sequence a_2 was introduced in [3] as $F : \mathbb{N} \to \mathbb{N}_0$, with $F(k) = k-1$ for $k \in \{1, 2\}$ and the same recursion law, which yields $F(k) = a_2(k-1)$ for $k \in \mathbb{N}$. Furthermore, using $F'(1) = F'(2) = 1$ as initial conditions does not change anything else, and this sequence is the OEIS entry `http://oeis.org/A046699`. It is shown (in our notation):

Lemma 16 ([3]). *For $k \in \mathbb{N}$ and $p := \lfloor \log_2(k+1) \rfloor$: $a_2(k) = 2^{p-1} + a_2(k+1-2^p)$.*

Lemma 16 yields a fast computation of $a_2(k)$. [10, Corollary 2.9, Equation (1)] determines the plateau sizes:

Lemma 17 ([10]). *a_2 is a 1-Delta sequence with $\mathfrak{P}(a_2(k))_{k \in \mathbb{N}} = \mathrm{ru}$.*

We can now show $a_2 = \frac{1}{2} S_2$, which has been conjectured on the OEIS (`http://oeis.org/A007843`, by Michel Marcus):

Theorem 18. *$\forall\, k \in \mathbb{N}_0 : S_2(k) = 2 \cdot a_2(k)$.*

Proof. By Lemma 14 and Lemma 17, together with $S_2(0) = a_2(0) = 0$. □

4 Recursions for Smarandache primitive numbers

In Subsection 4.1 we introduce the sequence S_2' via a recursive process, which directly ties into our main application in Theorem 38 for constructing unsatisfiable hitting clause-sets with many full clauses. This recursive definition uses an index, which is studied in Subsection 4.2. The central helper function is the "slack", studied in Subsection 4.3. We then prove a meta-Fibonacci recursion in Theorem 31, and obtain $S_2' = S_2$ in Theorem 33.

4.1 A simple course-of-values recursion

Definition 19. *For $k \in \mathbb{N}_0$ let*

1. *$S_2'(0) := 0$, $S_2'(1) := 2$; and for $k \geq 2$:*
2. *$S_2'(k) := 2 \cdot (k-i+1)$ for the minimal $i \in \{1, \dots, k-1\}$ with $k-i+1 \leq S_2'(i)$.*

Note that the recursion step is well-defined (the i exists), since for $i = k-1$ holds $k-i+1 = 2$, and $S_2'(k-1) = 2$ for $k = 2$, while for $k \geq 3$ holds $S_2'(k-1) = 2 \cdot ((k-1) - i' + 1) \geq 2 \cdot ((k-1) - ((k-1) - 1) + 1) = 4$. The condition "$k - i + 1 \leq S_2'(i)$" is equivalent to $k + 1 \leq i + S_2'(i)$. Some simple properties are that $S_2'(k)$ is divisible by 2, $S_2'(k) \geq 2$ for $k \geq 1$, and $S_2'(2) = 4$ and $S_2'(k) \geq 4$ for $k \geq 2$.

4.2 Analysing the index

Definition 20. *For $k \geq 0$ let* $\mathbf{i_S(}\boldsymbol{k}\mathbf{)} := k + 1 - \frac{S_2'(k)}{2} \in \mathbb{N}$.

Simple properties (for all $k \geq 0$):

1. $S_2'(k) = 2 \cdot (k - \mathrm{i_S}(k) + 1)$.
2. $\mathrm{i_S}(0) = \mathrm{i_S}(1) = \mathrm{i_S}(2) = 1$.
3. $\Delta\, \mathrm{i_S}(k) = 0 \Leftrightarrow \Delta S_2'(k) = 2$ and $\Delta\, \mathrm{i_S}(k) = 1 \Leftrightarrow \Delta S_2'(k) = 0$.

An alternative characterisation of $\mathrm{i_S}(k)$:

Lemma 21. *For $k \geq 0$: $\mathrm{i_S}(k)$ is the minimal $i \in \mathbb{N}_0$ with $i + S_2'(i) \geq k + 1$.*

Proof. The assertion follows by what has already been said above, plus the consideration of the corner cases: $0 + S_2'(0) = 0 < k + 1$ for all $k \geq 0$, while $1 + S_2'(1) = 3 \geq k + 1$ for $k \leq 2$. □

We obtain a method to prove lower bounds for $S_2'(k)$:

Corollary 22. *For $k, i \in \mathbb{N}_0$ with $S_2'(i) \geq k - i + 1$ holds $S_2'(k) \geq 2(k - i + 1)$.*

$\mathrm{i_S}(k)$ grows in steps of $+1$, while $S_2'(k)$ grows in steps of $+2$:

Lemma 23. *$\Delta S_2'(k) \in \{0, 2\}$ and $\Delta\, \mathrm{i_S}(k) \in \{0, 1\}$ for all $k \in \mathbb{N}_0$.*

Proof. Proof via (simultaneous) induction on k: The assertions hold for $k \leq 1$, and so consider $k \geq 2$. Now $\mathrm{i_S}(k)$ is the minimal $i \in \{1, \ldots, k-1\}$ with $k + 1 \leq i + S_2'(i)$, and due to $\Delta S_2'(i) \geq 0$ for all $i < k$ it follows $\Delta\, \mathrm{i_S}(k) \in \{0, 1\}$. □

We obtain a simple upper bound on $\mathrm{i_S}$:

Corollary 24. *For $k \geq 1$ holds $\mathrm{i_S}(k) \leq k$ and for $k \geq 2$ holds $\mathrm{i_S}(k) \leq k - 1$*

4.3 The "slack"

An important helper function is the "slack" $\mathrm{sl_S}(k)$:

Definition 25. *For $k \in \mathbb{N}_0$ let* $\mathbf{sl_S(}\boldsymbol{k}\mathbf{)} := (\mathrm{i_S}(k) + S_2'(\mathrm{i_S}(k))) - (k + 1) \in \mathbb{N}_0$.

