

An application of matroid theory to the SAT problem

Oliver Kullmann*

Department of Computer Science
University of Toronto
Toronto, Ontario M5S 3G4

e-mail: kullmann@cs.toronto.edu

<http://www.cs.utoronto.ca/~kullmann/>

Abstract

We consider the deficiency $\delta(F) := c(F) - n(F)$ and the maximal deficiency $\delta^*(F) := \max_{F' \subseteq F} \delta(F')$ of a clause-set F (a conjunctive normal form), where $c(F)$ is the number of clauses in F and $n(F)$ is the number of variables.

Combining ideas from matching and matroid theory with techniques from the area of resolution refutations, we prove that for clause-sets F with $\delta^*(F) \leq k$, where k is considered as a constant, the SAT problem, the minimally unsatisfiability problem and the MAXSAT problem are decidable in polynomial time (previously, only poly-time decidability of the minimally unsatisfiability problem was known, and that only for $k = 1$).

1. Introduction

Minimally unsatisfiable clause-sets

A conjunctive normal form F (represented as a clause-set in this paper) is called *minimally unsatisfiable* iff F is unsatisfiable, and any strict subset is satisfiable. Let $MUSAT$ denote the class of all minimally unsatisfiable clause-sets. It has been shown in [15] that $MUSAT$ is D^P -complete, where $D^P = \{L_1 \cap L_2 : L_1 \in NP \wedge L_2 \in co-NP\}$ is the class of languages representable as the intersection of a language in NP and a language in co-NP.

Applying Hall's theorem, in [1] the fundamental property $c(F) \geq n(F) + 1$ for $F \in MUSAT$ (attributing this lemma to M. Tarsi) is shown, where $c(F)$ is the number

*Supported by the Natural Sciences and Engineering Research Council of Canada and by the Communications and Information Technology Ontario.

of clauses of F , and $n(F)$ is the number of variables in F . Now two *special classes* of minimally unsatisfiable clause-sets are known, where the decision problem is in polynomial time:

- $SMUSAT(1)$, the class of *strongly* minimally unsatisfiable clause-sets F with $c(F) = n(F) + 1$, where “strongly” additionally requires, that adding any literal to any clause of F renders F satisfiable (this class has been introduced in [1], and from their characterisation of this class poly-time decidability follows);
- $MUSAT(1)$, the class of all minimally unsatisfiable F with $c(F) = n(F) + 1$ (characterised and shown to be poly-time decidable in [6] (using a quite complicated proof)).

At the SAT'98 workshop ([8]) Hans Kleine Büning has posed the question whether *all classes* $MUSAT(k) := \{F \in MUSAT : c(F) \leq n(F) + k\}$ for constant $k \geq 1$ are *poly-time decidable* (answered positively in this article). Furthermore, in [4] for $F \in MUSAT(k)$ the upper bound $2^{k-1} \cdot n(F)^2$ for the minimal number of resolution steps in a *resolution tree refutation* of F has been established (a strengthening of this fact is basic for our considerations).

In essence, most techniques from the work mentioned above is included in our approach. Additionally, the following notions have been of importance.

Autarkies and the deficiency

Straightening and generalising prior work, the use of the notion of an *autarky* for clause-sets F (partial assignments satisfying each clause of F they “touch”) and of the *deficiency* $\delta(F) := c(F) - n(F)$ are basic for this paper.

Being unaware of [1], in [7] an alternative (inductive) proof of “Tarsi’s lemma” is given, based on the (valuable) observation that substituting a truth value for any (single) variable in $F \in \mathcal{SMUSAT}$ results in a clause-set still in \mathcal{MUSAT} . They notice that a clause-set F with $1 \leq c(F) \leq n(F)$ must have a non-empty *autark subset* $F' \subseteq F$, that is, F' is satisfiable by an autarky (a partial assignment not using any variable appearing in $F \setminus F'$).

In [16] the property $c(F) \geq n(F) + 1$ has been shown for the more general class $F \in \mathcal{UMUSAT}$, the closure of \mathcal{MUSAT} under finite union. They also show for an arbitrary clause-set F and a subset $F' \subseteq F$ with the property that $c(F') - n(F')$ is maximal, that for every $F'' \subseteq F \setminus F'$ one has $c(\text{var}(F') * F'') - n(\text{var}(F') * F'') \leq 0$, where $\text{var}(F') * F''$ is obtained from F'' by crossing out all variables appearing in F' . *Actually* this says that $F \setminus F'$ is an *autark subset* of F , and thus for $F \in \mathcal{UMUSAT}$ we must have $F' = F$, and property (1) given in the next paragraph follows. Unfortunately the argumentation in [16] is very lengthy and does not arrive at such conceptual insights.

Independent of [16], the notion $\delta(F) := c(F) - n(F)$ of *deficiency* has been introduced in [9] and made fruitful for lucid argumentation. The class of *matched clause-sets* F , given by the property, that $\delta(F') \leq 0$ holds for all $F' \subseteq F$, is introduced, and it is shown that this class is poly-time decidable and contains only satisfiable clause-sets. A short proof of the statement

$$\forall F' \subseteq F : \delta(F') < \delta(F) \quad (1)$$

for $F \in \mathcal{MUSAT}$ is given (since $\delta(\top) = 0$, where \top is the empty clause-set, Tarsi’s lemma follows).

These attempts have been unified and strengthened in [11], starting a systematic investigation of the notion of autarkies and autark subsets. *Generalising MUSAT* (and *UMUSAT*), the class \mathcal{LEAN} of *lean clause-sets* F has been introduced, where all clauses of F can be used in a resolution tree refutation of F . It is shown, that a clause-set F is lean if and only if \top is the *only autark subset* of F .

In order to obtain special cases of autark subsets, which are computable in polynomial time, the notion of *linearly autark subsets* is introduced, and the class \mathcal{LEAN} has been enlarged to \mathcal{LLEAN} , the class of *linearly lean* clause-sets where \top is the only linearly autark subset. \mathcal{LLEAN} is poly-time decidable (using Linear Programming), and property (1) in fact holds true for all clause-sets $F \in \mathcal{LLEAN}$.

The *maximal deficiency* $\delta^*(F) := \max_{F' \subseteq F} \delta(F')$, also investigated in [11], is of basic importance (see at the end of the following subsection). Note that F is a matched clause-set iff $\delta^*(F) = 0$, and that for all minimally unsatisfiable F we have $\delta^*(F) = \delta(F)$ by (1).

