Relative efficiency of Propositional Proof Systems

Simone Kilian

Supervised by Prof. Dr. Jacobo Torán
Institute for Theoretical Computer Science, Ulm University

Seminar Introduction to Algorithmics
June 7, 2022

1. Motivation

The question whether $\mathbb{NP} = co\mathbb{NP}$ is interesting and important both for theoretical reasons as well as for practical application. To illustrate why, consider a problem P in $co\mathbb{NP}$ and an input x for that problem. If the answer to x is no, i.e., $x \notin P$, we can guess a negative witness and verify in polynomial time that it is indeed a negative witness to our problem. If we would like to verify that the answer to x is yes, i.e., $x \in P$, then we have to show that no such negative witness exists. An exponential amount of such witnesses could exist. This poses a problem if we want to practically solve such problems, as we cannot simply search for one "correct" witness, but have to exclude all of the exponentially many counter witnesses.

Further, knowing whether $\mathbb{NP} = co\mathbb{NP}$ is true or not also has implications on the $\mathbb{P} - \mathbb{NP}$ problem. If $\mathbb{NP} \neq co\mathbb{NP}$, then we know that $\mathbb{NP}$ is not closed under complement. On the other hand, we know that $\mathbb{P}$ is closed under complementation, which would imply that $\mathbb{P}$ and $\mathbb{NP}$ cannot be equal. I.e. showing that $\mathbb{NP} \neq co\mathbb{NP}$ would simultaneously solve the $\mathbb{P} - \mathbb{NP}$ problem.

If we could prove that $\mathbb{NP} = co\mathbb{NP}$, we would also be able to guess a positive witness, similar to the negative one, since the problem is now also in $\mathbb{NP}$. This could lead to better algorithms for these type of problems. If we could prove $\mathbb{NP} \neq co\mathbb{NP}$, then we would know that trying to construct such an algorithm searching for a positive witness would be futile. Whether $\mathbb{NP} = co\mathbb{NP}$, remains a currently unsolved problem.

One way of approaching the problem is by considering one particular problem in co$\mathbb{NP}$ – in our case **TAUT**. **TAUT** is the problem of deciding whether a propositional logic formula φ is a tautology, i.e., whether all truth assignments satisfy φ. Since **TAUT** is in co$\mathbb{NP}$ we could also ask us, what would follow if we could show that **TAUT** is also in $\mathbb{NP}$. If this were the case, then $\mathbb{NP} = co\mathbb{NP}$ would be true.

To prove this, let's assume that **TAUT** is in $\mathbb{NP}$. Then every problem P in $\mathbb{NP}$ would be polynomial-time-reducible to **TAUT**. Thus a function f exists such that $x \in P \Leftrightarrow f(x) \in$ **TAUT**. If we now negate both sides, we get $x \in \overline{P} \Leftrightarrow f(x) \in \overline{\textbf{TAUT}}$. The function f would then be a reduction function for the complement problem of P if and only if $\overline{\textbf{TAUT}}$ is in $\mathbb{NP}$. To see that this is the case, we can guess a truth assignment that does not satisfy the formula – which is possible in polynomial time. This truth assignment would be a witness for the fact that a formula is a non-tautology, i.e., in $\overline{\textbf{TAUT}}$. Thus $\overline{P}$ is in $\mathbb{NP}$, which implies that $\mathbb{NP}$ is closed under complement and thus $\mathbb{NP} = co\mathbb{NP}$.

Thus the central question is whether **TAUT** is in $\mathbb{NP}$. This is the case, if there would be a positive witness for each tautological formula in TAUT and that this proof would only be polynomial in size and could be verified in polynomial time.

A reasonable guess for the structure of such positive witnesses is that they are proofs with a polynomial amount of symbols in them. As such, the reformulated question would be whether every propositional tautology has a short proof showing that it is in fact a tautology. In the following, we have a closer look on the paper "The Relative Efficiency Of Propositional Proof Systems" by Cook and Reckhow [2]. They take the problem **TAUT** and ask the question if we can polynomially bound the length of a shortest proof for a propositional tautology. If we would know that such short, i.e., polynomially bounded, proofs exist, then we could verify a given tautology with this proof (i.e., a positive witness) in polynomial time. This would imply that we have showed that **TAUT** is in $\mathbb{NP}$. Hence we would have showed that $\mathbb{NP} = co\mathbb{NP}$.

1.1 Preliminaries

In this seminar-paper we consider only propositional logic. Throughout this seminar-paper we are going to consider logics with an arbitrary set of connectives κ. One common choice could be $\kappa = \{\wedge, \vee, \rightarrow, \leftrightarrow, \neg\}$ Some of the possible sets of connectives do not make sense, e.g., the set $\kappa = \{\neg\}$, as with them we cannot express all possible tautologies of propositional logic. Thus, we restrict ourselves only to sets of connectives κ that are complete in the sense that one can express all possible binary boolean functions with them in a compact way. One such example would be to choose κ as only containing the NAND operator.

1.2 Measuring The Difficulty Of Proving Propositional Tautologies

In order to decide whether a propositional formula is a tautology, we are interested in different proof methods and calculi. It might be the case that there are no short proofs with one type of proof methods, but there are with another. Thus it is relevant to compare different proof methods with regard to the length of their proofs.

A propositional proof system is a function that takes a proof and returns a formula that that is proven by the proof. As such, this must be a true statement namely a tautological formula and therefore be in **TAUT**. Only proof systems that can proof every propositional tautology are useful, i.e., where there is for every tautology at least one proof for it. This implies that proof systems are surjective functions.

In order to define this more adequate we give a definition of a proof system.

Definition 1.1 (Proof System [3]). Let Σ^* be a set of strings and **TAUT** the set of propositional tautologies. A propositional proof system (pps) is a surjective polynomial time computable function $f : \Sigma^* \rightarrow \textbf{TAUT}$.

For the remainder of this paper, we consider only a very specific type of proofs as inputs for the proof systems. We call them line-based proofs. Next, we define the restrictions and conditions we pose on these proofs and introduce measure to express the size of such proofs.

Definition 1.2 (Proof [2]). A *proof* π in the propositional proof system is a sequence of formulae (lines) separated by commas. As a proof is a finite sequence of symbols we consider a proof to be a string.

For such types of proofs, we define $f : \Sigma^* \rightarrow \textbf{TAUT}$ as in Def. 1.1 as follows. Let A_0 be some fixed tautology.[1]

$$f(\pi) = \begin{cases} A & \text{if } \pi \text{ proves A} \\ A_0 & \text{if } \pi \text{ is a string that does not correspond to a proof in the system} \end{cases}$$

Let A be a formula (or a sequence of formulae) and π be a proof. We define the following length measures:
$l(A) :=$ numbers of atoms and constants as $\{\top, \bot\}$ in A
$\lambda(\pi) :=$ number of lines in a proof π
$\rho(\pi) := max_i \; \{l(A_i) \mid \pi = (A_1, \ldots, A_n)\}$ (i.e. $\rho(\pi)$ is the longest line in the proof considering $l(A_i)$)

In order to show whether any propositional proof systems have polynomially bounded proofs for every tautology, we can try to put them into equivalence classes. I.e. we want to determine whether the proofs of two different proof systems are always of roughly equivalent size. Hence, if we show a bound for one system in an equivalence class we would show it for all. We define a concept that we call *p-simulation* with which we might show that two proofs are in the same equivalence class.