So $\mathrm{sl_S}(0) = (1 + 2) - (0 + 1) = 2$ and $\mathrm{sl_S}(1) = (1 + 2) - (1 + 1) = 1$. Directly from the definition follows:

Lemma 26. *For $k \geq 0$ holds $S_2'(\mathrm{i_S}(k)) = \frac{1}{2} S_2'(k) + \mathrm{sl_S}(k)$.*

We can characterise the cases $\Delta\, \mathrm{i_S}(k) = 1$ as the "slackless" k's:

Lemma 27. *For $k \geq 0$:*

1. $\Delta\, \mathrm{i_S}(k) = 1 \Leftrightarrow \mathrm{sl_S}(k) = 0 \Leftrightarrow \Delta S_2'(k) = 0$.
2. $\Delta\, \mathrm{i_S}(k) = 0 \Leftrightarrow \mathrm{sl_S}(k) \geq 1 \Leftrightarrow \Delta S_2'(k) = 2$.

Proof. If $\mathrm{sl}_{\mathrm{S}}(k) \geq 1$, then $\Delta\, \mathrm{i}_{\mathrm{S}}(k) = 0$ by Lemma 21, while for $\mathrm{sl}_{\mathrm{S}}(k) = 0$ we get $\Delta\, \mathrm{i}_{\mathrm{S}}(k) \geq 1$. □

Thus the slack determines the growth of S'_2:

Corollary 28. *For $k \geq 0$ holds $\Delta S'_2(k) = 2 \cdot \min(\mathrm{sl}_{\mathrm{S}}(k), 1)$.*

And plateaus of the slack happen only for slack zero, and from such a plateau the slack jumps to 2, and then is stepwise again decremented to zero:

Corollary 29. *For $k \geq 0$ holds:*

1. *If $\mathrm{sl}_{\mathrm{S}}(k) > 0$, then $\mathrm{sl}_{\mathrm{S}}(k+1) = \mathrm{sl}_{\mathrm{S}}(k) - 1$.*
2. *If $\mathrm{sl}_{\mathrm{S}}(k) = 0$, then $\mathrm{sl}_{\mathrm{S}}(k+1) \in \{0, 2\}$.*

4.4 A meta-Fibonacci recursion

We are ready to prove an interesting nested recursion for S'_2. First a combinatorial lemma, just exploiting the fact that the shape of the slack repeats the following pattern (Corollary 29): a plateau of zeros, followed by a jump to 2 and a stepwise decrement to 0 again (where right at $k = 0$ we start with $\mathrm{sl}_{\mathrm{S}}(0) = 2$):

Lemma 30. *For $k \geq 2$ holds $\sum_{i=1}^{2} \mathrm{sl}_{\mathrm{S}}(k-i) = \sum_{i=1}^{2} i \cdot \min(1, \mathrm{sl}_{\mathrm{S}}(k-i))$.*

Proof. There are $0 \leq p \leq 2$ and $1 \leq q \leq 3$ such that the left-hand side is

$$p + (p-1) + \cdots + 1 + 0 + \cdots + 0 + 2 + (2-1) + \cdots + q;$$

for $p = 0$ the initial part is empty, for $q = 3$ the final part is empty. Let $r \geq 0$ be the number of zeros; so $r = 0$ iff $p = 2$ (and then also $q = 3$). We have $p + r + (2 - q + 1) = 2$, i.e., $p + r + 1 = q$. Now the right-hand side is

$$1 + 2 + \cdots + p + 0 + \cdots + 0 + q + (q+1) + \cdots + 2,$$

and we see that both sides are equal. □

Theorem 31. *For $k \geq 2$ holds*

$$S'_2(k) = \sum_{i=1}^{2} S'_2(\mathrm{i}_{\mathrm{S}}(k-i))$$

(note that by Lemma 24 holds $\mathrm{i}_{\mathrm{S}}(k-i) < k$).