The contributions of this paper

Using $\mathcal{C}(k) := \{F \in \mathcal{C} : \delta^*(F) \leq k\}$ for any class \mathcal{C} contained in the set \mathcal{CLS} of all clause-sets (note the consistency with the prior use of this notation), we prove for any $k \geq 0$ (see Section 4):

- $\mathcal{SAT}(k)$ as well as the classes

$$\begin{aligned} \mathcal{SMUSAT}(k) &\subset \mathcal{MUSAT}(k) \subset \\ \mathcal{UMUSAT}(k) &\subset \mathcal{LEAN}(k) \subset \\ &\mathcal{USAT}(k) \cup \{\top\} \end{aligned}$$

are *poly-time decidable* (\mathcal{SAT} and \mathcal{USAT} are the satisfiable resp. unsatisfiable clause-sets).

- For $F \in \mathcal{CLS}(k)$ the two *decompositions*

$$F = F_1^i \cup F_2^i, F_1^i \cap F_2^i = \emptyset$$

($i = 1, 2$) given by

- F_1^1 is the *largest lean subset* of F , F_2^1 is the *largest autark subset* of F
- F_1^2 is the union of all *minimally unsatisfiable* sub-clause-sets of F , F_2^2 the intersection of all *maximally satisfiable* sub-clause-sets

are computable in *polynomial time*, as well as a *largest satisfiable subset*.

Furthermore the improved (and now tight) upper bound $2^{\delta^*(F)-1} \cdot n(F)$ for the minimal number of resolution steps needed in a *resolution tree refutation* is derived.

The basic new insights towards these results are:

1. In Theorem 3.2 lean sub-clause-sets of $F \in \mathcal{CLS}$ are shown to be *circuits* of some *matroids*, derived from the *transversal matroid* $T(F)$ given by the (natural) bipartite graph associated with F . The maximal deficiency $\delta^*(F')$ for subsets $F' \subseteq F$ is the *nullity* $c(F') - \text{rank}(F')$ in $T(F)$.

It follows that for fixed maximal deficiency of F a set of sub-clause-sets containing all lean sub-clause-sets can be computed in *polynomial time*. (See Appendix A, where also some basic notions from matroid theory are explained.)

2. In Theorem 3.11 it is shown that unsatisfiable clause-sets with fixed maximal deficiency have *small* “DLL-like refutations” (or “(generalised) semantic tree refutations”) of a *special form*, which can be searched for in polynomial time.

2. Notation

\mathcal{CLS} is the set of all clause-sets (conjunctive normal forms treated as sets), where we require clauses to be complement-free, \mathcal{SAT} is the set of all satisfiable clause-sets and \mathcal{USAT} the set of all unsatisfiable clause-sets. The set of minimally unsatisfiable clause-sets is denoted by \mathcal{MUSAT} (unsatisfiable clause-sets such that every strict subset is satisfiable), while \mathcal{SMUSAT} (called “strongly minimal” in [1] and “saturated” in [7]) is the set of minimally unsatisfiable clause-sets such that adding any literal to any clause renders them satisfiable. Finally \mathcal{UMUSAT} is the closure of \mathcal{MUSAT} under finite union (introduced in [16]). The complement of a literal x is denoted by \bar{x} , while for a clause C we set $\bar{C} := \{\bar{x} : x \in C\}$. Variables appearing in a literal x , a clause C or a clause-set F are denoted by $\text{var}(x)$, $\text{var}(C)$ and $\text{var}(F)$ respectively (the latter two being sets). For $F \in \mathcal{CLS}$ let $n(F)$ be the number of variables and $c(F)$ the number of clauses, while $\#_x(F)$ is the number of occurrences of literal x ; $\delta(F) := c(F) - n(F)$ and $\delta^*F := \max_{F' \subseteq F} \delta(F')$. And for any $\mathcal{C} \subseteq \mathcal{CLS}$ and $k \geq 0$ let $\mathcal{C}(k) := \{F \in \mathcal{C} : \delta^*(F) \leq k\}$.

For any partial assignment φ (assigning truth values to variables) we denote by $\varphi * F$ the result of applying φ to F , that is, all clauses satisfied by φ are removed, and from the remaining clauses all literals falsified by φ are cancelled. $\langle x \rightarrow \varepsilon \rangle$ denotes the partial assignment setting literal x to ε (and the complement \bar{x} to $\bar{\varepsilon}$). φ is called an **autarky** for F if each clause of F either is satisfied by φ or has no variables given a value by φ . And for a set V of variables let $V * F$ denote the clause-set resulting when crossing out all literals with variables in V . ($V * F$ is equal to the union of all $\varphi * F$ with $\text{var}(\varphi) = V$.)

3. The main lemmas

3.1. Circuits of the transversal matroid $T(F)$

Definition 3.1 For $F \in \mathcal{CLS}$ and $k \geq 0$ let $\mathfrak{C}_k(F)$ be the set of all sub-clause-sets $F' \subseteq F$ with $\delta(F') = k + 1$ such that for all $F'' \subset F'$ one has $\delta(F'') \leq k$. Furthermore $\mathfrak{C}(F) := \bigcup_{0 \leq k < \delta^*(F)} \mathfrak{C}_k(F)$. ■

Theorem 3.2 Consider $F \in \mathcal{CLS}$ and $k \geq 0$. Then \mathfrak{C}_{k+1} is the set of circuits of a matroid $\mathbf{T}_k(F)$ on F (the elements are the clauses of F), where the independent subsets of $\mathbf{T}_k(F)$ are those $F' \subseteq F$ with $\delta^*(F') \leq k$.

Proof: First we consider the case $k = 0$. The set $\{F' \subseteq F : \delta^*(F') = 0\}$ is the set of independent subsets of

the *transversal matroid* $T(F) = T(B(F))$ induced by the bipartite graph $B(F) = (F, R, \text{var}(F))$, where the relation $(C, v) \in R \subseteq F \times \text{var}(F)$ holds iff $v \in \text{var}(C)$ (see Appendix A).