Definition 1.3 (f_2 p-simulates f_1[2])**.** Let $L \subseteq \Sigma^*$, $f_1 : \Sigma_1^* \to L$ and $f_2 : \Sigma_2^* \to L$.
f_2 *p-simulates* f_1 iff there exists a polynomial time computable function $g : \Sigma_1^* \to \Sigma_2^*$ such that $f_2(g(x)) = f_1(x)$ for all $x \in \Sigma_1^*$

2. Frege Systems

Frege systems are a classical example for propositional proof systems. With them, we construct proofs showing that propositional formulae are tautologies. The base idea of Frege systems is to derive tautologies from other tautologies via deduction. The means by which such deductions happen are so-called Frege rules. A Frege rule is a scheme that allows to infer a new tautology (D) given that we have shown that some other formulae ($C_1, \ldots, C_n$) are already shown to be tautologies.

Definition 2.1 (Frege Rule[2])**.** A *Frege Rule* is a system of formulae

$$\frac{C_1, \ldots, C_n}{D}$$

where $C_1, \ldots, C_n \models D$.
In the case $n = 0$ we call the Frege Rule an *axiom scheme.* If $n > 0$ we call the Frege Rules *scheme rules.*

Example 2.2. *Modus Ponens* is a scheme rule for $n = 2$.

$$\frac{A, (A \to B)}{B}$$

[1] This is necessary as proof systems must be defined over Σ^* and we thus have to output a tautology that is proven even if the proof is not well formatted or wrong.

Applying Frege rules only with the variables they are defined with is relatively useless. For example we could need to apply the modus ponens rule with A and B replaced by any arbitrary formulae. This notion is formalised by substitutions that can be applied to Frege rules.

Definition 2.3 (Substitution [2])**.** Let $D_1, \dots, D_k$ be formulae and $P_1, \dots, P_k$ distinct atoms. Then

$$\sigma = [(P_1, \dots, P_k)/(D_1, \dots, D_k)]$$

is a *Substitution.* When we apply a substitution we replace the atoms P_i by the formulae D_i. After applying the substitution σ on the formula A we write σA for the result.

Let σ be a substitution. If we apply σ on $C_1, \dots, C_n$ we write $\sigma C_1, \dots, \sigma C_n$. We say that *$\sigma D$ follows from $\sigma C_1, \dots, \sigma C_n$* by the rule

$$\frac{C_1, \dots, C_n}{D}$$

If we perform proofs using a Frege System, we always use a given set of Frege rules with which we can derive new tautologies. This can be formalised as follows.

Definition 2.4. Let $\mathcal{F}$ be a set of *Frege Rules* and let further be π be a proof, i.e., a sequence of formulae. The proof is correct iff for every formula φ in π there is a Frege Rule $\frac{C_1, \dots, C_n}{D}$ in $\mathcal{F}$ and a substitution σ such that $\sigma(D) = \varphi$ and for every C_i there is a line ψ before φ in the proof π such that $\psi = \sigma(C_i)$. The statement proven by π is its last line.

A Frege proof is thus a sequence of tautologies that can be derived. For every such tautology, there has to be a Frege rule that allows us to infer the tautology. We perform this inference by applying a substitution to the rule. Then the substituted preconditions or antecedents of the rule have to already be tautologies, i.e., they have to be part of the proof already. Then the substituted consequence of the rule is also a tautology and can be a new line of the proof. The above definition deals only with deriving tautologies. If we want to derive logical consequences of a set of formulae $A_1, \dots, A_n$, then we also allow that a precondition of a Frege rule is satisfied by $\sigma(C_i) = A_j$ for some A_j.

Definition 2.5 (Inference system $\mathcal{F}$ [2, 3])**.** An *Inference System* $\mathcal{F}$ is a finite set of *Frege Rules (2.1).* That means it that consists of a finite set of *axiom schemes* and of a finite set of *scheme rules* of rules of inference (e.g. *modus ponens*).
We call $\mathcal{F}$ implicationally complete iff $A_1, \dots, A_n \models B$ implies $A_1, \dots, A_n \vdash_{\mathcal{F}} B$.

Implicational completeness is an important property for Frege systems. It states that every tautology can actually be derived using the system. Non-complete Frege systems are not of interest, as for them there are tautologies that cannot be derived.

Definition 2.6 (Frege System [2])**.** A *Frege System* is an *Inference System $\mathcal{F}$ (2.5)* that is implicationally complete (and sound). We will denote *Frege Systems* with $\mathcal{F}$ as well.

Thus a *Frege System* is a propositional proof system. To illustrate how Frege Systems operate, we give both an example for a Frege System as well as a derivation of a tautology in such a system.

Example 2.7 (Frege's system as a Frege System[2])**.** We use the logical connectives $\kappa = \{\neg, \to\}$ and as a scheme rule *modus ponens*

$$\frac{A, A \to B}{B} \tag{1}$$

That means if we have formulae of the form $A \to B$ and if we show that A holds we know that also B holds. And we could insert B as a new line to the proof.
As axiom schemes we use the following:

$$\frac{}{A \to (B \to A)} \tag{2}$$

$$\frac{}{(C \to (B \to A)) \to ((C \to B) \to (C \to A))} \tag{3}$$

$$\frac{}{(D \to (B \to A)) \to (B \to (D \to A))} \tag{4}$$

$$\frac{}{(B \to A) \to (\neg A \to \neg B)} \tag{5}$$

$$\frac{}{\neg\neg A \to A} \tag{6}$$

$$\frac{}{A \to \neg\neg A} \tag{7}$$

(We remark that the line above the axiom schemes is not the negation but the case $n = 0$ from (2.1).)

Example 2.8. We use the Frege system above and like to show that
$E \to (C \to \neg\neg C) \in \mathbf{TAUT}$.

Proof. From axiom scheme $\frac{}{A \to \neg\neg A}$ (7) we can derive under the substitution $\sigma = [A/C]$ that $C \to \neg\neg C$ is a tautology. If we substitute $\phi = [(A, B)\ /\ ((C \to \neg\neg C), E)]$ we can use the axiom scheme $\frac{}{A \to (B \to A)}$ (2) to proof that $(C \to \neg\neg C) \to (E \to (C \to \neg\neg C))$ is a tautology. To complete the proof we apply the modus-ponens scheme rule

$$\frac{A, A \to B}{B}$$

with the substitution $\tau = [(A, B)\ /\ ((C \to \neg\neg C), (E \to (C \to \neg\neg C)))]$:

$$\frac{(C \to \neg\neg C), (C \to \neg\neg C) \to (E \to (C \to \neg\neg C))}{(E \to (C \to \neg\neg C))}$$

We showed that τB follows from τA, $\tau(A \to B)$. Since we have proven that τA and
$\tau(A \to B)$ are tautologies it follows by modus ponens that $E \to (C \to \neg\neg C) \in \mathbf{TAUT}$. □

Example 2.9. Example (2.8) as proof in the sense of definition (1.2) would be written as follows:

Proof.

$$C \to \neg\neg C, \quad (C \to \neg\neg C) \to (E \to (C \to \neg\neg C)), \quad (E \to (C \to \neg\neg C))$$

□

Here $\pi := C \to \neg\neg C,\ (C \to \neg\neg C) \to (E \to (C \to \neg\neg C)),\ (E \to (C \to \neg\neg C))$ and $f(\pi) = E \to (C \to \neg\neg C)$.
How can we see now that π proves $E \to (C \to \neg\neg C)$?
The last line must be derived by the lines before by an axiom scheme or scheme rule. Hence, we can test in polynomial time if some *Frege Rule* fits under substitution. The same principle is applicable on every line of the prove only considering the lines before the current. Thus we can verify in polynomial time if π proofs $E \to (C \to \neg\neg C)$.