Proof. By Lemma 26 and Lemma 30 holds

$$\sum_{i=1}^{2} S'_2(\mathrm{i}_{\mathrm{S}}(k-i)) = (\sum_{i=1}^{2} \mathrm{sl}_{\mathrm{S}}(k-i)) + S'_2(k) - \frac{1}{2}\sum_{i=1}^{2}(S'_2(k) - S'_2(k-i)) =$$
$$S'_2(k) + (\sum_{i=1}^{2} i \cdot \min(1, \mathrm{sl}_{\mathrm{S}}(k-i))) - \frac{1}{2}\sum_{i=1}^{2}\sum_{j=0}^{i-1} \Delta S'_2(k+i-j),$$

where now by Corollary 28 holds $\sum_{i=1}^{2} \sum_{j=0}^{i-1} \Delta S'_2(k+i-j) = (\Delta S'_2(k-1)) + (\Delta S'_2(k-2) + \Delta S'_2(k-1)) = \sum_{i=1}^{2} i \cdot \Delta S'_2(k-1) = 2 \sum_{i=1}^{2} i \cdot \min(1, \mathrm{sl_S}(k))$, which completes the proof. $\square$

Now we see that S'_2 is basically the same as a_2 (recall Subsection 3.3):

Corollary 32. $\forall\, k \in \mathbb{N}_0 : S'_2(k) = 2 \cdot a_2(k)$.

Proof. For the purpose of the proof let $a_2(k) := \frac{1}{2} S'_2(k)$ for $k \in \mathbb{N}_0$. So we get $a_2(k) = k$ for $k \in \{0, 1\}$, while $\mathrm{i_S}(k) = k + 1 - a_2(k)$, and thus for $k \geq 2$:

$$a_2(k) = \frac{1}{2} S'_2(k) = \frac{1}{2} \sum_{i=1}^{2} S'_2(\mathrm{i_S}(k-i)) = \sum_{i=1}^{2} a_2(\mathrm{i_S}(k-i)) = \sum_{i=1}^{2} a_2(k-i+1-a_2(k-i)),$$

and so the assertion follows by the equations of Definition 15. $\square$

We obtain the main result of this section:

Theorem 33. *$S'_2 = S_2$ (recall Definition 9).*

Proof. By Corollary 32 and Theorem 18. $\square$

5 On the number of full clauses

First we review full subsumption resolution, $C \cup \{v\}, C \cup \{\overline{v}\} \rightsquigarrow C$, and its inversion, called "extension" in Section 5.1, where some care is needed, since we need complete control. From a clause-set F with "many" full clauses we can produce further clause-sets with "many" full clauses by full subsumption extension done in parallel, and this process of "full expansion" is presented in Definition 36. The recursive computation of S_2 via Definition 19 captures maximisation for this process, and so we can show in Theorem 38, that we can construct examples of unsatisfiable hitting clause-sets F_k of deficiency k and with $S_2(k)$ many full clauses. It follows that S_2 yields a lower bound on FCH (Conjecture 48 says this lower bound is actually an equality).

5.1 Full subsumption resolution

As studied in [23, Section 6] in some detail:

Definition 34 ([23]). *A **full subsumption resolution** for $F \in \mathcal{CLS}$ can be performed, if there is a clause $C \notin F$ with $C \cup \{v\}, C \cup \{\overline{v}\} \in F$ for some variable v, and replaces the two clauses $C \cup \{v\}, C \cup \{\overline{v}\}$ by the single clause C. For the **strict** form, there must exist a third clause $D \in F \setminus \{C \cup \{v\}, C \cup \{\overline{v}\}\}$ with $v \in \mathrm{var}(D)$, while for the **non-strict** form there must NOT exist such a third clause.*

If F' is obtained from F by one full subsumption resolution, then $c(F') = c(F) - 1$; we have the strict form iff $n(F') = n(F)$, or, equivalently, $\delta(F') = \delta(F) - 1$, while we have the non-strict form iff $n(F') = n(F) - 1$, or, equivalently, $\delta(F') = \delta(F)$. A very old transformation of a CNF (DNF) into an equivalent one uses the inverse of full subsumption resolution[4)]:

Definition 35 ([23]). *A **full subsumption extension** for $F \in \mathcal{CLS}$ and a clause $C \in F$ can be performed, if there is a variable $v \in \mathcal{VA} \setminus \operatorname{var}(C)$ with $C \cup \{v\}, C \cup \{\overline{v}\} \notin F$, and replaces the single clause C by the two clauses $C \cup \{v\}, C \cup \{\overline{v}\}$. For the **strict** form we have $v \in \operatorname{var}(F)$, while for the **non-strict** form we have $v \notin \operatorname{var}(F)$.*

If we consider $F \in \mathcal{MU}$ and $C \in F$, then we can always perform a non-strict full subsumption extension, while we can perform the strict form iff C is not full. If we denote the result by F', then for $F \in \mathcal{UHIT}$ we have again $F' \in \mathcal{UHIT}$, but for general $F \in \mathcal{MU}$ we might have $F' \notin \mathcal{MU}$; see [23, Lemma 6.5] for an exact characterisation.

