The Deduction Theorem for Strong Propositional Proof Systems

(Extended Abstract)

Olaf Beyersdorff*

Institut für Informatik, Humboldt-Universität zu Berlin, Germany
beyersdo@informatik.hu-berlin.de

Abstract. This paper focuses on the deduction theorem for propositional logic. We define and investigate different deduction properties and show that the presence of these deduction properties for strong proof systems is powerful enough to characterize the existence of optimal and even polynomially bounded proof systems. We also exhibit a similar, but apparently weaker condition that implies the existence of complete disjoint NP-pairs. In particular, this yields a sufficient condition for the completeness of the canonical pair of Frege systems and provides a general framework for the search for complete NP-pairs.

1 Introduction

The classical deduction theorem for propositional logic explains how a proof of a formula ψ from an extra hypothesis φ is transformed to a proof of $\varphi \to \psi$. While this property has been analysed in detail and is known to hold for Frege systems [3, 4], deduction has not been considered for stronger systems such as extensions of Frege systems, the apparent reason being that neither the extended Frege system EF nor the substitution Frege system SF satisfy the classical deduction theorem, as neither the extension nor the substitution rule is sound. We therefore relax the condition by requiring the extra hypothesis φ to be tautological. In this way we arrive at two weaker versions of the deduction property, for which we ask whether they are valid for strong proof systems with natural properties. It turns out that even these weaker versions of deduction are very powerful properties for strong proof systems as they allow the characterization of the existence of optimal and even polynomially bounded proof systems.

These characterizations are interesting as they relate to important concepts from different areas. The problem of the existence of polynomially bounded proof systems is known to be equivalent to the NP versus coNP question [6], while the existence of optimal proof systems is a famous and well-studied problem in proof complexity, posed by Krajíček and Pudlák [16], and with implications for a number of promise complexity classes (cf. [14, 19]). In particular, Sadowski [19] obtained different characterizations for the existence of optimal proof systems

* Supported by DFG grant KO 1053/5-1

in terms of optimal acceptors and enumerability conditions for easy subsets of TAUT. Earlier, Krajíček and Pudlák [16] established $\mathsf{NE} = \mathsf{coNE}$ as a sufficient condition for the existence of optimal proof systems, while Köbler et al. [14] showed that optimal proof systems imply complete sets for a number of other complexity classes like $\mathsf{NP} \cap \mathsf{coNP}$ and BPP.

On the other hand, we show that weak deduction combined with suitable closure properties of the underlying proof system implies the existence of complete disjoint NP-pairs. Although disjoint NP-pairs were already introduced into complexity theory in the 80's by Grollmann and Selman [12], it was only during recent years that disjoint NP-pairs have fully come into the focus of complexity-theoretic research [17, 8–11, 2, 1]. This interest mainly stems from the applications of disjoint NP-pairs to such different areas as cryptography [12, 13] and propositional proof complexity [18, 17, 2].

Similarly as for other promise classes it is not known whether the class of all disjoint NP-pairs contains pairs that are complete under the appropriate reductions. This question, posed by Razborov [18], is one of the most prominent open problems in the field. On the positive side, it is known that the existence of optimal proof systems suffices to guarantee the existence of complete pairs [18]. More towards the negative, a body of sophisticated relativization results underlines the difficulty of the problem. Glaßer et al. [8] provided an oracle under which complete disjoint NP-pairs do not exist. On the other hand, in [9] they also constructed an oracle relative to which there exist complete pairs but optimal proof systems do not exist.

Further information on the problem is provided by a number of different characterizations. Glaßer, Selman, and Sengupta [8] obtained a condition in terms of uniform enumerations of machines and also proved that the question of the existence of complete pairs receives the same answer under reductions of different strength. Additionally, the problem was characterized by provability conditions in propositional proof systems and shown to be robust under an increase of the number of components from two to arbitrary constants [1].

In this paper we exhibit several sufficient conditions for the existence of complete disjoint NP-pairs which involve properties of concrete proof systems such as Frege systems and their extensions. These results fall under a general paradigm for the search for complete NP-pairs, that asks for the existence of proof systems satisfying a weak version of the deduction theorem and moderate closure conditions. In particular, we provide two conditions that imply the completeness of the canonical pair of Frege systems and demonstrate that the existence of complete NP-pairs is tightly connected with the question whether EF is indeed more powerful than ordinary Frege systems.

The paper is organized as follows. In Sect. 2 we provide some background information on propositional proof systems and disjoint NP-pairs. In Sect. 3 we discuss various extensions of Frege systems that we investigate in Sect. 4 with respect to different versions of the deduction property. Section 5 contains the results connecting the deduction property for strong systems with the existence of complete NP-pairs. Finally, in Sect. 6 we conclude with some open problems.