The following Lemma will be useful in some of the following proofs. Its core statement is that we can apply substitutions to a proof as a whole, including its preconditions and the proven statements – and that the resulting substituted proof is still valid and now proofs the substituted proven statement σB.

Lemma 2.10. *(cf. [2]) If π is a derivation of B from $A_1, \ldots, A_n$ in a* Frege System *$\mathcal{F}$, then $\sigma(\pi)$ is a derivation of σB from $\sigma A_1, \ldots, \sigma A_n$ in $\mathcal{F}$, for any substitution σ.*

Proof. This follows by induction on the length of π. □

At this point, it is reasonable to believe that the choice of the Frege rules we employ to prove a statement influence the length of its proof. While this is the case, the next theorem shows that this influence is not significant. In particular, for every proof in one Frege System there is always a proof in any other Frege System with roughly equal length – in the sense that the length of the two proofs only differs by a constant factor. This implies that if we consider shortest proofs in any two Frege Systems, they differ in length by at most a constant factor. As such, the choice of the concrete Frege System does not influence the question whether short proofs exist or not. If they exist for any Frege System they exist for all others as well. To make this statement even clearer we can translate it into the contraposition: If we can show that there is one Frege system where all proofs for one theorem are not short then for all Frege Systems there is a theorem where all proofs for it are not short (and this theorem would be the same).

Theorem 2.11 (Existence of a function that translates proofs from one system to another [2])**.** *Let $\mathcal{F}_1, \mathcal{F}_2$ be two Frege Systems over κ:*
There exists a polynomial time computable function f and a constant $c \in \mathbb{N}$ such that for all formulae $A_1, \ldots, A_n,\ B$ and proofs π:

$$A_1, \ldots, A_n \vdash^{\pi}_{\mathcal{F}_1} B \quad \textit{implies} \quad A_1, \ldots, A_n \vdash^{f(\pi)}_{\mathcal{F}_2} B$$

with the property that $\lambda(f(\pi)) \leq c\lambda(\pi)$ and $\rho(f(\pi)) \leq c\rho(\pi)$.

Proof. (cf. [2]) Let $\mathcal{F}_1, \mathcal{F}_2$ be Frege systems over some logical connectives set κ. Since a Frege Rule in one Frege System $\mathcal{F}_1$ is correct and the other Frege System $\mathcal{F}_2$ is complete we are able to derive its implication as a proof in $\mathcal{F}_2$. Thus, for each Frege Rule $R := \frac{C_1,\ldots,C_n}{D}$ in $\mathcal{F}_1$ let π_R be a derivation of D from $C_1, \ldots, C_n$ in $\mathcal{F}_2$.

Suppose π is a derivation of B_k from $A_1, \ldots, A_n$ in $\mathcal{F}_1$ and $\pi = (B_1, \ldots, B_k)$.
We construct the $\mathcal{F}_2$ derivation $f(\pi)$ from π: Consider that B_i follows from B_a with $1 \leq a \leq i$ by the rule R_i of $\mathcal{F}_1$ and substitution σ_i.
In F_2 there is the derivation π_{R_i}. If we apply σ_i to this derivation, we obtain a derivation of B_i from B_a in F_2 (as in 2.10). The correctness of this statement can be proven by induction over the length of π_{R_i}. Replace B_i by the derivation $\sigma_i(\pi_{R_i})$ (with deleted hypothesis). $\sigma_i(\pi_{R_i})$ is a derivation of B_i from the same earlier B_a's.

Let $c \in \mathbb{N}$ be the number of lines that must be included in the longest derivation in π_R in $\mathcal{F}_2$, then it holds that $\lambda(f(\pi)) \leq c\lambda(\pi)$.
And let $d \in \mathbb{N}$ be an upper bound on $l(A)$ with A ranges over all formulae in all derivations π_R, such that $\rho(f(\pi)) \leq d\rho(\pi)$. □

The following corollary follows from the proof above and formalised the intuitive notion given before the theorem.

Corollary 2.12 (Equivalence on Frege Systems [2])**.** *Let f_1, f_2 be the pps-function (1.2) and $\mathcal{F}_1, \mathcal{F}_2$ be two Frege systems over κ. It holds that f_1 p-simulates f_2.*

As an immediate consequence of (Thm. 2.12) it follows from the equivalence of Frege Systems a stronger property as the implication above:

Corollary 2.13. *There exists a Frege system over κ that is polynomially bounded iff all Frege systems κ are polynomially bounded.*

2.1 Pigeon-Hole Principle

We take now a deeper look on the question if Frege Systems are good suited to leed short propositional proofs. As an example we try to simulate a proof in a Frege System. For this purpose we use an informal proof of the pigeon-hole principle.

The *Pigeon-Hole Principle* is whether n pigeons could be put into $(n-1)$ holes by placing only one pigeon in one hole. As we see this will not work. This is a famous example of a non-injective function on finite sets.
An injective function has the following property.

Definition 2.14 (Injective Function)**.** Let $f : A \to B$ be a function. f is called injective iff it holds for all a_1, a_2 that $f(a_1) = f(a_2)$ implies $a_1 = a_2$

A function is not injective if the cardinality of set A is greater than the cardinality of set B. Thus, some $x \in im(f)$ has more than one inverse image $f^{-1}(x)$ as f is a function and maps all elements of A to B. For the pigeon-hole principle we consider only finite sets.

Theorem 2.15 (Pigeon-Hole Principle)**.** *Let $f : A \to B$ and A, B finite sets with $|A| > |B|$. There exists no injective function that maps from A to B.*

In the following we consider only the case, that $|A| = n$ and $|B| = n - 1$.