5.2 Full expansions

We now perform full subsumption extensions in parallel to m full clauses of F, first using a non-strict extension, and then reusing the extension variable via strict extensions:

Definition 36. *For $F \in \mathcal{CLS}$ and $m \in \mathbb{N}$, where $\operatorname{fc}(F) \geq m$, a **full m-expansion of** F is some $G \in \mathcal{CLS}$ obtained by*

1. *choosing some $F' \subseteq F \cap A(\operatorname{var}(F))$ with $c(F') = m$,*
2. *choosing some $v \in \mathcal{VA} \setminus \operatorname{var}(F)$ (the **extension variable**),*
3. *and replacing the clauses $C \in F'$ in F by their full subsumption extension with v (recall Definition 35).*

The choice of v in Definition 36 is irrelevant, while the choice of F' might have an influence on further properties of G, but is irrelevant for our uses. The following basic properties all follow directly from the definition:

Lemma 37. *Consider the situation of Definition 36.*

1. *There is always a full m-expansion G (unique for any fixed F', v).*
2. *If $F \in \mathcal{UHIT}$, then $G \in \mathcal{UHIT}$.*
3. *$n(G) = n(F) + 1$, $c(G) = c(F) + m$.*
4. *$\delta(G) = \delta(F) + m - 1$.*
5. *$\operatorname{fc}(G) = 2 \cdot m$.*

We turn to the construction of unsatisfiable hitting clause-sets with many full clauses (for a given deficiency):

---

4) Boole introduced in [2], Chapter 5, Proposition II, the general "expansion" $f(v, \boldsymbol{x}) = (f(0, \boldsymbol{x}) \wedge \overline{v}) \vee (f(1, \boldsymbol{x}) \wedge v)$ for boolean functions f, where for our application $f(v, \boldsymbol{x}) \approx C$. This was taken up by [26], and is often referred to as "Shannon expansion".

Theorem 38. *For $k \in \mathbb{N}$ we recursively construct $F_k \in \mathcal{UHIT}_{\delta=k}$ as follows:*

1. $F_1 := \{\{1\}, \{-1\}\}$.
2. *For $k \geq 2$ let F_k be a full $a_2(k)$-expansion of $F_{\mathrm{i_S}(k)}$.*

Then we have $\mathrm{fc}(F_k) = S_2(k)$. Thus $\forall\, k \in \mathbb{N} : S_2(k) \leq \mathrm{FCH}(k)$.

Proof. If the construction is well-defined, then we get $\mathrm{fc}(F_k) = 2 \cdot a_2(k) = S_2(k)$ and $\delta(F_k) = \delta(F_{\mathrm{i_S}(k)}) + a_2(k) - 1 = \mathrm{i_S}(k) + a_2(k) - 1 = k$ for $k \geq 2$ by Lemma 37 (using Theorem 33 freely), while these two properties hold trivially for $k = 1$.

It remains to show that $1 \leq \mathrm{i_S}(k) \leq k - 1$ and $a_2(k) \leq \mathrm{fc}(F_{\mathrm{i_S}(k)})$ for $k \geq 2$. The first statement follows by Corollary 24, while the second statement follows by Lemma 21. □

6 Applications

We start by sharpening the upper bound from Lemma 12:

Theorem 39. *For $k \in \mathbb{N}$ holds $S_2(k) \leq \mathrm{nM}(k) \leq k + 1 + \lfloor \log_2(k) \rfloor$.*

Proof. By Theorem 38 and Theorem 3. □

We can also provide an independent proof of the lower bound of Lemma 12:

Lemma 40. *For $k \in \mathbb{N}$ holds $S_2(k) \geq k + 1$.*

Proof. We prove the assertion by induction. For $k = 1$ we have $S_2(1) = 2$, so consider $k \geq 2$. We use Corollary 22, and so we need $i \in \mathbb{N}$ with $k+1 \leq 2(k-i+1)$, i.e., $i \leq \frac{k+1}{2}$. So we choose $i := \lfloor \frac{k+1}{2} \rfloor \in \mathbb{N}$. We have $i < k$, and so we can apply the induction hypothesis to i: $i + S_2(i) = \lfloor \frac{k+1}{2} \rfloor + S_2(\lfloor \frac{k+1}{2} \rfloor) \geq \lfloor \frac{k+1}{2} \rfloor + \lfloor \frac{k+1}{2} \rfloor + 1 = 2\lfloor \frac{k+1}{2} \rfloor + 1 > 2(\frac{k+1}{2} - 1) + 1 = k$, and thus $i + S_2(i) \geq k + 1$. □

When upper and lower bound coincide, then we know all four fundamental quantities; first we name the sets of deficiencies (recall Theorems 3, 4):

Definition 41. $\boldsymbol{\mathcal{SNM}} := \{k \in \mathbb{N} : S_2(k) = \mathrm{nM}(k)\}$, $\boldsymbol{\mathcal{SNM}_1} := \{k \in \mathbb{N} : S_2(k) = \mathrm{nM}_1(k)\}$.

By $S_2 \leq \mathrm{VDM} \leq \mathrm{nM}_1 \leq \mathrm{nM}$ we get $\mathcal{SNM} \subseteq \mathcal{SNM}_1$ and:

Theorem 42. *For $k \in \mathcal{SNM}_1$ holds $S_2(k) = \mathrm{FCH}(k) = \mathrm{FCM}(k) = \mathrm{VDH}(k) = \mathrm{VDM}(k) = \mathrm{nM}_1(k)$.*

We prove now that the special deficiencies $2^n - n, 2^n - n - 1$ ($n \geq 1$; note $\delta(A_n) = 2^n - n$) considered in [23, Lemmas 12.10, 12.11], where we have shown that for them the four fundamental quantities coincide, are indeed in $\mathcal{SNM}$, and that furthermore the special deficiencies $2^n - n + 1$ ($n \geq 3$), where nM_1 differs from nM, are in $\mathcal{SNM}_1$:

Lemma 43. *Consider $n \in \mathbb{N}$.*

1. *$S_2(2^n - n) = 2^n$, and for $k \in \mathbb{N}_0$ holds $S_2(k) = 2^n \Leftrightarrow 2^n - n \le k \le 2^n - 1$.*
2. *$2^n - n \in \mathcal{SNM}$, while $2^n - n + 1, \dots, 2^n - 1 \notin \mathcal{SNM}$.*
3. *Assume $n \ge 2$ now. Then $2^n - n - 1 \in \mathcal{SNM}$ with $S_2(2^n - n - 1) = 2^n - 2$.*
4. *For $n \ge 3$ holds $2^n - n + 1 \in \mathcal{SNM}_1$.*

Proof. By [23, Corollary 7.24] we have $\mathrm{nM}(2^n - n) = 2^n$, while $\mathrm{nM}(2^n - n - 1) = 2^n - 2$ (remember that the jumps for nM happens at the deficiencies $2^n - n$). Thus $S_2(2^n - n) \le 2^n$ and $S_2(2^n - n - 1) \le 2^n - 2$. Since for the value 2^n the sequence S_2 has a plateau of length n (Lemma 10), while nM is strictly increasing, for Parts 1, 2, 3 it remains to show $S_2(2^n - n) \ge 2^n$. We show this by induction: For $n = 1$ we have $S_2(1) = 2 = 2^1$, while for $n \ge 2$ by induction hypothesis we have $(2^n - n) - (2^{n-1} - (n-1)) + 1 = 2^{n-1} \le S_2(2^{n-1} - (n-1))$, thus by Corollary 22 $S_2(2^n - n) \ge 2 \cdot 2^{n-1} = 2^n$. Finally, for Part 4 we note $S_2(2^n - n + 1) = S_2(n) = 2^n$ by Part 1, while $\mathrm{nM}_1(k)$ differs from $\mathrm{nM}(k)$ exactly at the positions $k = 2^n - n + 1$ for $n \ge 3$, where then $\mathrm{nM}_1(k) = \mathrm{nM}(k) - 1 = 2^n$ ([23, Theorem 14.7]). □

So the lower bound of Lemma 40 is sharp for infinitely many deficiencies:

Corollary 44. *We have $S_2(k) = k + 1$ for all $k = 2^n - 1$, $n \in \mathbb{N}$.*

7 Initial values of the four fundamental quantities

The task of this penultimate section is to prove the values in Table 1 (in Theorem 47; of course, only the four fundamental quantities are open).

k	1	2	**3**	4	5	6	**7**	**8**	9	**10**	11	12	13
$\mathrm{nM}(k)$	2	4	**5**	6	8	9	**10**	**11**	12	**13**	14	16	17
$\mathrm{nM}_1(k)$	2	4	**5**	6	8	8	**10**	**11**	12	**13**	14	16	16
$\mathrm{VDM}(k)$	2	4	**5**	6	8	8	**10**	**11**	12	**13**	14	16	16
$\mathrm{VDH}(k)$	2	4	**5**	6	8	8	**10**	**11**	12	**13**	14	16	16
$\mathrm{FCM}(k)$	2	4	**4**	6	8	8	**9**	**10**	12	**12**	14	16	16
$\mathrm{FCH}(k)$	2	4	**4**	6	8	8	**8**	**10**	12	**12**	14	16	16
$S_2(k)$	2	4	**4**	6	8	8	**8**	**10**	12	**12**	14	16	16

Table 1. Values for the fundamental quantities for $1 \le k \le 13$; in bold the columns not in $\mathcal{SNM}_1$, while the vertical bars are left of the special deficiencies $2^n - n$, $n \ge 2$.

Strengthening [23, Corollary 12.13], first we show strong properties for minimally unsatisfiable clause-sets F such that the number of full clauses equals the min-var-degree, i.e., there is a variable which occurs only in the full clauses. We use $\mathrm{var}_{\mu\mathrm{vd}}(F) := \{v \in \mathrm{var}(F) : \mathrm{vd}_F(v) = \mu\mathrm{vd}(F)\}$ for $F \in \mathcal{CLS}$ with $n(F) > 0$ (the set of variables with minimal degree). Furthermore we use $\mathrm{DP}_v(F)$

for $F \in \mathcal{CLS}$ and $v \in \mathrm{var}(F)$ for the result of replacing the clauses containing variable v by their resolvents on v; indeed the special use in Lemma 45 will be just the inverse of the expansion process from Definition 36.