The set of circuits of $T(F)$ is $\mathfrak{C}_0(F)$. For the rank in $T(F) = T_0(F)$ we have $\text{rank}(F') = c(F') - \delta^*(F')$ for all $F' \subseteq F$. The function $f(F') := \text{rank}(F') + k$ is sub-modular, increasing and integer-valued, and thus the set \mathfrak{C}'_k consisting of all $F' \subseteq F$ which are minimal w.r.t. the property “ $f(F') < c(F')$ ” is the set of circuits of a matroid $T_k(F)$, whose set of independent subsets is

$$\mathfrak{I}'_k = \{F' \subseteq F : \forall F'' \subseteq F' [f(F'') \geq c(F'')]\}$$

(see [14], page 380, 12.1.1 and 12.1.2). Since for $C \in F$ and $F' \subseteq F$ we have $f(F' \cup \{C\}) \leq f(F') + 1$, in fact $\mathfrak{I}'_k = \{F' \subseteq F : f(F') \geq c(F')\} = \{F' \subseteq F : \delta^*(F') \leq k\}$ and $\mathfrak{C}'_k = \mathfrak{C}_k$ holds. ■

Using the enumeration of the circuits of $T_k(F)$ in Appendix A, we get

Corollary 3.3 For $k \geq 0$ and $F \in \mathcal{CLS}(k)$ the sets $\mathfrak{C}_0(F), \dots, \mathfrak{C}_{k-1}(F)$ can be computed in polynomial time (k is treated as constant). ■

3.2. Lean clause-sets

For a clause-set $F \in \mathcal{CLS}$ we call $F' \subseteq F$ an **autark subset** of F iff there is an autarky φ for F with $F' = \{C \in F : \varphi(C) = 1\}$. We make frequent use of the following reformulation:

$$F' \text{ autark subset of } F \Leftrightarrow \text{var}(F \setminus F') * F' \in \mathcal{SAT}. \quad (2)$$

Generalising minimally unsatisfiable clause-sets, in [11] the notion of *lean clause-sets* has been introduced, which are clause-sets such that every clause *can* be used in a resolution refutation. It has been shown, that a clause-set is lean if and only if it has no non-trivial autarky. In our context, we take this as the definition of lean clause-sets.

Definition 3.4 \mathcal{LEAN} is the set of autarky-free clause-set $F \in \mathcal{CLS}$, that is, the only autark subset of F is \top . ■

Simple properties (for more information see [11]):

1. $\mathcal{UMUSAT} \subset \mathcal{LEAN} \subset \mathcal{USAT} \cup \{\top\}$.
2. \mathcal{LEAN} is stable under crossing out variables and union, that is, for $F, F_1, F_2 \in \mathcal{LEAN}$ and $V \subseteq \mathcal{VA}$ we have $V * F \in \mathcal{LEAN}$ as well as $F_1 \cup F_2 \in \mathcal{LEAN}$.
3. $F \in \mathcal{LEAN} \Leftrightarrow F \setminus \{\perp\} \in \mathcal{LEAN}$.

For detecting autarkies in our context, the following inequality (with straightforward proof) from [11] is basic.

Lemma 3.5 For any $F \in \mathcal{CLS}$ and any $F' \subseteq F$ we have $\delta^*(\text{var}(F') * (F \setminus F')) + \delta(F') \leq \delta^*(F)$. ■

Since $\delta^*(F) = 0$ implies $F \in \mathcal{SAT}$ (every maximum matching in $B(F)$ induces a satisfying assignment; clause-sets F with $\delta^*(F) = 0$ are called “matched clause-sets” in [9]), by (2) we get the following sufficient condition for a subset being autark (also included in [16], while a fore-runner can be found in [7]; for a refinement using linear programming see [11]).

Corollary 3.6 For $F \in \mathcal{CLS}$ and $F' \subseteq F$ with $\delta(F') = \delta^*(F)$ the set $F \setminus F'$ is an autark subset of F . ■

From Corollary 3.6 and Definition 3.4 we obtain the following fundamental lemma (see [11] for a generalisation; in [9] Lemma 3.7 is proved for the special case of minimally unsatisfiable clause-sets).

Lemma 3.7 For $F \in \mathcal{LEAN}$ we have $\delta(F') < \delta(F)$ for all $F' \subset F$. It follows $\delta(F) \geq 1$ for $F \neq \top$. ■

Hence all lean sub-clause-sets of F (and thus all minimally unsatisfiable sub-clause-sets) are contained in $\mathfrak{C}(F)$:

$$\forall F \in \mathcal{CLS} \forall F' \subseteq F : F' \in \mathcal{LEAN} \Rightarrow F' \in \mathfrak{C}(F). \quad (3)$$

See also the forthcoming paper [12], where clause-sets F with property (1) are characterised as “matching lean clause-sets”, which in turn is equivalent to the property that $T(F)$ is “cyclic” (see [3]).

Corollary 3.8 $\mathcal{MUSAT}(1) = \mathcal{LEAN}(1) \setminus \{\top\}$. ■

3.3. Splitting up lean clause-sets

Now we turn to the problem of investigating the deficiency of clause-sets when performing splitting, which is, besides Corollary 3.3, the basis for our decision algorithms, and also establishes the connection to tree-like resolution. For the proofs of Lemmas 3.9 and 3.10 see Appendix D.

In general for any clause-set $F \in \mathcal{CLS}$ and any literal x we have

$$\delta^*(\langle x \rightarrow 1 \rangle * F) \leq \delta^*(F) + 1 \quad (4)$$

since for every $F' \subseteq \langle x \rightarrow 1 \rangle * F$ there is a $F'' \subseteq F$ with $c(F'') = c(F')$ and $n(F'') \leq n(F') + 1$. Our problem is to actually *reduce* the deficiency by splitting.

Part 3 of the following lemma generalises (and simplifies) the “splitting theorem” from [4].

Lemma 3.9 Consider $F \in \mathcal{LEAN}$ such that for all variables $v \in \text{var}(F)$ we have $(\#_v + \#\bar{v})(F) \geq 3$.

1. $c(\{C \in F : \text{var}(C) \cap V \neq \emptyset\}) \geq |V| + 2$ holds for any non-empty set $V \subseteq \text{var}(F)$ of variables of F .
2. $\delta(F') \leq \delta(F) - 2$ holds for any subset $F' \subseteq F$ with $\text{var}(F') \subset \text{var}(F)$.
3. For all literals x with $\#_x(F) \geq 2$ we have

$$\delta^*(\langle x \rightarrow 1 \rangle * F) \leq \delta(F) - 1.$$

Using “Davis-Putnam resolution” (see Appendix B) we get

Lemma 3.10 Consider $F \in \mathcal{USAT}$ and $F' \subseteq F$ with $F' \in \mathcal{MUSAT}$ and $\delta(F') \geq 2$. Then there is a variable $v \in \text{var}(F)$ such that for both $\varepsilon \in \{0, 1\}$ the clause-set $\langle v \rightarrow \varepsilon \rangle * F$ contains some $F'_\varepsilon \in \mathcal{MUSAT}(\delta(F') - 1)$.