Proof. (cf. [2]) We proceed an informal proof via induction on n.
It clearly holds that for $n = 2$ f is not an injective function.
IH: For $n \in \mathbb{N}$ it holds that there is no injective function from $\{1, \ldots, n\} \to \{1, \ldots, (n-1)\}$
Instead of showing that if it holds for $\{1, \ldots, (n-1)\} \to \{1, \ldots, (n-2)\}$ it also holds for $\{1, \ldots, n\} \to \{1, \ldots, (n-1)\}$ we will do the induction step by a proof by contradiction:
Suppose that there exists an injective function $f : \{1, \ldots, n\} \to \{1, \ldots, (n-1)\}$.
Based on f we construct $f' : \{1, \ldots (n-1)\} \to \{1, \ldots (n-2)\}$ that is injective. Let us define f' with

$$f'(i) = \begin{cases} f(i) & , f(i) \neq (n-1) \\ f(n) & , otherwise \end{cases}$$

f' is injective by construction. This contradicts the induction hypothesis and thus it follows that f was not injective. □

Example 2.16. We formulate a propositional formula φ that is satisfiable iff a function $f : A \to B$ with $|A| = 3$ and $|B| = 2$ is injective. $P_{i,j}$ be an atom with the meaning that pigeon i is in hole j for $1 \leq i \leq 3$ and $1 \leq j \leq 2$. Thus, we obtain the following formula:

$$\begin{aligned} \varphi := \; & (P_{1,1} \lor P_{1,2}) \land (P_{2,1} \lor P_{2,2}) \land (P_{3,1} \lor P_{3,2}) \land \\ & \neg(P_{1,1} \land P_{2,1}) \land \neg(P_{1,1} \land P_{3,1}) \land \neg(P_{2,1} \land P_{3,1}) \land \\ & \neg(P_{1,2} \land P_{2,2}) \land \neg(P_{1,2} \land P_{3,2}) \land \neg(P_{2,2} \land P_{3,2}) \\ \equiv \; & (P_{1,1} \lor P_{1,2}) \land (P_{2,1} \lor P_{2,2}) \land (P_{3,1} \lor P_{3,2}) \land \\ & (\neg P_{1,1} \lor \neg P_{2,1}) \land (\neg P_{1,1} \lor \neg P_{3,1}) \land (\neg P_{2,1} \lor \neg P_{3,1}) \land \\ & (\neg P_{1,2} \lor \neg P_{2,2}) \land (\neg P_{1,2} \lor \neg P_{3,2}) \land (\neg P_{2,2} \lor \neg P_{3,2}) \end{aligned}$$

In the following we will use Frege Systems to show that φ is unsatifiable. This would imply that f is not injective. φ is unsatisfiable iff there exists no truth assignment that satisfies φ iff every truth assignment is a model for $\neg\varphi$ iff $\neg\varphi$ is a tautology. Thus, if we can prove using a Frege System that $\neg\varphi$ is a tautology, f is not injective.

We generalize this example and formulate the formal proof for (2.15). When we translate this proof into a formal Frege System proof, we will see that its length is not polynomial in n.

Proof. (cf. [2]) Let $f : A \to B$, with $A = \{1, \ldots, n\}$ and $B = \{1, \ldots, (n-1)\}$ and $P_{i,j}$ be an atom with the meaning that pigeon i is in hole j for $1 \leq i \leq n$ and $1 \leq j \leq (n-1)$. Let $\mathcal{S}_n := \{P_{i,1} \lor \ \ldots \ \lor P_{i,(n-1)} \mid 1 \leq i \leq n\} \cup \{\neg P_{k,j} \lor \neg P_{l,j} \mid 1 \leq k \leq l \leq n, 1 \leq j \leq (n-1)\}$ is the set of formulae or the conjunction of them.
$\mathcal{S}_n$ is satisfiable iff there exists one truth assignment that satisfies all formulae $\rho \in \mathcal{S}_n$. Hence, there exists an injective function from $\{1, \ldots, n\}$ to $\{1, \ldots (n-1)\}$. $\mathcal{S}_n$ is unsatisfiable iff the formula $\neg\mathcal{S}_n$ is a tautology.

We are going to mimic the proof (2.1) in a more formal way. Suppose for a contradiction that $\mathcal{S}_n$ is satisfiable. For each i, j we introduce a formula $B_{i,j}$ which means $f'(i) = j$ such that $B_{i,j} = P_{i,j} \vee (P_{i,n-1} \wedge P_{n,j})$ with $1 \leq i \leq (n-1)$, $1 \leq j \leq (n-2)$. This simulates the definition of the function f' in (2.1).
We apply now this definition by a substitution on $\mathcal{S}_n$. Let σ_{n-1} be this substitution with $[P_{i,j}/B_{i,j}]$ with $1 \leq i \leq (n-1)$, $1 \leq j \leq (n-2)$.
The argument that f is injective implies f' to be injective translates to $\mathcal{S}_n \models \sigma_{n-1}\mathcal{S}_{n-1}$. By the completeness of Frege Systems it follows that $\mathcal{S}_n \vdash \sigma_{n-1}\mathcal{S}_{n-1}$. And by the same argumentation it follows $\mathcal{S}_{n-1} \vdash \sigma_{n-2}\mathcal{S}_{n-2}$. With Lemma 2.10 there is a derivation of the same number of lines, that shows $\sigma_{n-1}\mathcal{S}_{n-1} \vdash \sigma_{n-1}\sigma_{n-2}\mathcal{S}_{n-2}$. Thus, we obtain $\mathcal{S}_n \vdash \sigma_{n-1}\sigma_{n-2}\mathcal{S}_{n-2}$.
If we continue proceeding like this we finally get $\mathcal{S}_n \vdash \sigma_{n-1}\sigma_{n-2}\ldots\sigma_2\mathcal{S}_2$. But the case $\mathcal{S}_2$ is easy $\mathcal{S}_2 = \{P_{1,1}, P_{2,1}, \neg P_{1,1} \vee \neg P_{2,1}\} \equiv P_{1,1} \wedge P_{2,1} \wedge (\neg P_{1,1} \vee \neg P_{2,1}) =: \varphi$ and if we leave out the substitutions, short. In order to show that this formula is unsatisfialbe we will show that φ is a tautology by using Frege Rules. For

$$\begin{aligned} \neg\varphi &= \neg(P_{1,1} \wedge P_{2,1} \wedge (\neg P_{1,1} \vee \neg P_{2,1}) \\ &\equiv (P_{1,1} \wedge P_{2,1}) \rightarrow (P_{1,1} \wedge P_{2,1}) \end{aligned}$$

this would be easy to see. But as we are using Frege Systems to proof that this is a tautology, we consider $\mathcal{S}_2 = \{P_{1,1}, P_{2,1}, \neg P_{1,1} \vee \neg P_{2,1}\}$ as $P_{1,1} \rightarrow (P_{2,1} \rightarrow (\neg(P_{1,1} \rightarrow \neg P_{2,1})))$.
This suffices as if it follows by induction on the length of π that if π is a derivation of $\mathcal{S}_2$ from $\mathcal{S}_n$ then $\sigma_{n-1}\sigma_{n-2}\ldots\sigma_2\pi$ is a derivation from $\sigma_{n-1}\sigma_{n-2}\ldots\sigma_2\mathcal{S}_2$ from $\sigma_{n-1}\sigma_{n-2}\ldots\sigma_2\mathcal{S}_2\mathcal{S}_n$. Hence we showed that $\vdash \neg\mathcal{S}_n$.