Lemma 45. *Consider $F \in \mathcal{MU}$ with $\mathrm{fc}(F) = \mu\mathrm{vd}(F)$ (and thus $n(F) > 0$).*

1. *$\mathrm{var}_{\mu\mathrm{vd}}(F)$ is the set of all $v \in \mathrm{var}(F)$ which occur only in full clauses of F.*
2. *$\mathrm{fc}(F)$ is even.*
3. *For $v \in \mathrm{var}_{\mu\mathrm{vd}}(F)$ and $F' := \mathrm{DP}_v(F)$ we have $F' \in \mathcal{MU}_{\delta=\delta(F)-\frac{\mathrm{fc}(F)}{2}+1}$.*
4. *$\mathrm{fc}(F) \le 2 \cdot \mathrm{FCM}(\delta(F) - \frac{\mathrm{fc}(F)}{2} + 1)$.*

Proof. Consider $v \in \mathrm{var}(F)$ with $\mathrm{vd}_F(v) = \mu\mathrm{vd}(F)$. The occurrences of v are now exactly in the full clauses of F (Part 1). Every full clauses must be resolvable on v, and thus the full clauses of F can be partitioned into pairs $\{v\} \cup C, \{\overline{v}\} \cup C$ for $\frac{\mathrm{fc}(F)}{2}$ many clauses C. This shows Part 2. Parts 3, 4 now follow by considering $F' := \mathrm{DP}_v(F)$: F' is obtained by replacing the full clauses of F by the clauses C (i.e., performs a full subsumption resolution, which are all strict except of the last one, which is non-strict). The new clauses C are full in F' (though there might be other full clauses in F'). Obviously $F' \in \mathcal{MU}$ and $\delta(F') = \delta(F) - \frac{\mathrm{fc}(F)}{2} + 1$. □

For deficiency $k = 7$ we have the first case of $\mathrm{FCH}(k) < \mathrm{FCM}(k)$:

Lemma 46. $\mathrm{FCM}(7) = 9 = \mathrm{nM}(7) - 1$*, while* $\mathrm{FCH}(7) = 8 = S_2(7)$*.*

Proof. By $S_2(7) = 8$ we have $\mathrm{FCH}(7) \ge 8$. By Lemma 45, Part 4 and by $\mathrm{FCM}(3) = 4$ the assumption of $\mathrm{FCM}(7) = 10 = \mathrm{nM}(7)$ yields the contradiction $10 \le 2\,\mathrm{FCM}(7 - 5 + 1) = 2 \cdot 4 = 8$, and thus $\mathrm{FCM}(7) \le 9$. By Lemma 1 we obtain $\mathrm{FCH}(7) = 8$. A clause-set $F \in \mathcal{MU}_{\delta=7}$ with $\mathrm{fc}(F) = 9$ (and $n(F) = 4$) is given by the following variable-clause-matrix:

$$\begin{pmatrix} - & - & + & + & - & - & + & - & - & + & 0 \\ + & + & - & - & - & - & + & - & + & - & 0 \\ + & - & + & - & + & - & 0 & + & + & + & - \\ + & + & + & + & + & + & 0 & - & - & - & - \end{pmatrix}$$

Let the variables be $1, \dots, 4$, as indices of the rows. Now setting variable 4 to `false` yields A_3, where one non-strict subsumption resolution has been performed, while setting variable 4 to `true` followed by unit-clause propagation of $\{-3\}$ yields A_2. So both instantiations yield minimally unsatisfiable clause-sets, whence by [23, Lemma 3.15, Part 2] $F \in \mathcal{MU}$.[5)] □

We are ready to prove the final main result of this paper:

Theorem 47. *Table 1 is correct.*

[5)] [23, Lemma 3.15] contains a technical correction over [21, Lemma 1].

Proof. The values for $1 \leq k \leq 6$ have been determined in [23, Section 14]. We observe that $1, 2, 4, 5, 6, 9, 11, 12, 13 \in \mathcal{SNM}_1$, and thus by Theorem 42 nothing is to be done for these values, and only the deficiencies $7, 8, 10$ remain.

By Lemma 45, Part 2, we get that $\mathrm{FCH}(8) = \mathrm{FCM}(8) = 10$ (since $\mathrm{nM}(8) = 11$ is odd), and also $\mathrm{FCH}(10) = \mathrm{FCM}(10) = 12$. By Lemma 46 it remains to provide unsatisfiable hitting clause-sets witnessing $\mathrm{VDH}(7) = 10$, $\mathrm{VDH}(8) = 11$ and $\mathrm{VDH}(10) = 13$. For deficiency 7 consider

$$F_7 := \begin{pmatrix} 0 & + & - & + & - & + & - & - & + & - & + \\ 0 & - & + & + & - & - & + & - & - & + & + \\ - & + & + & + & - & - & - & + & + & + & 0 \\ - & - & - & - & + & + & + & 0 & + & + & + \end{pmatrix}.$$

F_7 has 4 variables and 11 clauses, thus $\delta(F_7) = 11 - 4 = 7$; the hitting property is checked by visual inspection, and F_7 is unsatisfiable due to $8 \cdot 2^{-4} + 2 \cdot 2^{-3} + 2^{-2} = \frac{1}{2} + \frac{1}{4} + \frac{1}{4} = 1$, while finally every row contains exactly one 0, and thus F_7 is variable-regular of degree $10 = \mathrm{nM}(7)$.

Finally consider A_4 with $\delta(A_4) = 16 - 4 = 12$ and $\mu\mathrm{vd}(A_4) = 16$: perform four strict full subsumption resolutions on variables $1, 2, 3, 4$, and obtain elements of $\mathcal{UHIT}$ of deficiency $11, 10, 9, 8$ with min-var-degree $14, 13, 12, 11$. □

8 Conclusion and Outlook

In this paper we have improved the understanding of the four fundamental quantities, by supplying the lower bound $S_2 \leq \mathrm{FCH}$. The recursion defining S'_2 sheds also light on $S_2 = S'_2$, and we gained a deeper understanding of $S_2 = 2a_2$. Moreover we believe

Conjecture 48. $\forall\, k \in \mathbb{N} : S_2(k) = \mathrm{FCH}(k)$.