Due to space limitations we only sketch proofs or omit them in this extended abstract. Proofs of the main results are contained in the appendix.

2 Preliminaries

Propositional Proof Systems. Propositional proof systems were defined in a very general way by Cook and Reckhow [6] as polynomial-time functions P which have as its range the set of all tautologies. A string π with $P(\pi) = \varphi$ is called a P-proof of the tautology φ. By $P \vdash_{\leq m} \varphi$ we indicate that there is a P-proof of φ of size $\leq m$. We write $P \vdash_* \varphi_n$ if φ_n is a sequence of tautologies with polynomial-size P-proofs.

Proof systems are compared according to their strength by simulations introduced in [6] and [16]. A proof system Q *simulates* a proof system P (denoted $P \leq Q$), if there exists a polynomial p such that $P \vdash_{\leq m} \varphi$ implies $Q \vdash_{\leq p(m)} \varphi$ for all formulas φ. A proof system is called *optimal* if it simulates all proof systems.

In the following sections simple closure properties of propositional proof systems will play an important role. We say that a proof system P is *closed under modus ponens* if there exists a constant c such that $P \vdash_{\leq m} \varphi$ and $P \vdash_{\leq n} \varphi \to \psi$ imply $P \vdash_{\leq m+n+|\psi|+c} \psi$ for all formulas φ and ψ. Similarly, we say that P is *closed under substitutions of variables with respect to the polynomial* q if $P \vdash_{\leq m} \varphi(\bar{x})$ implies $P \vdash_{\leq q(m)} \varphi(\bar{y})$ for all formulas $\varphi(\bar{x})$ and propositional variables $\bar{y}$ that are distinct from $\bar{x}$. Not specifying the polynomial explicitly, we say that P is *closed under substitutions of variables* if there exists a polynomial q with this property. Likewise, P is *closed under substitutions by constants* if there exists a polynomial q such that $P \vdash_{\leq m} \varphi(\bar{x}, \bar{y})$ implies $P \vdash_{\leq q(m)} \varphi(\bar{a}, \bar{y})$ for all formulas $\varphi(\bar{x}, \bar{y})$ and constants $\bar{a} \in \{0,1\}^{|\bar{x}|}$.

Disjoint NP-Pairs. A pair (A, B) is called a *disjoint* NP*-pair* if $A, B \in \mathsf{NP}$ and $A \cap B = \emptyset$. Grollmann and Selman [12] defined the following reduction between disjoint NP-pairs (A, B) and (C, D): $(A, B) \leq_p (C, D)$ if there exists a polynomial-time computable function f such that $f(A) \subseteq C$ and $f(B) \subseteq D$.

The connection between disjoint NP-pairs and propositional proof systems was established by Razborov [18], who associated a *canonical disjoint* NP*-pair* $(\mathrm{Ref}(P), \mathrm{SAT}^*)$ with a proof system P, where the first component $\mathrm{Ref}(P) = \{(\varphi, 1^m) \mid P \vdash_{\leq m} \varphi\}$ contains information about proof lengths in P and the second component $\mathrm{SAT}^* = \{(\varphi, 1^m) \mid \neg\varphi \in \mathrm{SAT}\}$ is a padded version of SAT. This canonical pair is linked to the automatizablility and the reflection property of the proof system [17]. More information on the connection between disjoint NP-pairs and propositional proof systems can be found in [17, 2, 10].

3 Extensions of Frege Systems

A prominent example of a class of proof systems is provided by *Frege systems* which are usual textbook proof systems based on axioms and rules. In the context

of propositional proof complexity these systems were first studied by Cook and Reckhow [6] and it was proven there that all Frege systems, i.e., systems using different axiomatizations and rules, are polynomially equivalent.

Augmenting Frege systems by the possibility to abbreviate complex formulas by propositional variables we arrive at the *extended Frege proof system* EF. This extension rule might further reduce the proof size, but it is not known whether EF is really stronger than ordinary Frege systems. Both Frege and the extended Frege system are very strong systems for which no non-trivial lower bounds to the proof size are currently known.

Another way to enhance the power of Frege systems is to allow substitutions not only for axioms but also for all formulas that have been derived in Frege proofs. Augmenting Frege systems by this substitution rule leads to the *substitution Frege system* SF. The extensions EF and SF were introduced by Cook and Reckhow [6]. While it was already proven there that EF is simulated by SF, the converse simulation is considerably more involved and was shown independently by Dowd [7] and Krajíček and Pudlák [16]. For more detailed information on Frege systems and its extensions we refer to the monograph [15].