□

Corollary 2.17. *(cf. [2]) The derivation of $\sigma_{n-1}\mathcal{S}_{n-1}$ from $\mathcal{S}_n$ has $\mathcal{O}(n^3)$ lines. Thus, the entire proof of $\neg S_n$ has $\mathcal{O}(n^4) = \mathcal{O}(|\neg\mathcal{S}_n|^{\frac{4}{3}})$ where $|\neg\mathcal{S}_n| = n^3$ lines.*
But each application of a substitution σ_i triples the length of a formula. This yields for the longest formulae in the proof of $\neg\mathcal{S}_n$ that they grow exponentially in n.

We have thus seen that the formalisation of the intuitive proof of the Pigeon-Hole Principle leads to a long, i.e., non-polynomial proof. This points us to assume that Frege Systems might not be the best means for proving tautologies in the sense that they can actually have long proofs. Note that this example is not a proof that Frege Systems do not have short proof or equivalently that the proof of the Pigeon-Hole principle is necessarily long in a Frege System – it is just evidence that this might be the case and that we may want to consider stronger proof systems. In fact, it is possible to prove the Pigeon-Hole principle using Frege Systems with short proofs [1].

2.2 Extended Frege Systems

The fact that the intuitive proof of the Pigeon-Hole principle is long begs the question whether the *Frege System* could be adjusted in order to reduce formula length. We will call this adjusted version *Extended Frege System*. Its core idea is that we are allowed to introduced abbreviations for long formulae by introducing new atoms and defining them appropriately. This should allow to compact proofs and especially the Pigeon-Hole proof. It is formalized as follows.

Definition 2.18 (Extension Rule (cf. [2])). Let A be a formula over some connective set κ and P a new atom.
The *Extension Rule* allows us to add *defining formulae* as $P \equiv A$ to the derivation. We say P is a *defined* atom.
P has the following properties:

(i) P must not occur in A

(ii) P must not occur in any lines preceding $P \equiv A$ in the derivation or proof

(iii) P must not occur in any hypothesis to the derivation

(iv) P may occur or be used in subsequent lines

(v) P must not occur in the last line of the derivation or proof

In the case that $\equiv \notin \kappa$ we have to show first how to represent *semantic equivalence* with only the connectives from κ.
For this purpose we introduce the function $\sim$ that returns an arbitrary but short formula $\sim (P, Q)$ over κ that is equivalent to $P \equiv Q$. Let now $\sim (P, A)$, or in infix notation $P \sim A$, be the *defining formula* for P.

Remember that we consider only sets κ that are complete in the sense that we can express all possible binary boolean functions with them in a compact way. Thus, the existence of a short formula that $\sim$ could return follows.

Example 2.19. We take a closer look at our example (2.9) where we used the tautology $C \to \neg\neg C$ as a line. Note that we used this formula multiple times in the other lines as well. For our example, we ignore the last condition (v). Let now P be a defining atom.
We defined κ as $\kappa = \{\neg, \to\}$. Since $\equiv \notin \kappa$ holds we have to consider $\sim$. For κ, $P \sim Q$ is $\neg((P \to Q) \to \neg(Q \to P))$, which is semantically equivalent to $(P \to Q) \wedge (Q \to P)$ and thus to $P \equiv Q$. Thus, we add $\neg((P \to (C \to \neg\neg C)) \to \neg((C \to \neg\neg C) \to P))$ to the lines of the derivation.
We now consider the case that $\equiv \in \kappa$:
With the *Extension Rule* we could add $P \equiv (C \to \neg\neg C)$ as a new line to the derivation.
More informally: We could say we would define $P := (C \to \neg\neg C)$. Hence, we would now use P instead of $C \to \neg\neg C$ in the derivation. We could also interpret this as a substitution $[P/(C \to \neg\neg C)]$ that we implicitly apply to the subsequent lines.

Using the notion of an extension rule, we can now formally define what an extended Frege System is.

Definition 2.20 (Extended Frege System (cf. [2])). An *Extendended Frege System* $e\mathcal{F}$ over a connective set κ is a *Frege System* (2.6) over κ and an *Extension Rule*.

Before we show the Soundness and Completeness of Extended Frege Systems we give a short example on how the Extension Rule has impact on the formula length considering the Pigeon-Hole Principle.

Example 2.21. If we consider the atoms $B_{i,j}$, we had the problem that we had to "replace" them by the formula $P_{i,j} \vee (P_{i,n-1} \wedge P_{n,j}$ from **??**. Since we had to do this recursively, the size of this expression grows exponentially. With Extended Frege Systems we solve this problem by applying Extension Rules instead of replacing the $B_{i,j}$ with the subformula. This will lead to more, but overall shorter lines.

First of all, we need to show that Extended Frege Systems are sound and complete. Completeness follows directly from the fact that every Frege System proof is also a valid proof in an Extended Frege System. We show soundness in the next theorem.

Proposition 2.22 (Soundness of $e\mathcal{F}$ [2])**.** *If $A_1, \ldots, A_n \vdash_{e\mathcal{F}} B$, then $A_1, \ldots, A_n \models B$.*

Proof. (cf. [2]) Let $\mathcal{A}$ be a truth assignment to the atoms of $A_1, \ldots, A_n$ that satisfies B. Then $\mathcal{A}$ can be extended to make every line in the derivation true. Particularly, if $P \equiv A$ is a *defining formula* implies that P has not occurred earlier in the derivation. Thus, we are free to extend $\mathcal{A}$ such that $\mathcal{A}(P) = \mathcal{A}(A)$ □

Next we consider the relation between proofs in Extended Frege Systems and regular Frege Systems. To start, we look at the size of these proofs in terms of the number of lines in their respective proofs. The following theorem states that Extended Frege Systems do not provide a significant reduction in the number of lines, as for every Extended Frege Proof there is a pure Frege Proof that is only slightly (namely by cm lines) longer.

Proposition 2.23. *(cf. [2]) If π is a derivation of B from $A_1, \ldots, A_n$ in $e\mathcal{F}$, then there is a derivation π' of B from $A_1, \ldots, A_n$ in $\mathcal{F}$ with the following property considering the number of lines: $\lambda(\pi') \leq \lambda(\pi) + cm$ where c depends only on $\mathcal{F}$ and m is the number of defining formulae in π.*