The following theorem now immediately follows.

Theorem 3.11 Consider $F \in \mathcal{USAT}(k)$ for some $k \geq 1$. Then there is a binary tree T of height at most $k - 1$ with

- the root of T is labelled by F ;
- the two childs of an inner node labelled by F' are labelled by $\langle v \rightarrow 0 \rangle * F'$ resp. $\langle v \rightarrow 1 \rangle * F'$ for some variable $v \in \text{var}(F')$;
- every clause-set G labelling some leaf contains some $G' \subseteq G$ with $G' \in \mathcal{MUSAT}(1)$. ■

From the results in [12] it follows that the class

$$\mathcal{C} := \{F \in \mathcal{CLS} \mid \exists F' \subseteq F : F' \in \mathcal{MUSAT}(1)\}$$

of possible leaf labels of trees T is (properly) contained in the class of clause-set refutable by “read-once” resolution (resolution refutations in tree form, where every input clause can be used at most once), shown to be NP-complete in [10]. It is an interesting question whether also the problem “ $F \in \mathcal{C}$?” is NP-complete.

We conclude this section by an application to resolution complexity. Theorem 3.11 together with Lemma C.3 yields immediately

Corollary 3.12 $2^{\delta^*(F)-1} \cdot n(F)$ is an upper bound for the minimal number of inner nodes in resolution tree refutations of $F \in \mathcal{USAT}$. ■

4. The poly-time computability results

Theorem 4.1 For fixed $k \geq 0$ the class $SAT(k)$ is decidable in polynomial time. It follows that also $MUSAT(k)$ and $SMUSAT(k)$ are poly-time decidable. Furthermore for $F \in \mathcal{CLS}(k)$ a satisfiable $F' \subseteq F$ of largest size (containing at least $c(F) - \delta^*(F)$ clauses) is poly-time computable (that is, the MAXSAT problem is poly-time solvable for $\mathcal{CLS}(k)$).

Proof: Given $F \in \mathcal{CLS}$, enumerate all “splitting trees T ” of F of height at most $k - 1$ (fulfilling the first two conditions of Theorem 3.11). Now for a leaf label G of T we have $\delta^*(G) \leq k + (k - 1) = 2k - 1$ by (4). Hence the property “ $\exists G' \subseteq G : G' \in MUSAT(1)$?” is decidable in poly-time by Corollary 3.3 and Lemma C.3.

It follows that the existence of a proof T of unsatisfiability fulfilling the conditions from Theorem 3.11 can be decided in poly-time. For $F \in \mathcal{USAT}(k)$ such a proof must exist and hence SAT decision for $F \in \mathcal{CLS}(k)$ is poly-time.

W.r.t the properties “minimally unsatisfiable” and “saturated minimally unsatisfiable” now we can simply check whether the definitions hold. And w.r.t. MAXSAT note that a basis F' of the transversal matroid $T(F)$ is computable in polynomial time. We have $F' \in SAT$ and $c(F') = c(F) - \delta^*(F)$, and thus the question is whether a larger satisfiable subset exists. So we simply search through all (polynomially many) subset $F'' \subseteq F$ with $c(F'') > c(F) - \delta^*(F)$ and decide whether $F'' \in SAT$ holds. ■

The unique “autarky-decomposition”

From [11] we get the following unique decomposition of $F \in \mathcal{CLS}$ into the largest lean subset F_1 and the largest autark subset F_a

$$F = F_1 \cup F_a, \quad F_1 \cap F_a = \emptyset \quad (5)$$

where F_1 has the following characterisations:

- F_1 is the union of all lean subsets $F' \in \mathcal{LEAN}$ of F ;
- F_1 is the largest lean subset $F' \in \mathcal{LEAN}$ of F ;
- F_1 is the set of all clauses $C \in F$ such that a resolution tree refutation of F exists using C as an axiom;

on the other side, F_a has the following characterisations:

- F_a is the union of all autark subsets of F ;
- F_a is the largest autark subset of F .

Theorem 4.2 For fixed $k \geq 0$ and $F \in \mathcal{CLS}(k)$ the unique decomposition of F according to equation (5) is computable in polynomial time. Thus the class $\mathcal{LEAN}(k)$ is poly-time decidable.

Proof: Since F_a is the largest autark subset of F , by (2) and (3) we can characterise F_1 as the smallest subset (w.r.t. inclusion) $F' \subseteq F$ with

$$F' \in \mathcal{C}(F) \text{ and } \text{var}(F') * (F \setminus F') \in SAT.$$

Lemmas 3.5, 3.7 give $\delta^*(\text{var}(F') * (F \setminus F')) \leq \delta^*(F)$. Hence by Corollary 3.3 and Theorem 4.1 we can find F_1 in polynomial time. ■

”Minimally unsatisfiable” vs. ”maximally satisfiable”

Lemma 4.3 For a clause-set F let

$$\begin{aligned} \min(F) &:= \{F' \subseteq F : F' \text{ minimally unsatisfiable}\} \\ \max(F) &:= \{F' \subseteq F : F' \text{ maximally satisfiable}\}, \end{aligned}$$

where $F' \subseteq F$ is maximally satisfiable iff F' is satisfiable and every $F'' \subseteq F$ with $F' \subset F''$ is unsatisfiable. Then

$$\begin{aligned} \bigcup \min(F) \cup \bigcap \max(F) &= F \\ \bigcup \min(F) \cap \bigcap \max(F) &= \emptyset. \end{aligned}$$

Proof: Consider $C \in F$:

If there is a minimally unsatisfiable subset $F' \subseteq F$ with $C \in F'$, then consider a maximally satisfiable $F'' \subseteq F$ with $F' \setminus \{C\} \subseteq F''$ — here we have $C \notin F''$.

If there is a maximally satisfiable subset $F' \subseteq F$ with $C \notin F'$, consider a minimally unsatisfiable $F'' \subseteq F$ with $F'' \subseteq F' \cup \{C\}$ — now $C \in F''$. ■

Note that the largest autark subset F_a from equation (5) is contained in $\bigcap \max(F)$ (since for an autark subset F_1 of F and a satisfiable subset F_2 of F also $F_1 \cup F_2$ is satisfiable).

Theorem 4.4 For fixed $k \geq 0$ and $F \in \mathcal{CLS}(k)$ the unique decomposition of F according to Lemma 4.3 is computable in polynomial time. Thus the class $\mathcal{UMUSAT}(k)$ is poly-time decidable.