Under the notion of *Hilbert-style proof systems* we subsume all proof systems that have as proofs sequences of formulas, and formulas in such a sequence are derived from earlier formulas in the sequence by the rules available in the proof system. In particular, Frege systems and its extensions are Hilbert-style systems. Hilbert-style proof systems P can be enhanced by additional axioms in two different ways. Namely, we can form a proof system $P + \Phi$ augmenting P by a polynomial-time computable set Φ of tautologies as new axiom schemes. This means that formulas from Φ as well as substitution instances of these formulas can be freely introduced as new lines in $P + \Phi$ -proofs. In contrast to this we use the notation $P \cup \Phi$ for the proof system that extends P only by formulas from Φ but not by their substitution instances as new axioms. In our applications the set Φ will mostly be *printable*, meaning that Φ can both be decided and generated in polynomial time.

For EF there are two canonical ways how to define the extensions $EF \cup \Phi$ and $EF + \Phi$, where these two possibilities differ in the use of the extension axioms. In the first method we will allow the introduction of extension axioms $p \equiv \varphi$ only for extension variables p not occurring in Φ, whereas in the second method we can freely use extension axioms that also involve variables from Φ. For the first weaker notion we will use the notation $EF^- \cup \Phi$ and $EF^- + \Phi$, or only EF^- when we augment EF in this manner by different sets of tautologies Φ, whereas the stronger second way is indicated by the usual notation $EF \cup \Phi$, $EF + \Phi$, or simply EF. We will use the same notation $(EF + \Psi)^-$ when we use an extension $EF + \Psi$ as the base system and augment this with further axioms Φ to systems $(EF + \Psi)^- \cup \Phi$.

In principle, this gives four possible types of extensions of EF, but it is easily seen that the distinction between EF and EF^- becomes irrelevant when we augment these systems by axiom schemes Φ:

Proposition 1. *Let Φ be a polynomial-time decidable set of tautologies. Then the proof systems $EF + \Phi$ and $EF^- + \Phi$ are polynomially equivalent.*

These extensions of EF are particularly important as every proof system P is simulated by a proof system of the form $EF + \Phi$ where the axioms Φ provide a propositional description of the reflection principle of P, expressing a strong form of the consistency of P (cf. [15] for details).

In addition, also the systems $EF \cup \Phi$ and $EF + \Phi$ appear to be very close to each other, as also $EF \cup \Phi$ can use substitution instances of Φ in its proofs. Namely, if $\varphi(p_1, \ldots, p_n)$ is a formula from Φ and $\theta_1(\bar{q}), \ldots, \theta_n(\bar{q})$ are propositional formulas in the variables $\bar{q}$ that are disjoint from $\bar{p}$, then we can deduce $\varphi(\theta_1, \ldots, \theta_n)$ in $EF \cup \Phi$ as follows: we start with the extension axioms $p_1 \equiv \theta_1(\bar{q}), \ldots, p_n \equiv \theta_n(\bar{q})$ and use these formulas to show the equivalence $\varphi(p_1, \ldots, p_n) \equiv \varphi(\theta_1, \ldots, \theta_n)$ by induction on the formula φ. Using the original axiom $\varphi(p_1, \ldots, p_n)$ from Φ we arrive with modus ponens at the substitution instance $\varphi(\theta_1, \ldots, \theta_n)$. We leave it open, whether this idea can be extended to a full simulation of $EF + \Phi$ by $EF \cup \Phi$, but the argument shows that also the system $EF \cup \Phi$ is quite natural, as it is equivalent to the proof system $P = EF + \Phi$ where formulas from Φ use pairwise distinct variables and each P-proof may contain at most one substitution instance of each formula from Φ.

For SF the situation becomes even simpler, as there is only one sensible way to define extensions of SF. Namely, because SF can immediately generate substitution instances, we have $SF \cup \Phi \equiv_p SF + \Phi$. In total the following picture of possible extension of Frege systems emerges:

Proof system	Extensions by polynomial-time decidable axioms Φ
F	$F \cup \Phi \ \leq_p \ F + \Phi$
EF	$EF^- \cup \Phi \ \leq_p \ EF \cup \Phi \ \leq_p \ EF^- + \Phi \ \equiv_p \ EF + \Phi$
SF	$SF \cup \Phi \ \equiv_p \ SF + \Phi$

In the above table all shown simulation relations are probably strict in each line (except for $EF \cup \Phi \leq_p EF + \Phi$ as mentioned above), because the converse simulations have unlikely consequences, as we will show in the sequel of this paper, or easily follows from known results. The next table gives an overview of these consequences, ranging in strength from the existence of complete disjoint NP-pairs to the existence of optimal proof systems.

Assumption		Consequence
$F \equiv F^- \cup \Phi$	*)	EF is optimal (cf. [15] and the Appendix)
$F \cup \Phi \equiv F + \Phi$	*)	Complete disjoint NP-pairs exist (Corollary 15)
$EF \equiv EF^- \cup \Phi$	*)	EF is optimal (cf. [15])
$EF^- \cup \Phi \equiv EF \cup \Phi$	*)	EF is optimal (Theorem 8)
$SF \equiv SF \cup \Phi$	*)	SF is optimal (cf. [15])

*) for all polynomial-time decidable sets of tautologies Φ

In contrast, we do not seem to have such indication for separating the systems in the vertical columns of the first table, as even the relation between F and $EF \equiv_p SF$ is not settled.