Proof. (cf. [2]) Suppose $P_i \sim C_i$ with $1 \leq i \leq m$ are the defining formulae in π are the *defining formulae* in π. The index refers to the order in which they occur in π. $P_1 \sim C_1, \ldots, P_m \sim C_m$ are not allowed in $\mathcal{F}$-proofs. Thus, we have to add them to the preconditions: π is a derivation in $\mathcal{F}$ of B from $A_1, \ldots, A_n, P_1 \sim C_1, \ldots, P_m \sim C_m$.
In the lines of the π some defining formulae might have been used. Hence, we have to substitute them. Let σ be the composed substitution

$$[P_m/C_m] \circ [P_{m-1}/C_{m-1}] \circ \ldots [P_1/C_1]$$

Thus, we obtain $\sigma(\pi)$. The preconditions of the proof in $\mathcal{F}$ still include the defining formulae. Hence we have to substitute P_i.
By Lemma 2.10 $\sigma(\pi)$ is a derivation of σB from $\sigma A_1, \ldots, \sigma A_n, \sigma(P_1 \sim C_1), \ldots, \sigma(P_m \sim C_m)$. The defining formulae $P_j \sim C_j$ could not have been added before A_i and not before $P_i \sim C_i$ for $1 < i < j$. The substitution σ does not affect them. Thus, it follows by the restrictions on the *defined atoms* P_i (see 2.18) that $\sigma(\pi)$ is a derivation in $\mathcal{F}$ of B from $A_1, \ldots, A_n, \sigma(C_1 \sim C_1), \ldots, \sigma(C_m \sim C_m)$ and even $A_1, \ldots, A_n, (C_1 \sim C_1), \ldots, (C_m \sim C_m)$. Still it is not possible that $\sigma(C_1 \sim C_1), \ldots, \sigma(C_m \sim C_m)$ remains in the preconditions since they were not part of the preconditions of π. Consider now $D \sim D$ has some fixed proof in $\mathcal{F}$ of say c lines. By Lemma 2.10, each $\sigma C_i \sim \sigma C_i$ has a proof in $\mathcal{F}$ of c lines. Thus, $\lambda(\sigma(\pi)) = \lambda(\pi)$. Hence we construct π' from $\sigma(\pi)$ together with this m substituted equivalence proofs. And we obtain $\lambda(\pi') \leq \lambda(\pi) + cm$. □

We have seen by the pigeon-hole principle that - also if the formulae of π are short - the formulae of π' can grow exponentially in m.

The following theorem shows a connection between two measures of size for proofs in Extended Frege Systems. It states that if we can bound the number of lines of a proof, then this also bounds the length of the longest line in that proof and consequently also the overall size of the proof in terms of number of characters. More formally, if we can bound the number of lines in a proof by some function $L(n)$ where n is the size of the tautology to prove, then for any Extended Frege System, the number of lines of a proof is also bounded by a constant multiple of this function and the maximum size of every line in the proof is bounded by a constant multiple of the length of the tautology to be proven. This property is in contrast to the pure Frege Systems where we have seen with the example of the Pigeon-Hole principle that proofs with small (i.e., polynomial) number of lines can have extremely and especially exponentially long lines. As such, this property makes Extended Frege System a interesting starting point for investigating whether all tautologies have shorts proofs.

Theorem 2.24 (An upper bound for shortest proofs in Extended Frege Systems (cf. [2]))**.** *Suppose that $e\mathcal{F}$ and $e\mathcal{F}'$ are* Extended Frege Systems *over some connective sets κ and κ' and suppose that $L(n) \geq n$ be a natural number function such that every tautology A over κ has a proof π in $e\mathcal{F}$ with the property $\lambda(\pi) \leq L(l(A))$, i.e., that $L(l(A))$ is an upper bound for the length of a shortest proof in terms of number of lines.*
Then every tautology A' over κ' has a proof π' in $e\mathcal{F}'$ such that $\lambda(\pi') \leq cL(cl(A'))$ and $\rho(\pi') \leq cl(A')$, where the constant c depends only on $\mathcal{F}$ and $\mathcal{F}'$.

Theorem 2.25 (An upper bound that refers to [4] in [2])**.** *For any* Extended Frege System *$e\mathcal{F}$ and tautology A holds: if π is a proof of A in $e\mathcal{F}$, then there is a proof π' of A in $e\mathcal{F}$ such that $\lambda(\pi') \leq c(\lambda(\pi) + l(A))$ and $\rho(\pi') \leq cl(A)$, where the constant c depends only on $\mathcal{F}$.*

Corollary 2.26. *(cf. [2]) A given* Extended Frege System *$e\mathcal{F}$ is polynomially bounded iff all* Extended Frege Systems *over all connective sets κ are polynomially bounded.*
An Extended Frege System *$e\mathcal{F}$ is polynomially bounded iff there is a polynomial bound on the number of lines in proofs in $e\mathcal{F}$.*

In the following we will proof the theorems 2.24 and 2.25. For this purpose we will need some results that we provide now.

Lemma 2.27. *(cf. [2]) Suppose $e\mathcal{F}$ and $L(n)$ satisfy the hypothesis of 2.24. If $A_1, \ldots, A_m, B$ are formulae over some connective set κ such that $A_1, \ldots, A_m \models B$, then there is a derivation π in $e\mathcal{F}$ of B from $A_1, \ldots, A_m$ with the property that $\lambda(\pi) \leq cL(cn)$, where $n = l(A_1) + \cdots + l(A_m) + l(B)$, and c depends only on $\mathcal{F}$.*

Proof. (cf. [2]) Suppose that $\vee, \neg \in \kappa$ of $\mathcal{F}$. Since $A_1, \ldots, A_m \models B$ holds, $(A_1 \wedge \cdots \wedge A_n) \rightarrow B$ is tautological. In order to use *Modus Ponens* in a recursive way, we can write
$\models (A_1 \rightarrow (A_2 \rightarrow \ldots (A_n \rightarrow B) \ldots))$ or $\models (\neg A_1 \vee (\neg A_2 \vee \ldots (\neg A_n \vee B) \ldots))$ instead.
Hence this formula has a proof π' in $e\mathcal{F}$ with $\lambda(\pi') \leq L(n)$ with $n = l(\neg A_1 \vee (\neg A_2 \vee$

$\dots(\neg A_n \vee B)\dots)) = l(A_1) + \dots + l(A_m) + l(B)$. If we assume that $\mathcal{F}$ has the *Modus Ponens scheme rule*

$$\frac{P, \neg P \vee Q}{Q}$$

We can apply this rule m times to π' and obtain a derivation π of B from $A_1, \dots, A_m$ with the property that $\lambda(\pi) \leq L(n) + m \leq 2L(n)$. This property suffices the property of the lemma.
If the *Modus Ponens scheme rule* is not in $\mathcal{F}$ then by Theorem 2.11 there is a polynomial time computable function that simulates the rules of one Frege System in another. In Theorem 2.11 we showed that the simulation f has the following property: $\lambda(f(\pi)) \leq c\lambda(\pi)$. We can apply the same transformation here and thus also obtain $\lambda(\pi) \leq cL(\pi)$.
If $\vee \notin \kappa$ or $\neg \notin \kappa$, we will find formulae with the functions $or(P, Q)$ and $neg(P)$ over κ that return the equivalent formulae to $P \vee Q$ and $\neg P$, respectively. We remark that for size purposes $or(P, Q)$ and $neg(P)$ have at most one occurrence each of P and Q. Hence, we obtain the at most polynomial bound: $\lambda(\pi) \leq cL(cn)$. □

Next, we provide some definitions that are also useful for the overall proof.

Definition 2.28. (cf. [2])

(i) For non-atomic formulae A we define an Atom P_A that is associated with the formula.

(ii) We assume that there are infinite many atoms P that we call *admissible atoms.* To be precise, these atoms are not of the form P_B for any non-atomic B.