This would indeed give an unexpected precise connection of combinatorial SAT theory and elementary number theory. On the upper bound side, by Conjectures 12.1, 12.6 in [23] (see Figure 1 there for a summary of the relations between the four fundamental quantities) we get:

Conjecture 49. $\forall\, k \in \mathbb{N} : \mathrm{nM}(k) - 1 \leq \mathrm{FCM}(k) \leq \mathrm{VDM}(k) = \mathrm{VDH}(k)$.

Recall that $\mathrm{VDM}(k) \leq \mathrm{nM}(k)$; so we believe that three of the four fundamental quantities are very close to $\mathrm{nM}(k)$. This is in contrast to $\mathrm{nM}(k) - S_2(k)$ being unbounded, and indeed $S_2(k) = k + 1$ for infinitely many k (Corollary 44), while by Lemma 43 we also know $S_2(k) = \mathrm{nM}(k)$ for infinitely many k, and thus S_2 oscillates between the linear function $k+1$ and the quasi-linear function $\mathrm{nM}(k)$. To eventually determine the four fundamental quantities (which, if our conjectures are true, boil down to VDM and FCM, while VDH = VDM and FCH = S_2), detailed investigations like those in Section 7 need to be continued.

As $\mathrm{FCH}(k)$ and $S_2(k)$ are closely related via (boolean) hitting clause-sets, via generalised (non-boolean) (hitting-)clause-sets (see [19,20] for the basic theory) we can establish a close connection to the $S_p(k)$ for all prime numbers p in forthcoming work ($S_p(k)$ is the smallest $n \in \mathbb{N}_0$ such that p^k divides $n!$, introduced in [28, Unsolved Problem 49]).

References

1. Ron Aharoni and Nathan Linial. Minimal non-two-colorable hypergraphs and minimal unsatisfiable formulas. *Journal of Combinatorial Theory, Series A*, 43(2):196–204, November 1986. doi:10.1016/0097-3165(86)90060-9.
2. George Boole. *An Investigation of The Laws of Thought, on which are founded The Mathematical Theorie of Logic and Probabilities.* Dover Publication, Inc., first published in 1958. ISBN 0-486-60028-9; printing of the work originally published by Macmillan in 1854, with all corrections made within the text. Available from: http://www.gutenberg.org/ebooks/15114.
3. B.W. Conolly. Meta-Fibonacci sequences (a letter from B.W. Conolly). In *Fibonacci and Lucas Numbers and the Golden Section*, chapter 12, pages 127–138. Dover Publications, Inc., 2008. Unabridged republication of the work originally published in 1989 by Ellis Horwood Limited, Chichester, England.
4. Yves Crama and Peter L. Hammer. *Boolean Functions: Theory, Algorithms, and Applications*, volume 142 of *Encyclopedia of Mathematics and Its Applications*. Cambridge University Press, 2011. ISBN 978-0-521-84751-3.
5. Gennady Davydov, Inna Davydova, and Hans Kleine Büning. An efficient algorithm for the minimal unsatisfiability problem for a subclass of CNF. *Annals of Mathematics and Artificial Intelligence*, 23(3-4):229–245, 1998. doi:10.1023/A:1018924526592.
6. Herbert Fleischner, Oliver Kullmann, and Stefan Szeider. Polynomial–time recognition of minimal unsatisfiable formulas with fixed clause–variable difference. *Theoretical Computer Science*, 289(1):503–516, November 2002. doi:10.1016/S0304-3975(01)00337-1.
7. W.M. Gorman. More scope for qualitative economics. *The Review of Economics Studies*, 31(1):65–68, January 1964. Available from: http://www.jstor.org/pss/2295936.
8. Douglas R. Hofstadter. *Gödel, Escher, Bach: An eternal golden braid.* Basic Books, 1979. Pdf version with 801 pages, md5sum="0cb32e8ea5dd2485f63842f5acffb3f0 GEBen.pdf". Available from: http://www.physixfan.com/wp-content/files/GEBen.pdf.
9. Henry Ibstedt. *Computer Analysis of Number Sequences.* American Research Press, 1998. Available from: http://www.gallup.unm.edu/~smarandache/Ibstedt-computer.pdf.
10. Brad Jackson and Frank Ruskey. Meta-Fibonacci sequences, binary trees and extremal compact codes. *The Electronic Journal of Combinatorics*, 13(1), March 2006. R26; Corrigendum November 2007. Available from: http://www.emis.de/journals/EJC/Volume_13/Abstracts/v13i1r26.html.
11. Victor Klee. Recursive structure of S-matrices and an $O(m^2)$ algorithm for recognizing strong sign solvability. *Linear Algebra and its Applications*, 96:233–247, November 1987. doi:10.1016/0024-3795(87)90347-8.
12. Victor Klee and Richard Ladner. Qualitative matrices: Strong sign-solvability and weak satisfiability. In Harvey J. Greenberg and John S. Maybee, editors, *Computer-Assisted Analysis and Model Simplification*, pages 293–320, 1981. Proceedings of the First Symposium on Computer-Assisted Analysis and Model Simplification, University of Colorado, Boulder, Colorado, March 28, 1980. doi:10.1016/B978-0-12-299680-1.50022-7.
13. Victor Klee, Richard Ladner, and Rachel Manber. Signsolvability revisited. *Linear Algebra and its Applications*, 59:131–157, June 1984. doi:10.1016/0024-3795(84)90164-2.