4 Deduction Properties for Frege Systems

The deduction theorem of propositional logic states that in a Frege system F a formula ψ is provable from a formula φ if and only if $\varphi \to \psi$ is provable in F. Because proof complexity is focusing on the length of proofs it is interesting to analyse how the proof length is changing in the deduction theorem. An F-proof of $\varphi \to \psi$ together with the axiom φ immediately yields the formula ψ with one application of modus ponens. Therefore it is only interesting to ask for the increase in proof length when constructing a proof of $\varphi \to \psi$ from an F-proof of ψ with the extra axiom φ. This was analysed in detail in [3, 4].

The main application of the deduction property is to simplify proofs of complex formulas. Namely, to prove an implication $\varphi \to \psi$ it suffices to construct a proof of ψ from φ. In particular, φ can be any formula and is not necessarily a tautology. It is clear that such a deduction property is doomed to fail for strong systems like EF or SF that can immediately produce substitution instances from φ. For instance, by one application of the substitution rule we get $SF \cup \{p\} \vdash q$, whereas $p \to q$ is not even a tautology. Similarly, we get $EF \cup \{p\} \vdash q$ by introducing the extension axiom $p \equiv q$ with extension variable p as the first line of the proof, and then derive q by modus ponens. This example, however, does not work for EF^- as we have used the variable p from the extra assumption as an extension variable. In fact, such an example cannot be found as the classical deduction theorem is valid for EF^- (Theorem 3).

Aiming in particular at strong proof systems like EF we therefore restrict φ to tautologies and make the following general definition.

Definition 2. *A Hilbert-style proof system P allows* efficient deduction *if there exists a polynomial p such that for all finite sets Φ of tautologies $P \cup \Phi \vdash_{\leq m} \psi$ implies $P \vdash_{\leq p(m+m')} (\bigwedge_{\varphi\in\Phi} \varphi) \to \psi$ where $m' = |\bigwedge_{\varphi\in\Phi} \varphi|$.*

If this even holds for all finite sets Φ of propositional formulas, then we say that P has the classical deduction property*.*

This classical deduction property is known to hold for Frege systems (cf. [4]), but actually almost the same proof also holds for the presumably stronger system EF^-.

Theorem 3 (Deduction theorem for Frege systems). *Let Ψ be a polynomial-time decidable set of tautologies. Then every Frege system $F + \Psi$ and every extended Frege system of the form $(EF + \Psi)^-$ has the classical deduction property.*

A still weaker form of the deduction property is given in the next definition.

Definition 4. *A Hilbert-style proof system P allows* weak deduction *if the following condition holds. For all printable sets $\Phi \subseteq$ TAUT there exists a polynomial p such that for all finite subsets $\Phi_0 \subseteq \Phi$ we can infer from $P \cup \Phi_0 \vdash_{\leq m} \psi$ that $P \vdash_{\leq p(m+m')} (\bigwedge_{\varphi\in\Phi_0} \varphi) \to \psi$ where $m' = |\bigwedge_{\varphi\in\Phi_0} \varphi|$.*

In Definition 2 we allowed a fixed polynomial increase for the proof size in the transformation of a proof from ψ to the implication $(\bigwedge_{\varphi\in\Phi_0}\varphi)\to\psi$, whereas in the weak deduction property this polynomial might depend on the choice of the extra axioms Φ. This weakening of the deduction property allows us to show the following proposition.

Proposition 5. *Optimal Hilbert-style proof systems have the weak deduction property. Similarly, polynomially bounded Hilbert-style proof systems have the efficient deduction property.*

Proof. (Idea) Let Φ be a printable set of tautologies and let π be a $P\cup\Phi$-proof of ψ. If P is optimal (or even polynomially bounded), then we can first devise polynomial-size P-proofs of the extra assumptions Φ_0 in π and thus construct a P-proof of $(\bigwedge_{\varphi\in\Phi_0}\varphi)\to\psi$. □

The following theorem provides a form of a converse to the last proposition. This shows that even the weak deduction property is a very strong assumption for natural proof systems.

Theorem 6. *Let $P\geq EF$ be a Hilbert-style proof system that fulfills the following conditions:*

1. *P is closed under modus ponens and substitutions by constants.*
2. *For all printable sets of tautologies Φ the proof system $P\cup\Phi$ is closed under substitutions of variables.*
3. *P has the weak deduction property.*

Then P is an optimal proof system.