(iii) We call a formula A *admissible* if all its atoms are admissible.

(iv) If A is admissible, then every truth assignment $\mathcal{A}$ to the atoms of A has a unique extension $\mathcal{A}'$ to the atoms P_B that are associated to the subformulae of A, such that $\mathcal{A}'(P_B) = \mathcal{A}(B)$.

(v) We denote the *defining set of formulae* for a formula A with $def_\kappa(A)$. κ can be any adequate set of connectives and does not necessarily have to refer to the set κ of A. $def_\kappa(A)$ is defined such that any extension $\mathcal{A}''$of $\mathcal{A}$ satisfies $def(A)$ iff $\mathcal{A}''$ agrees with $\mathcal{A}'$ on the atoms P_B.

Let κ_1 and κ_2 be connective sets.

(vi) Corresponding to each *nullary* connective K_1 (as $\top, \bot$) in κ_1 we associate a fixed formula K_2 over κ_2 that is equivalent to K_1.

(vii) Corresponding to each *unary* connective u_1 over κ_1 we associate a fixed formula u_2P over κ_2 equivalent to u_1P

(viii) Corresponding to each *binary* connective $\circ_1$ in κ_1 we associate a fixed formula $P \circ_2 Q$ over κ_2 equivalent to $P \circ_1 Q$.

(ix) We assume for the defining formulae $P \sim_1 Q$ over κ_1 and $P \sim_2 Q$ over κ_2 are each equivalent to $P \equiv Q$.

(x) For each formula A_1 over κ_1 we associate a set $def_{\kappa_2}(A_1)$ over κ_2 defined by induction on the length of A_1 as follows:
$def_{\kappa_2}(P) = \emptyset$ for each atom P that is *admissible* (see (ii))
$def_{\kappa_2}(K_1) = \{P_{K_1} \sim_2 K_2\}$ for each constant K_1 in κ_1
$def_{\kappa_2}(u_1 A) = def_{\kappa_2}(A) \cup def_{\kappa_2}(P_{u_1 A} \sim_2 u_2 P_A)$ for each unary connective u_1 in κ_1
$def_{\kappa_2}(A \circ_1 B) = def_{\kappa_2}(A) \cup def_{\kappa_2}(B) \cup \{P_{A \circ_1 B} \sim_2 (P_A \circ_2 P_B)\}$ for each binary connective $\circ_1$ in κ_1

In the case that $\kappa_1 = \kappa_2$ we assume that $K_1 = K_2$, $u_1 = u_2$ and $\circ_1 = \circ_2$.
(Remark: the total number of occurrences of atoms in def_{K_2} is bounded by a linear function of $l(A)$.

Example 2.29 (to (v) (see [2])). Let A be $Q \vee (R \wedge S)$. Then $def_\kappa(A) = \{(P_{R\wedge S} \equiv (R \wedge S)), (P_A \equiv (Q \vee P_{R\wedge S}))\}$.

Lemma 2.30. *(cf. [2]) Suppose $e\mathcal{F}$ is an* Extended Frege System *over κ, A is an admissible formula over κ and $def_\kappa(A) \vdash^{\pi}_{e\mathcal{F}} P_A$. Then for some π' we have $\vdash^{\pi}_{e\mathcal{F}} A$ with the properties that $\lambda(\pi') \leq \lambda(\pi) + cl(A)$ and $\rho(\pi') \leq (\rho(\pi) + c)$ and c depends only on $\mathcal{F}$.*

Proof. (cf. [2]) Let σ be the substitution $[P_E/E]$ for all non-atomic subformulae E of A. Particularly, this means $\sigma P_A = A$. Then every formula in $\sigma(def_\kappa(A))$ is an instance of $P \sim P$. Each of this instances will have a proof in $\mathcal{F}$ of some fixed number of lines, and a number of atoms bounded by a constant times $l(A)$. These proofs, together with $\sigma(\pi)$, contain π'. □

Lemma 2.31. *Let $e\mathcal{F}$ and $e\mathcal{F}'$ be two Extended Frege systems with the connective sets κ and κ'. Let A' be a formula that is valid over κ', i.e., that only contains connectives in κ'. Let further be $def_\kappa(A') \vdash^{\pi}_{e\mathcal{F}} P_{A'}$. Then there is a derivation π' in $e\mathcal{F}'$ such that $def_{\kappa'}(A') \vdash^{\pi'}_{e\mathcal{F}'} P_{A'}$. Further $\lambda(\pi') \leq c\lambda(\pi)$ and $\rho(\pi') \leq d$ where c and d are constants that only depend on the choice of $e\mathcal{F}$ and $e\mathcal{F}'$.*

Proof. Assume that the proof π consists of the lines $B_1, \ldots, B_m$. We can assume that the proof π is correct, i.e., all its lines are admissible except for those of the form $P_{C'}$ (i.e., this single atom) where C' is a subformula of A'.

For the proof, we construct a derivation π' in $e\mathcal{F}'$. We do this by starting with a skeleton derivation $P_{B_1}, \ldots, P_{B_m}$ and then fill in the required proof between these lines. Note that all these lines of the skeleton contain only single atoms and that the last one is identical to $P_{A'}$. The core of the proof will then be that we show that each P_{B_i} can be derived from P_{B_j}'s occurring earlier in the proof and from $def_{\kappa'}(A')$. Such proofs will have at most c lines and only formulae C of length $l(C) \leq d$ for two constants c and d that only depend on $e\mathcal{F}$ and $e\mathcal{F}'$.

For the proof, we have do distinguish by which means B_i was derived in the proof π. For all three cases we will consider, we will assume that some of the formulae of $def_{\kappa'}(B_i)$ are available in π'. This is either because they are part of $def_{\kappa'}(A')$ of π' or because they have been introduced at the beginning of π' by the extension rule. For any C the defining formula for P_C is either a subformula if $def_{\kappa'}(A')$ if C is a subformula of A' or the defining formula for P_C can be introduced at the beginning of π'.

Case 1: B_i might be a hypothesis, i.e., it might be in $def_\kappa(A')$. Wlog, we can assume that B_i has the form $P_{C'} \sim (P_{D'} \circ P_{E'})$ for some 2-ary boolean connective $\circ$, where C', D', and E' are subformulae of A' and $P \circ Q$ is a fixed formula over κ that is equivalent to $P \circ' Q$ and that $C' = D' \circ' E'$. Then $P_{C'} \sim' (P_{D'} \circ' P_{E'})$ is in $def_{\kappa'}(A')$ and thus it is a hypothesis of π. Let $H(\circ')$ be the formula $P \sim (Q \circ R)$ over κ. The value of $H(\circ')$ only depends on the 2-ary connective $\circ'$ but not on the concrete choice of B_i.

From this, we know that the rule

$$R = \frac{P \sim' (Q \circ' R), def_{\kappa'}(H(\circ'))}{P_{H(\circ')}}$$

is correct. Based on Theorem 2.11, we can now assume that R is a rule in $e\mathcal{F}'$. We can no construct an extension of the following substitution σ such that $\sigma(def_{\kappa'}(H(\circ'))) = def_{\kappa'}(B_i)$:

$$\sigma = \frac{P_{C'}}{P}\frac{P_{D'}}{D}\frac{P_{E'}}{R}\frac{P_B}{P_{H(\circ')}}$$

We can then apply the rule R with σ to $def_{\kappa'}(A')$ and $def_{\kappa'}(B_i)$ and conclude P_{B_i} as required.