Proof: By Corollary 3.3, (3) in Subsection 3.2 and Theorem 4.1 all minimally unsatisfiable subsets of F can be enumerated in poly-time. ■

To conclude, it seems to me an interesting observation that \mathcal{UMUSAT} is the closure of \mathcal{MUSAT} under crossing out variables, that is

$$\mathcal{UMUSAT} = \{V * F : V \subseteq \mathcal{VA} \wedge F \in \mathcal{MUSAT}\}.$$

References

- [1] R. Aharoni and N. Linial. Minimal non-two-colorable hypergraphs and minimal unsatisfiable formulas. *Journal of Combinatorial Theory*, A 43:196–204, 1986.
- [2] M. Aigner. *Kombinatorik; II. Matroide und Transversaltheorie*. Springer-Verlag, 1976.
- [3] R. A. Brualdi. Transversal matroids. In N. White, editor, *Combinatorial Geometries*, volume 29 of *Encyclopedia of mathematics and its applications*, chapter 5, pages 72–97. Cambridge University Press, 1987.
- [4] H. K. Büning. An upper bound for minimal resolution refutations. In G. Gottlob, E. Grandjean, and K. Seyr, editors, *Computer Science Logic 12th International Workshop, CSL'98*, volume 1584 of *Lecture Notes in Computer Science*, pages 171–178. Springer, 1999.
- [5] S. A. Cook. An exponential example for analytic tableaux. Manuscript (see [17], page 432), 1973.
- [6] G. Davydov, I. Davydova, and H. K. Büning. An efficient algorithm for the minimal unsatisfiability problem for a subclass of CNF. *Annals of Mathematics and Artificial Intelligence*, 23:229–245, 1998.
- [7] T. Fliti and G. Reynaud. Sizes of minimally unsatisfiable conjunctive normal forms. Faculté des Sciences de Luminy, Dpt. Mathématique-Informatique, 13288 Marseille, France, November 1994.
- [8] J. Franco, G. Gallo, H. K. Büning, and E. Speckenmeyer, editors. *Second Workshop On The Satisfiability Problem*, 1998. Eringerfeld (Germany), May 10 - 14 (1998).
- [9] J. Franco and A. V. Gelder. A perspective on certain polynomial time solvable classes of satisfiability. To appear in *Discrete Applied Mathematics*, 2000.
- [10] K. Iwama and E. Miyano. Intractability of read-once resolution. In *Proceedings Structure in Complexity Theory, Tenth Annual Conference*, pages 29–36. IEEE, 1995.
- [11] O. Kullmann. Investigations on autark assignments. Submitted to *Discrete Applied Mathematics*, 1998.
- [12] O. Kullmann. On structural properties of minimally unsatisfiable clause-sets. In preparation, January 2000.
- [13] L. Lovász and M. D. Plummer. *Matching Theory*, volume 121 of *Mathematics Studies*. North-Holland, 1986.
- [14] J. G. Oxley. *Matroid Theory*. Oxford University Press, 1992.
- [15] C. H. Papadimitriou and D. Wolfe. The complexity of facets resolved. *Journal of Computer and System Sciences*, 37:2–13, 1988.
- [16] M. Shaohan and L. Dongmin. A polynomial-time algorithm for reducing the number of variables in MAX SAT problem. *Science in China (Series E)*, 40(3):301–311, June 1997.
- [17] A. Urquhart. The complexity of propositional proofs. *The Bulletin of Symbolic Logic*, 1(4):425–467, 1995.
- [18] H. Whitney. On the abstract properties of linear dependence. *American Journal of Mathematics*, 57:509–533, 1935.

APPENDIX

A. The induced transversal matroid of a bipartite graph

We recall some basic matroid theory (see any introductory text on the subject, like [14], Chapter 1). Let X be a finite set. A set $\mathcal{I} \subseteq \mathbb{P}(X)$ of subsets of X is the set of *independent subsets* of a matroid $M(X)$ on X iff $\emptyset \in \mathcal{I}$, with $A \in \mathcal{I}$ also $A' \in \mathcal{I}$ holds for all $A' \subseteq A$, and for all $A_1, A_2 \in \mathcal{I}$ with $|A_1| > |A_2|$ there is $a \in A_1$ fulfilling $A_2 \cup \{a\} \in \mathcal{I}$. $M(X)$ is uniquely determined by \mathcal{I} . A *circuit* of $M(X)$ is a minimal *dependent* (that is, not independent) subset of X . The *rank* of $A \subseteq X$ is the size of a maximal independent subset of A (called a *basis* of A — all basis' have the same cardinality).

A *submodular* function on X is a map $f : \mathbb{P}(X) \rightarrow \mathbb{R}$ from subsets of X to real numbers fulfilling

$$\forall A, B \in \mathbb{P}(X) : f(A \cap B) + f(A \cup B) \leq f(A) + f(B).$$

Consider a bipartite graph $G = (X, R, Y)$ ($R \subseteq X \times Y$). The function $A \subseteq X \mapsto |R(A)|$ is submodular (on X), increasing, integer-valued, and fulfils $f(\emptyset) = 0$, and thus

$$\mathcal{I}_G := \{ A \subseteq X \mid \forall A' \subseteq A : |A'| \leq |R(A')| \}$$

is the set of independent sets of a matroid on X , called the *transversal matroid* $T(G)$ (w.r.t. X). For a proof of this basic property of submodular functions see for example [14], Chapter 12.1 or [2], Chapter VI.1.

By Hall's theorem, the independent subsets of $T(G)$ are exactly the “partial transversals” $A \subseteq X$, that is, subsets A of X which can be covered by a matching in G , which in turn means, that there is an injective mapping $m : A \rightarrow Y$ with $\forall a \in A : (a, m(a)) \in R$.

For the rank in $T(G)$ we have the formula

$$\text{rank}(A) = |A| - \delta^*(A)$$

for $A \subseteq X$, where $\delta(A)^* := \max_{A' \subseteq A} \delta(A')$, and $\delta(A) := |A| - |R(A)|$ is the *deficiency* of A (see also [13], Chapter 1).¹⁾ $\text{rank}(A)$ is the size of a largest matching in the subgraph $(A, R|(A \times Y), Y)$, and thus $\text{rank}(A)$ as well as $\delta^*(A)$ are computable in polynomial time. Note that $A \subseteq X$ is independent iff $\text{rank}(A) = |A|$ iff $\delta^*(A) = 0$.