Proof. To obtain the optimality of a proof system $P\geq EF$ that is closed under modus ponens, it suffices to show $P\vdash_* \varphi_n$ for all printable sequences of tautologies φ_n (cf. [15] Theorem 14.2.2). Let $\varphi_n(\bar{p})$ be a printable sequence in the variables $\bar{p}$, and let $\bar{q}$ be a sequence of propositional variables that is disjoint from $\bar{p}$. We consider the proof system $P'=P\cup\{\varphi_n(\bar{q})\mid n\geq 0\}$ where the variables $\bar{p}$ from $\varphi_n(\bar{p})$ are substituted by $\bar{q}$. By assumption P' is closed under substitutions of variables and hence we have $P'\vdash_*\varphi_n(\bar{p})$. By the weak deduction property for P we get $P\vdash_*\bigwedge_{i\in I}\varphi_i(\bar{q})\to\varphi_n(\bar{p})$ for some finite set I. Using closure under substitutions by constants we derive $P\vdash_*\bigwedge_{i\in I}\varphi_i(1,\ldots,1)\to\varphi_n(\bar{p})$ where we have substituted all variables $\bar{q}$ in $\varphi_i(\bar{q})$ by constants 1. Because all φ_i are tautologies, the formulas $\varphi_i(1,\ldots,1)$ are true formulas without variables and therefore admit polynomial-size P-proofs, as $P\geq EF$. Using modus ponens for P we arrive at polynomial-size P-proofs of $\varphi_n(\bar{p})$, as desired. □

Polynomially bounded proof systems P can be characterized by $P\vdash_{\leq p(n)}\varphi_n$ for all printable sequences of tautologies φ_n and a fixed polynomial p. In the definition of the efficient deduction property and the other closure properties we have also bounded the increase in the proof length by fixed polynomials. Hence the proof of the previous theorem yields the following result.

Theorem 7. *Let $P \geq EF$ be a Hilbert-style proof system that fulfills the following conditions:*

1. *P is closed under modus ponens and substitutions by constants.*
2. *There exists a polynomial p such that for all printable sets of tautologies Φ the proof system $P \cup \Phi$ is closed under substitutions of variables with respect to p.*
3. *P has the efficient deduction property.*

Then P is a polynomially bounded proof system.

In comparison to Theorem 6 we replaced the hypothesis of weak deduction for P by the stronger notion of efficient deduction and arrive at the stronger consequence of the polynomial boundedness of P.

Examining the situation for extensions of EF we obtain the following result.

Theorem 8. *Let Ψ be a polynomial-time decidable set of tautologies. Then the following conditions are equivalent:*

1. *$EF + \Psi$ has the weak deduction property.*
2. *$EF + \Psi$ is an optimal proof system.*
3. *For all polynomial-time decidable sets $\Phi \subset TAUT$ the systems $(EF+\Psi)^- \cup \Phi$ and $(EF + \Psi) \cup \Phi$ are equivalent.*
4. *For all polynomial-time decidable sets $\Phi \subset TAUT$ the proof system $(EF + \Psi)^- \cup \Phi$ is closed under substitutions of variables.*

In particular, the last theorem yields two seemingly unrelated characterizations for the optimality of EF, namely weak deduction for EF and closure of $EF^- \cup \Phi$ under substitutions of variables for arbitrary tautologies Φ.

Similarly, we obtain the following characterizations for the efficient deduction property of extensions of EF.

Theorem 9. *Let Ψ be a polynomial-time decidable set of tautologies. Then the following conditions are equivalent:*

1. *$EF + \Psi$ has the efficient deduction property.*
2. *$EF + \Psi$ is polynomially bounded.*
3. *There exists a polynomial p such that for all polynomial-time decidable sets $\Phi \subset TAUT$ the proof system $(EF + \Psi)^- \cup \Phi$ is closed under substitutions with respect to p.*

While one might have objections on the naturality of the above systems $(EF + \Psi) \cup \Phi$, the same results are also valid for substitution Frege systems. In particular, we obtain from Theorems 6 and 7 the following characterizations.

Theorem 10. *Let Ψ be a polynomial-time decidable set of tautologies. Then the proof system $SF + \Psi$ is optimal if and only if $SF + \Psi$ has the weak deduction property. Further, the system $SF + \Psi$ is polynomially bounded if and only if $SF + \Psi$ has the efficient deduction property.*

As we know that every proof system P is simulated by a proof system of the form $EF + \Psi$ with printable $\Psi \subset \mathrm{TAUT}$ (for instance we can take Ψ as translations of the reflection principle of P), we can deduce the following characterization of the existence of optimal proof systems.

Corollary 11. *There exists an optimal proof system if and only if there exists a polynomial-time decidable set $\Psi \subset \mathrm{TAUT}$ such that $EF + \Psi$ has the weak deduction property.*

Similarly, we can characterize the existence of polynomially bounded proof systems by the efficient deduction property.