Case 2: B_i may have been introduced into π via the extension rule. In this case B_i has the from $P \sim C$ for some atom P that did not previously occur in the proof. In particular, P does not occur in the hypothesis or conclusion of π'. Further, we could at this point in the proof (i.e., between the sections deriving B_{i-1} and B_i) first introduce any formula of $def_{\kappa *}(B_i)$ that did not previously occur in the proof via the extension rule. Next we can also introduce $P \sim' P_C$ via the extension rule. Lastly, we add the formula $P_{B_i} \sim' (P \sim' P_C)$.

Now if $\equiv \in \kappa$, then B_i will be of the form $P \equiv C$. Further $P_{B_i} \sim' (P \sim' P_C)$ will already be part of $def_{kappa'}(B_i)$. Using Theorem 2.11, we know that P_{B_i} can be deduced in π' from $P \sim' P_C$ and $P_{B_i} \sim' (P \sim' P_C)$ with a bounded number of steps.
If in the contrary $\equiv \notin \kappa$, such a proof still exists. It may be longer as the representation of the proof depends on the concrete κ and κ'. It can however only be longer by a constant factor not depending on B_i.
In total, B_i can thus be deduced in a bounded number of steps.

Case 3: Lastly, B_i might have been derived from the earlier lines by a true Frege Rule. Let this rule be $R = \frac{C_1,\ldots,C_n}{D}$. It will have been applied under some substitution σ. Since R was a Frege Rule, $C_1, \ldots, C_n \models D$ – i.e., the rule is correct. From this, we can now construct a new rule by splitting up the formulae C_i into their definitions and the truth of their representative atom P_{C_i} and move the definition of D into the antecedent of the rule as follows:

$$R' = \frac{def_{\kappa'}(D), def_{\kappa'}(C_1), \ldots def_{\kappa'}(C_n), P_{C_1}, \ldots, P_{P_n}}{P_D}$$

Using Theorem 2.11, we can assume wlog that this rule R' is in $e\mathcal{F}'$. We can next construct a new substitution as $\sigma = [\sigma(E)/P_E \mid \forall E \in \{C_1, \ldots, C_n, D\}]$. It then holds that $\forall j \in \{1, \ldots, n\} : \sigma'(def_{\kappa'}(C_j)) \subseteq def_{\kappa'}(\sigma(C_j))$ and $\sigma'(def_{\kappa'}(D)) \subseteq def_{\kappa'}(\sigma(D))$. Naturally each $\sigma(C_j)$ yields some B_j with $j < i$ – as the original rule R was applicable. Further $\sigma(D) = B_i$ as we used R to conclude B_i. $P_{\sigma(C_j)}$ must occur before the lines we are currently constructing in the new proof – as we have already modified these lines and $\sigma(C_j)$ was used in the original proof. Finally, we can now conclude that we can apply R' under σ' and that all its antecedents are either prior formulae in the proof or hypothesis. The number of these formulae is bounded, which concludes the proof of the lemma. □

Now we prove Theorem 2.24 with the recent results and definitions.

Proof. (cf. [2]) Suppose the hypothesis of 2.24. Let A' be an arbitrary but valid formula over κ'. Suppose now that A' is *admissible.* (If not, we rename the atoms of A' in order to make it admissible. Then we prove what we obtained and finally rename the atoms of the proof in order to get a suitable proof of A'). Thus, $def_\kappa(A') \models P_{A'}$. By the hypothesis it follows that $l(def_\kappa(A'))$ is an upper bound to the number of lines.
By Lemma 2.27 we obtain the existence of a derivation π in $e\mathcal{F}$ of $P_{A'}$ from $def_\kappa(A')$ with the property that $\lambda(\pi) \leq c_1 L(c_1 l(A'))$ for some constant c_1.
After Lemma 2.31 exists a derivation π' in $e\mathcal{F}'$ of $P_{A'}$ from $def_{\kappa'}(A')$ with the property that $\lambda(\pi') \leq c_2 L(c1l(A'))$ for some constant c_2 and $\rho(\pi') \leq d$ for some constant d. Finally, by Lemma 2.30 Theorem 2.24 follows. □

Next, we may prove Theorem 2.25.

Proof. (cf. [2]) Assume again that A is *admissible* (as above). It follows by induction on the length of B that for every *admissible* formula B over some connective set κ (that is also the connective set for $e\mathcal{F}$) there is a derivation π_B in $e\mathcal{F}$ of $P_B \sim B$ from $def_\kappa(B)$ with the properties $\lambda(\pi_B) \leq c_1 l(B)$ and $\rho(\pi_B) \leq c_2 l(B)$. By concatenating π_A with π of the theorem this yields to the derivation π_1 of P_A from $def_\kappa(A)$ such that $\lambda(\pi_1) \leq c_3(\lambda(\pi) + l(A))$ and $\rho(\pi_1) \leq c_4 l(\pi)$. By applying Lemma 2.31 with $\kappa = \kappa'$, $e\mathcal{F} = e\mathcal{F}'$ and $\pi = \pi_1$ we modify π_1. Thus, its formulas have bounded length. Next, we apply Lemma 2.30 to the resulting derivation. □

3. Conclusion

Investigating whether tautologies have short proofs is an important task as this may lead us to better understanding of the relationship of $\mathbb{NP}$ and $co\mathbb{NP}$. We have formalised the notion of proof length and p-simulation, which allows us to make formal statements of the relative length of proofs for different proof systems. For Frege systems, we have seen that the concrete choice of Frege rules does not make a difference with respect to the length of the proofs. I.e. if one has short proofs all others do as well. We have then discussed the Pigeon-Hole principle and a possible proof in Frege Systems. This proof led us to suggesting that Frege Systems are not good enough in the sense that they might have long proofs for some tautologies. We have thus introduced Extended Frege Systems, in which the Pigeon-Hole principle has a short proof. Lastly, we have shown that for extended Frege Systems the choice of the Frege rules is again irrelevant and that it suffices to show that the number of lines in the proof is polynomially bounded – as a consequence the length of the longest line of the proof is also short.

References

[1] Samuel R Buss. Polynomial size proofs of the propositional pigeonhole principle. *The Journal of Symbolic Logic*, 52(4):916–927, 1987.

[2] Stephen A Cook and Robert A Reckhow. The relative efficiency of propositional proof systems. *The journal of symbolic logic*, 44(1):36–50, 1979.

[3] Nicola Galesi and Jacobo Torán. *A Gentle Introduction to Propositional Proof Complexity.* 2022.

[4] Richard Statman. Complexity of derivations from quantifier-free horn formulae, mechanical introduction of explicit definitions, and refinement of completeness theorems. In *Studies in Logic and the Foundations of Mathematics*, volume 87, pages 505–518. Elsevier, 1977.