For a clause-set $F \in \mathcal{CLS}$ the bipartite graph $B(F)$ associated with F is $(F, R, \text{var}(F))$ where $(C, v) \in R \Leftrightarrow v \in \text{var}(C)$, that is, the “left side” of the bipartite graph are the clauses of F , the “right side” are the variables of F , and an edge joins clause C and variable v if v is contained in C (positively or negatively).

¹⁾For general matroids $M(X)$ the quantity $|A| - \text{rank}(A)$ for $A \subseteq X$ has been called the “nullity” of A in the fundamental paper [18], which generalises the notion of deficiency.

Computing the set of all circuits of a matroid

The set of all circuits of a matroid can be computed in polynomial time as follows (pointed out to me by Günter M. Ziegler).

Given a basis B of X and an element $x \in X \setminus B$, there is a unique circuit $C_x(B) \subseteq B \cup \{x\}$ (see [14], page 18, 1.2.6), and since every independent subset can be extended to a basis, it is easy to see that for every circuit C there is a basis B and an element $x \in X \setminus B$ with $C = C_x(B)$.

Consider any matroid $M(X)$ such that the property “ $A \subseteq X$ is independent” is decidable in polynomial time (w.r.t. $|X|$). (Now for a given dependent subset A a circuit $C \subseteq A$ can be found in poly-time, and for $B \subseteq X$ we can decide in poly-time whether B is a basis of $M(X)$ or not.) Assume furthermore that the corank of $M(X)$ is bounded by a constant k , that is $|X| - \text{rank}(X) \leq k$. Then by running through all subsets of X of size $\text{rank}(X)$ we can enumerate all bases B in polynomial time, and then considering all $x \in X \setminus B$ yields all circuits $C_x(B)$ of $M(X)$ (in poly-time).

B. Reduction by ”Davis-Putnam resolution”

For clauses C, D with exactly one “clashing literal” x , that is $C \cap \overline{D} = \{x\}$, let $C \diamond D := (C \setminus \{x\}) \cup (D \setminus \{\overline{x}\})$ denote their *resolvent*.

For $a, b \in \mathbb{N} \cup \{\infty\}$ and clause-sets $F, F' \in \mathcal{CLS}$ the relation $F \xrightarrow{a,b} F'$ holds if there is a variable v with

$$\#_v(F) \leq a, \#_{\overline{v}}(F) \leq b \text{ or } \#_v(F) \leq b, \#_{\overline{v}}(F) \leq a$$

such that F' is obtained from F by *Davis-Putnam resolution* on v without any “blockings” or *contractions*, that is for

$$P := \{(C, D) \in F^2 : v \in C \wedge \overline{v} \in D\}$$

the following conditions hold:

- $P \neq \emptyset$ (neither v or \overline{v} is pure)
- $(C, D) \in P \implies C \cap \overline{D} = \{v\}$ and $C \diamond D \notin F$ (all pairs $(C, D) \in P$ are resolvable, and none of their resolvents is already contained in F)
- $(C, D), (C', D') \in P$ and $(C, D) \neq (C', D') \implies C \diamond D \neq C' \diamond D'$ (any two different pairs from P yield different resolvents)

and we have

$$F' = \{C \in F : v \notin \text{var}(C)\} \cup \{C \diamond D : (C, D) \in P\}.$$

It is well known that F' and F are satisfiability equivalent. The following invariance properties are basic.

Lemma B.1 For any clause-sets F, F' with $F \xrightarrow{1, \infty} F'$ we have

1. $\delta(F) = \delta(F')$ (and $n(F') = n(F) - 1$)
2. $F \in \mathcal{MUSAT} \Leftrightarrow F' \in \mathcal{MUSAT}$
3. $F \in \mathcal{LEAN} \Leftrightarrow F' \in \mathcal{LEAN}$.

Proof: Part 1 and 2 follow directly from the definition of $\xrightarrow{1, \infty}$. For part 3 consider v and P according to the above definition and assume w.l.o.g. that $P = \{C\} \times \{D_1, \dots, D_m\}$.

If φ is a (non-trivial) autarky for F (satisfying at least one clause) then φ is also a (non-trivial) autarky for F' .

If on the other side φ is a non-trivial autarky for F' , then, assuming w.l.o.g. $v \notin \text{var}(\varphi)$, we construct a non-trivial autarky φ' for F as follows:

- if $\varphi(C) = 1$, then $\varphi' := \varphi \cup \langle \overline{v} \rightarrow 1 \rangle$;
- if $\text{var}(C) \cap \text{var}(\varphi) = \emptyset$, then $\varphi' := \varphi$;
- otherwise $\varphi(D_i) = 1$ for all D_i , and we set $\varphi' := \varphi \cup \langle v \rightarrow 1 \rangle$. ■

Note that if for some variable v , occurring in one sign at most once in F , the additional conditions on P in the definition of $\xrightarrow{1, \infty}$ are violated, then F is not minimally unsatisfiable.

The proof of Lemma 3.10 uses the following compatibility properties (here the proofs are somewhat technical, and the reader may skip them).

Lemma B.2 To unify notations, for clause-sets F, F' the relation $F \xrightarrow{\text{del}} F'$ holds iff there is a clause $C \in F$ with $F' = F \setminus \{C\}$. (As usual, by “ $\xrightarrow{\text{del}}$ ” we denote the reflexive-transitive closure of the relation “ \rightarrow ”.)

For all clause-sets F, F_1, F_2 and partial assignments φ we have

1. $F \xrightarrow{\text{del}} F_1 \implies \varphi * F \xrightarrow{\text{del}} \varphi * F_1$.
2. If $\text{var}(\varphi) \subseteq \text{var}(F_1)$, then there is $F' \in \mathcal{CLS}$ with $F \xrightarrow{1, \infty} F_1 \implies \varphi * F \xrightarrow{\text{del}} F' \xrightarrow{1, \infty} \varphi * F_1$.
3. There is $F' \in \mathcal{CLS}$ with $F \xrightarrow{1, \infty} F_1 \xrightarrow{\text{del}} F_2 \implies F' \xrightarrow{\text{del}} F' \xrightarrow{1, \infty} F_2$.

Proof: We only have to consider $\varphi = \langle x \rightarrow 1 \rangle$ for some literal x , and also only single applications of the reductions.