Corollary 12. *There exists a polynomially bounded proof system if and only if there exists a polynomial-time decidable set $\Psi \subset \mathrm{TAUT}$ such that $EF + \Psi$ has the efficient deduction property.*

5 Deduction Properties and Complete NP-Pairs

In this section we link the deduction property to the problem of the existence of complete disjoint NP-pairs. In this analysis properties of proof systems are transferred to properties of the corresponding canonical pairs of the systems.

Augmenting Hilbert-style proof systems P by additional axioms Φ will usually enhance the power of the proof system. The following lemma shows, however, that if P has the weak deduction property, then the canonical pair of $P \cup \Phi$ will not be more difficult than the canonical P-pair. In particular, combined with Theorem 3 the next lemma shows that the canonical pairs of F and its extensions $F \cup \Phi$ are equivalent for printable sets $\Phi \subseteq \mathrm{TAUT}$.

Lemma 13. *Let Φ be a printable set of tautologies and let P be a proof system with the weak deduction property. Then* $(\mathrm{Ref}(P \cup \Phi), \mathrm{SAT}^*) \leq_p (\mathrm{Ref}(P), \mathrm{SAT}^*)$.

Proof. (Idea) The reduction is performed by the mapping

$$(\psi, 1^m) \mapsto ((\bigwedge_{\varphi \in \Phi_m} \varphi) \to \psi, 1^{p(m)})$$

where $\Phi_m = \Phi \cap \Sigma^{\leq m}$, and p is the polynomial from the weak deduction property of P. □

In the next theorem we formulate a sufficient condition for the existence of complete NP-pairs. The hypotheses in this theorem are very similar to the hypotheses in Theorem 6, which gave a sufficient condition for the existence of optimal proof systems. The decisive difference between the two theorems is that in Theorem 6 we needed closure of P under substitutions of variables, whereas in the following theorem closure under substitutions by constants suffices.

Theorem 14. *Let P be a Hilbert-style proof system that simulates the truth-table system and fulfills the following three conditions:*

1. *P is closed under modus ponens.*
2. *For all printable sets of tautologies Φ the proof system $P \cup \Phi$ is closed under substitutions by constants.*
3. *P has the weak deduction property.*

Then the canonical pair of P is a complete disjoint NP*-pair.*

Proof. The idea of the proof is to construct suitable propositional representations of disjoint NP-pairs (A, B). Such representations for A and B can be obtained similarly as in Cook's proof of the NP-completeness of SAT [5]. We then form a proof system $P' = P \cup \Phi$ extending P, where Φ are new axioms expressing the disjointness of (A, B) with respect to the above representations. This allows to reduce (A, B) to the canonical pair of P'. As P has weak deduction, we can use Lemma 13 to reduce the canonical pair of P' to the canonical pair of P, and hence (A, B) is $\leq_p$-reducible to $(\mathrm{Ref}(P), \mathrm{SAT}^*)$. □

The decisive hypotheses in Theorem 14 are assumptions 2 and 3. For Frege systems property 3 of Theorem 14 is fulfilled but property 2 is not clear. For EF and SF, however, we have property 2, but whether property 3 holds is open. To find out whether some strong proof system fulfills both conditions 2 and 3 remains as a challenging task.

Instantiating Theorem 14 for Frege systems leads to the following corollary which asks, in principle, whether the systems $F \cup \Phi$ and $F + \Phi$ are equivalent.

Corollary 15. *Assume that for all printable sets of tautologies Φ the system $F \cup \Phi$ is closed under substitutions by constants. Then the canonical F-pair is a complete disjoint* NP*-pair.*

By Theorem 3 and Lemma 13 the same corollary also holds for the proof system EF^-.

Our last result shows that the existence of complete NP-pairs is tightly connected with the question whether F and EF are indeed proof systems of different strength.

Corollary 16. *Assume that for all printable sequences Φ of tautologies the proof systems $F \cup \Phi$ and $EF \cup \Phi$ are equivalent. Then the canonical pair of the Frege proof system is complete for the class of all disjoint* NP*-pairs.*

In Table 1 we have summarized the different deduction properties and their implications for the existence of complete NP-pairs for Frege systems and their extensions.

6 Conclusion

In this paper we have brought attention to the question whether strong proof systems such as extensions of Frege systems have some kind deduction property. On the one hand, we have shown that optimal proof systems can be characterized by the weak deduction property. On the other hand, weak deduction combined with a moderate amount of closure properties yields complete disjoint NP-pairs. It therefore seems to be interesting to investigate the following problem:

Proof system P	Frege/EF^-	EF/SF
classical deduction	yes	no
efficient deduction	yes	no, unless P is optimal
weak deduction	yes	no, unless P is polynomially bounded
weakest known condition for the completeness of the canonical pair of P	closure of $P \cup \Phi$ under substitutions by constants for all printable Φ	optimality of P

Table 1. Deduction properties for different types of proof systems

Problem 17. Are there natural strong proof systems besides Frege systems that satisfy the weak deduction property?