14. Hans Kleine Büning. On subclasses of minimal unsatisfiable formulas. *Discrete Applied Mathematics*, 107(1-3):83–98, 2000. doi:10.1016/S0166-218X(00)00245-6.
15. Hans Kleine Büning and Oliver Kullmann. Minimal unsatisfiability and autarkies. In Armin Biere, Marijn J.H. Heule, Hans van Maaren, and Toby Walsh, editors, *Handbook of Satisfiability*, volume 185 of *Frontiers in Artificial Intelligence and Applications*, chapter 11, pages 339–401. IOS Press, February 2009. doi:10.3233/978-1-58603-929-5-339.
16. Oliver Kullmann. An application of matroid theory to the SAT problem. In *Proceedings of the 15th Annual IEEE Conference on Computational Complexity*, pages 116–124, July 2000. doi:10.1109/CCC.2000.856741.
17. Oliver Kullmann. Investigations on autark assignments. *Discrete Applied Mathematics*, 107:99–137, 2000. doi:10.1016/S0166-218X(00)00262-6.
18. Oliver Kullmann. Lean clause-sets: Generalizations of minimally unsatisfiable clause-sets. *Discrete Applied Mathematics*, 130:209–249, 2003. doi:10.1016/S0166-218X(02)00406-7.
19. Oliver Kullmann. Constraint satisfaction problems in clausal form I: Autarkies and deficiency. *Fundamenta Informaticae*, 109(1):27–81, 2011. doi:10.3233/FI-2011-428.
20. Oliver Kullmann. Constraint satisfaction problems in clausal form II: Minimal unsatisfiability and conflict structure. *Fundamenta Informaticae*, 109(1):83–119, 2011. doi:10.3233/FI-2011-429.
21. Oliver Kullmann and Xishun Zhao. On variables with few occurrences in conjunctive normal forms. In Laurent Simon and Karem Sakallah, editors, *Theory and Applications of Satisfiability Testing - SAT 2011*, volume 6695 of *Lecture Notes in Computer Science*, pages 33–46. Springer, 2011. doi:10.1007/978-3-642-21581-0_5.
22. Oliver Kullmann and Xishun Zhao. On Davis-Putnam reductions for minimally unsatisfiable clause-sets. *Theoretical Computer Science*, 492:70–87, June 2013. doi:10.1016/j.tcs.2013.04.020.
23. Oliver Kullmann and Xishun Zhao. Bounds for variables with few occurrences in conjunctive normal forms. Technical Report arXiv:1408.0629v3 [math.CO], arXiv, November 2014. Available from: http://arxiv.org/abs/1408.0629.
24. Frank Ruskey and Chris Deugau. The combinatorics of certain k-ary meta-Fibonacci sequences. *Journal of Integer Sequences*, 12, May 2009. Article 09.4.3. Available from: http://cs.uwaterloo.ca/journals/JIS/VOL12/Ruskey/ruskey6.html.
25. Winfrid G. Schneeweiss. *Boolean Functions with Engineering Applications and Computer Programs*. Springer-Verlag, 1989. ISBN 3-540-18892-4. doi:10.1007/978-3-642-45638-1.
26. Claude Elwood Shannon. A symbolic analysis of relay and switching circuits. Master's thesis, Massachusetts Institute of Technology. Dept. of Electrical Engineering, 1940. Available from: http://dspace.mit.edu/handle/1721.1/11173.
27. Neil J.A. Sloane. The On-Line Encyclopedia of Integer Sequences (OEIS), 2008. Available from: http://oeis.org/.
28. Florentin Smarandache. *Only problems, not solutions!* Xiquan Publishing House, fourth edition, 1993. Available from: http://www.gallup.unm.edu/~smarandache/OPNS.pdf.
29. Zhang Wenpeng, editor. *Research on Smarandache Problems in Number Theory (Collected papers)*. Hexis, Phoenix, USA, 2004. Available from: http://fs.gallup.unm.edu/wenpeng-book.pdf.

30. Zhang Wenpeng and Liu Duansen. On the primitive numbers of power P and its asymptotic property. In Jack Allen, Feng Liu, and Dragoş Constantinescu, editors, *Smarandache Notions*, pages 173–175. American Research Press, Rehoboth, NM, USA, 2002. Available from: http://dl.acm.org/citation.cfm?id=773309.
31. Doron Zeilberger. How Berger, Felzenbaum and Fraenkel revolutionized covering systems the same way that George Boole revolutionized logic. *The Electronic Journal of Combinatorics*, 8(2):1–10, 2001. The Fraenkel Festschrift volume; Article #A1. Available from: http://www.combinatorics.org/ojs/index.php/eljc/article/view/v8i2a1.