Part 1: For $F_1 = F \setminus \{C_0\}$ either $\langle x \rightarrow 1 \rangle * F = \langle x \rightarrow 1 \rangle * (F \setminus \{C_0\})$ holds or

$$\langle \langle x \rightarrow 1 \rangle * F \rangle \setminus \{C_0 \setminus \{\bar{x}\}\} = \langle x \rightarrow 1 \rangle * (F \setminus \{C_0\}).$$

For parts 2 and 3 let

$$F_1 = \{C \in F : v \notin \text{var}(C)\} \cup \{C \diamond D : (C, D) \in P\}$$

according to the definition of $\xrightarrow{1, \infty}$ from the beginning of this section. We have to take care that also for the modified application of $\xrightarrow{1, \infty}$ the three conditions on P (applied accordingly) are fulfilled. W.l.o.g. we may assume $\#_v(F) = 1$. So there is exactly one $C_0 \in F$ with $v \in C_0$.

Part 2: Since $\text{var}(x) \in \text{var}(F_1)$, we have $\text{var}(x) \neq v$. Let

$$P' := \{(C, D) \in (\langle x \rightarrow 1 \rangle * F)^2 : v \in C \wedge \bar{v} \in D\}.$$

If $P' = \emptyset$ then $\langle x \rightarrow 1 \rangle * F_1$ is equal to

$$\langle \langle x \rightarrow 1 \rangle * F \rangle \setminus \{C \in \langle x \rightarrow 1 \rangle * F : v \in \text{var}(C)\}.$$

So assume $P' \neq \emptyset$. Let

$$G := \left(\{C \diamond D : (C, D) \in P'\} \cap \langle x \rightarrow 1 \rangle * F \right) \cup \left\{ D \mid \exists D' : (C_0, D), (C_0, D') \in P' \wedge D \neq D' \wedge C_0 \diamond D = C_0 \diamond D' \wedge (C_0, D \cup \{\bar{x}\}) \in P \right\}.$$

Using Davis-Putnam resolution on v we get

$$\langle \langle x \rightarrow 1 \rangle * F \rangle \setminus G \xrightarrow{1, \infty} \langle x \rightarrow 1 \rangle * F_1.$$

Part 3: Let $F_2 = F_1 \setminus \{E\}$. In case of $E \in F$ we get $F \setminus \{E\} \xrightarrow{1, \infty} F_2$. Otherwise let $E = C_0 \diamond D$ for $(C_0, D) \in P$. In case of $\#_{\bar{v}}(F) = 1$ we have $F \setminus \{C_0, D\} = F_2$. So assume $\#_{\bar{v}}(F) \geq 2$. Now $F \setminus \{D\} \xrightarrow{1, \infty} F_2$ holds. ■

C. Minimally unsatisfiable clause-sets with deficiency $\delta = 1$

Here simple proofs of the characterisations of the classes $\mathcal{MUSAT}(1)$ and $\mathcal{SMUSAT}(1)$ obtained in [6] resp. [1] are given. The following “starting lemma” motivated the use of \mathcal{SMUSAT} in [7] (called “saturated minimally unsatisfiable clause-sets” there).

Lemma C.1 [7] *For any clause-set $F \in \mathcal{SMUSAT}$ and any literal x we have $\langle x \rightarrow 1 \rangle * F \in \mathcal{MUSAT}$.*

Proof: Consider $F' \subseteq \langle x \rightarrow 1 \rangle * F$, $F' \in \mathcal{MUSAT}$ and assume there is $C \in \langle x \rightarrow 1 \rangle * F$, $C \notin F'$. Since F is

minimally unsatisfiable we must have $C \in F$ (otherwise for $C_0 \in F$ with $C_0 \setminus \{\bar{x}\} = C$ both $\langle x \rightarrow 1 \rangle * (F \setminus \{C_0\})$ and $\langle x \rightarrow 0 \rangle * (F \setminus \{C_0\})$ would be unsatisfiable). But now replacing C in F by $C \cup \{x\}$ maintains unsatisfiability (since application of $\langle x \rightarrow 1 \rangle$ and $\langle x \rightarrow 0 \rangle$ still yield unsatisfiable clause-sets), contradicting the assumption that F is saturated. ■

As an easy application, in [7] an alternative proof (by induction on $n(F)$) is given for the fact $\delta(F) \geq 1$ if F is minimally unsatisfiable:

If $n(F) = 0$ then $F = \{\perp\}$ and thus $\delta(F) = 1$. Otherwise saturate F and obtain $F' \in \mathcal{SMUSAT}$. Choose a variable $v \in \text{var}(F)$ with $\#_v(F') + \#_{\bar{v}}(F')$ minimal. Now by induction hypothesis and the previous lemma we get $1 \leq \delta(\langle x \rightarrow 1 \rangle * F') \leq (c(F) - 1) - (n(F) - 1) = \delta(F') = \delta(F)$. ✓

For the special case $k = 1$ the next lemma has been shown in [6] (using a quite complicated proof).

Lemma C.2 *For any clause-set $F \in \mathcal{MUSAT}(k)$ with $F \neq \{\perp\}$ there exists a variable $v \in \text{var}(F)$ with $\#_v(F), \#_{\bar{v}}(F) \leq k$.*

Proof: Assume the contrary and saturate F , obtaining $F' \in \mathcal{SMUSAT}$. Consider a variable $v \in \text{var}(F)$ with $\#_v(F') + \#_{\bar{v}}(F')$ minimal. W.l.o.g. we assume $\#_v(F') \geq k + 1$. Now $\langle v \rightarrow 1 \rangle * F' \in \mathcal{MUSAT}$ and thus $\delta(\langle x \rightarrow 1 \rangle * F') \geq 1$, but $c(\langle x \rightarrow 1 \rangle * F') \leq c(F) - (k + 1)$ and $n(\langle x \rightarrow 1 \rangle * F') = n(F) - 1$. ■

Lemma C.2 together with Lemma B.1 yields

Lemma C.3 *For $F \in \mathcal{CLS}$ the following conditions are equivalent:*

1. $F \in \mathcal{MUSAT}(1)$
2. $F \xrightarrow{1, \infty}^* \{\perp\}$ (for some or any order of applications of $\xrightarrow{1, \infty}$)
3. $F \xrightarrow{1, 1}^* \{\perp\}$ (for some or any order of applications of $\xrightarrow{1, 1}$).