Given the implications above, we expect, however, that neither proving nor disproving this question will be an easy task.

It would also be interesting to know whether the condition in Corollary 15 also characterizes the completeness of the canonical Frege pair, similarly as in Corollaries 11 and 12. A more general program is to determine which consequences of the completeness of the canonical pair of some proof system P are to expect for the system P itself.

Acknowledgements. I am indebted to Emil Jeřábek, Johannes Köbler, and Pavel Pudlák for helpful suggestions on this work.

References

1. O. Beyersdorff. Tuples of disjoint NP-sets. *Theory of Computing Systems*. To appear.
2. O. Beyersdorff. Classes of representable disjoint NP-pairs. *Theoretical Computer Science*, 377:93–109, 2007.
3. M. L. Bonet. Number of symbols in Frege proofs with and without the deduction rule. In P. Clote and J. Krajíček, editors, *Arithmetic, Proof Theory and Computational Complexity*, pages 61–95. Oxford University Press, Oxford, 1993.
4. M. L. Bonet and S. R. Buss. The deduction rule and linear and near-linear proof simulations. *The Journal of Symbolic Logic*, 58(2):688–709, 1993.
5. S. A. Cook. The complexity of theorem proving procedures. In *Proc. 3rd Annual ACM Symposium on Theory of Computing*, pages 151–158, 1971.
6. S. A. Cook and R. A. Reckhow. The relative efficiency of propositional proof systems. *The Journal of Symbolic Logic*, 44:36–50, 1979.
7. M. Dowd. Model-theoretic aspects of P≠NP. Unpublished manuscript, 1985.
8. C. Glaßer, A. L. Selman, and S. Sengupta. Reductions between disjoint NP-pairs. *Information and Computation*, 200(2):247–267, 2005.

9. C. Glaßer, A. L. Selman, S. Sengupta, and L. Zhang. Disjoint NP-pairs. *SIAM Journal on Computing*, 33(6):1369–1416, 2004.
10. C. Glaßer, A. L. Selman, and L. Zhang. Survey of disjoint NP-pairs and relations to propositional proof systems. In O. Goldreich, A. L. Rosenberg, and A. L. Selman, editors, *Essays in Theoretical Computer Science in Memory of Shimon Even*, pages 241–253. Springer-Verlag, Berlin Heidelberg, 2006.
11. C. Glaßer, A. L. Selman, and L. Zhang. Canonical disjoint NP-pairs of propositional proof systems. *Theoretical Computer Science*, 370:60–73, 2007.
12. J. Grollmann and A. L. Selman. Complexity measures for public-key cryptosystems. *SIAM Journal on Computing*, 17(2):309–335, 1988.
13. S. Homer and A. L. Selman. Oracles for structural properties: The isomorphism problem and public-key cryptography. *Journal of Computer and System Sciences*, 44(2):287–301, 1992.
14. J. Köbler, J. Messner, and J. Torán. Optimal proof systems imply complete sets for promise classes. *Information and Computation*, 184:71–92, 2003.
15. J. Krajíček. *Bounded Arithmetic, Propositional Logic, and Complexity Theory*, volume 60 of *Encyclopedia of Mathematics and Its Applications*. Cambridge University Press, Cambridge, 1995.
16. J. Krajíček and P. Pudlák. Propositional proof systems, the consistency of first order theories and the complexity of computations. *The Journal of Symbolic Logic*, 54:1963–1079, 1989.
17. P. Pudlák. On reducibility and symmetry of disjoint NP-pairs. *Theoretical Computer Science*, 295:323–339, 2003.
18. A. A. Razborov. On provably disjoint NP-pairs. Technical Report TR94-006, Electronic Colloquium on Computational Complexity, 1994.
19. Z. Sadowski. On an optimal propositional proof system and the structure of easy subsets of TAUT. *Theoretical Computer Science*, 288(1):181–193, 2002.

Appendix

The appendix contains all proofs that are omitted or only briefly sketched in the extended abstract.

Proposition 1. *Let Φ be a polynomial-time decidable set of tautologies. Then the proof systems $EF + \Phi$ and $EF^- + \Phi$ are polynomially equivalent.*

Proof. By definition the system $EF + \Phi$ simulates the system $EF^- + \Phi$. For the converse simulation let π be an $EF + \Phi$-proof of a formula φ. In order to convert π into an $EF^- + \Phi$-proof we only have to check, whether π contains extension axioms $p \equiv \theta$ with extension variables p that occur in the axiom set Φ. If this is not the case, then π is already an $EF^- + \Phi$-proof. Otherwise, we just rename every occurrence of p in π to a new variable q, which neither appears in π nor in Φ. Performing this step for every extension atom in π, we already arrive at an $EF^- + \Phi$-proof of φ. This is correct, because the proven formula φ may not contain any extension variables, and renaming variables in axioms from Φ in the proof still results in valid substitution instances of Φ, which we are permitted to use in $EF^- + \Phi$-proofs. □

Next we want to argue for the implications of the following table from Sect. 3.