Thus “ $F \in \mathcal{MUSAT}(1)$?” is poly-time decidable. ■

To characterize $\mathcal{SMUSAT}(1)$, we refine the relation $F \xrightarrow{1, 1} F'$ by the restriction, that the two clauses $C, D \in F$ with $F' = (F \setminus \{C, D\}) \cup \{C \diamond D\}$ are as large as possible, that is, for $C \cap \bar{D} = \{x\}$ we have

$$C \setminus \{x\} = D \setminus \{\bar{x}\}. \quad (6)$$

Let “ $\xrightarrow{s1,1}$ ” denote this restricted form (“s” for “strong”). Easily we get for clause-sets F, F' with $F \xrightarrow{s1,1} F'$ that F is saturated minimally unsatisfiable iff F' is. Observing that in the situation of Lemma C.2 for $F \in \text{SMUSAT}(1)$ the two occurrences C, D of variable v must fulfill (6) (with $x = v$), we immediately get the next lemma.

Lemma C.4 For $F \in \text{CLS}$ we have

$$F \in \text{SMUSAT}(1) \Leftrightarrow F \xrightarrow{s1,1} \{\perp\}.$$

Since the order of applications of $\xrightarrow{s1,1}$ does not matter here, $\text{SMUSAT}(1)$ is poly-time decidable. ■

Lemma C.5 Consider a binary tree T , where each inner node w is labeled by a distinct variable v_w . We label the “left” outgoing edge of an inner node w by the literal v_w and the “right” outgoing edge by $\overline{v_w}$. For each node w let the clause C_w be the set of all literals along the path from the root to w . Now let

$$F(T) := \{C_w : w \text{ leaf of } T\}.$$

Furthermore we say, F' is derived from $F(T)$ by “literal elimination” if one can obtain F' from $F(T)$ by repeatedly crossing out literal occurrences without ever producing a pure literal (that is, for each inner node w the “left” (resp. “right”) subtree at w still has at least one clause containing v_w (resp. $\overline{v_w}$)). Now the following holds.

1. The class $\text{SMUSAT}(1)$ is the class of all $F(T)$ for T as above. (Thus any two different clauses in $F \in \text{SMUSAT}(1)$ “clash” (they contain at least one complementary literal pair), any clause of F contains at least one literal occurring only once in F , and there is a variable occurring in all clauses of F .)
2. The class $\text{MUSAT}(1)$ is the class of all F' derived from some $F(T)$ by “literal elimination.”

Proof: Part 1 follows from Lemma C.4. Analogously, part 2 follows from Lemma C.3 and Lemma B.1, part 2. (“Attacking” for both parts any inner node next to a leaf and applying $\xrightarrow{s1,1}$ resp. $\xrightarrow{1,1}$.) ■

Thus saturated minimally unsatisfiable clause-sets of deficiency 1 are exactly the formulas introduced in [5].

D. The proof of the Splitting theorem

Proof of Lemma 3.9: For part 1 let

$$\begin{aligned} V^* &:= \text{var}(F) \setminus V \\ F^* &:= (V^* * F) \setminus \{\perp\} \\ F' &:= \{C \in F : \text{var}(C) \cap V \neq \emptyset\}. \end{aligned}$$

By definition we have $c(F^*) \leq c(F')$ and $\text{var}(F^*) = V$. Lemma 3.7 gives $\delta(F^*) \geq 1$. If there is a variable $v \in V$ with $\#_v(F^*) + \#\overline{v}(F^*) \leq 2$ then in fact $c(F') \geq c(F^*) + 1$, and thus $c(F') - |V| \geq \delta(F^*) + 1 = 2$. Otherwise Corollary 3.8 and Lemma C.2 yield $c(F') - |V| \geq \delta(F^*) \geq 2$.

For part 2 let

$$\begin{aligned} V^* &:= \text{var}(F) \setminus \text{var}(F') \\ F^* &:= \{C \in F : \text{var}(C) \cap V^* \neq \emptyset\}. \end{aligned}$$

Because of $F^* \cap F' = \emptyset$ we have $c(F') \leq c(F) - c(F^*)$, while part 1 gives $c(F^*) \geq |V^*| + 2 = n(F) - n(F') + 2$, which altogether yield $\delta(F') \leq \delta(F) - 2$.

For part 3 consider $F' \subseteq \langle x \rightarrow 1 \rangle * F$. If $n(F') = n(F) - 1$, then $\delta(F') \leq (c(F) - 2) - (n(F) - 1) = \delta(F) - 1$.

Otherwise take $F'_0 \subseteq F$ with $c(F'_0) = c(F')$ and $\langle x \rightarrow 1 \rangle * F'_0 = F'$. Now $\text{var}(F'_0) \subset \text{var}(F)$, and thus by part 2: $\delta(F') \leq \delta(F'_0) + 1 \leq \delta(F) - 1$. ■

Proof of Lemma 3.10: Consider a clause-set F'' such that $F' \xrightarrow{1, \infty} F''$ and every variable occurs in F'' in both signs at least twice.

By Lemma 3.9, part 3, for any variable $v \in \text{var}(F'')$ and both $\varepsilon \in \{0, 1\}$ there is F''_ε with

$$\langle v \rightarrow \varepsilon \rangle * F'' \xrightarrow{\text{del}} F''_\varepsilon \in \text{MUSAT}(\delta(F') - 1)$$

(using the relation $\xrightarrow{\text{del}}$ from Lemma B.2). Altogether:

$$F \xrightarrow{\text{del}} F' \xrightarrow{1, \infty} F'' \rightarrow \langle v \rightarrow \varepsilon \rangle * F'' \xrightarrow{\text{del}} F''_\varepsilon.$$

By Lemma B.2, part 2 there is G with

$$\begin{aligned} F \xrightarrow{\text{del}} F' \rightarrow \langle v \rightarrow \varepsilon \rangle * F' \xrightarrow{\text{del}} G \\ \xrightarrow{1, \infty} \langle v \rightarrow \varepsilon \rangle * F'' \xrightarrow{\text{del}} F''_\varepsilon, \end{aligned}$$

and part 1 of this lemma now yields

$$\langle v \rightarrow \varepsilon \rangle * F' \xrightarrow{\text{del}} G \xrightarrow{1, \infty} \langle v \rightarrow \varepsilon \rangle * F'' \xrightarrow{\text{del}} F''_\varepsilon.$$

Finally by part 3 of that lemma there is F'_ε with

$$\langle v \rightarrow \varepsilon \rangle * F' \xrightarrow{\text{del}} F'_\varepsilon \xrightarrow{1, \infty} F''_\varepsilon,$$

and Lemma B.1 gives $F'_\varepsilon \in \text{MUSAT}(\delta(F') - 1)$. ■