Assumption		Consequence
$F \equiv F^- \cup \Phi$	*)	EF is optimal
$F \cup \Phi \equiv F + \Phi$	*)	Complete disjoint NP-pairs exist (Corollary 15)
$EF \equiv EF^- \cup \Phi$	*)	EF is optimal
$EF^- \cup \Phi \equiv EF \cup \Phi$	*)	EF is optimal (Theorem 8)
$SF \equiv SF \cup \Phi$	*)	SF is optimal

*) for all polynomial-time decidable sets of tautologies Φ

The implication in lines 2 and 4 follow from Corollary 15 and Theorem 8 below. All other implications can be derived from the following result from [15]:

Theorem (Krajíček [15]). *Let $P \geq EF$ be a proof system that is closed under substitutions and modus ponens. Then P is optimal if and only if $P \vdash_* \varphi_n$ for all printable sequences φ_n of tautologies.*

To derive the first line from the above table, assume that $F \equiv F^- \cup \Phi$ for all polynomial-time decidable sets of tautologies Φ. In particular, this means $F \vdash_* \varphi_n$ for all printable sequences φ_n, and hence also $EF \vdash_* \varphi_n$. As EF has the necessary closure properties, the optimality of EF follows by the above theorem. Lines 4 and 5 are deduced analogously.

Theorem 3 (Deduction theorem for Frege systems). *Let Ψ be a polynomial-time decidable set of tautologies. Then every Frege system $F + \Psi$ and every extension Frege system of the form $(EF+\Psi)^-$ has the classical deduction property.*

Proof. For every F-rule

$$R_i = \frac{\psi_1 \quad \dots \quad \psi_r}{\psi}$$

we fix an F-proof π_i of the tautology

$$((q \to \psi_1) \wedge \cdots \wedge (q \to \psi_r)) \to (q \to \psi) \ .$$

In particular, for $r = 0$ this also includes the case that R_i is an axiom scheme.

Let $\varphi_1, \dots, \varphi_n$ be tautologies and let $(\theta_1, \dots, \theta_k)$ be a proof of ψ of size m in the system $P \cup \{\varphi_1, \dots, \varphi_n\}$, where P is $F+\Psi$ or $(EF+\Psi)^-$. Let $m' = \sum_{i=1}^n |\varphi_i|$. By induction on j we construct proofs of the implications

$$(\bigwedge_{i=1}^n \varphi_i) \to \theta_j \ .$$

We distinguish three cases on how the formula θ_j was derived.

If θ_j is one of the formulas from $\{\varphi_1, \dots, \varphi_n\}$ or a substitution instance from Ψ, then we get $(\bigwedge_{i=1}^n \varphi_i) \to \theta_j$ in a proof of size $O(m')$.

If θ_j was inferred from $\theta_{j_1}, \dots, \theta_{j_r}$ by the F-rule R_i, then we can get from π_i an F-proof of size $O(m' + |\theta_j| + \sum_{l=1}^r |\theta_{j_l}|)$ of the tautology

$$((\bigwedge_{i=1}^n \varphi_i) \to \theta_{j_1}) \wedge \cdots \wedge ((\bigwedge_{i=1}^n \varphi_i) \to \theta_{j_r}) \to ((\bigwedge_{i=1}^n \varphi_i) \to \theta_j).$$

Combining all the earlier proved implications $(\bigwedge_{i=1}^n \varphi_i) \to \theta_{j_l}$, $l = 1, \dots, r$ by conjunctions and using modus ponens we get the desired implication $(\bigwedge_{i=1}^n \varphi_i) \to \theta_j$ in a proof of size $O(m + m')$.

If in the case of $(EF+\Psi)^-$ the formula θ_j was derived by the extension rule, i.e.,

$$\theta_j = (q \equiv \theta)$$

with a new variable q, then we can also use the extension rule to get $q \equiv \theta$ and derive

$$(\bigvee_{i=1}^n \neg\varphi_i) \vee (q \equiv \theta) = (\bigwedge_{i=1}^n \varphi_i) \to (q \equiv \theta) \ .$$

in a proof of size $O(m'+|\theta|)$. Here it is important that by the definition of $(EF+\Psi)^-$ the extension variable q does not occur in the formulas φ_i, as otherwise we would not be able to use q as an extension variable in an $EF + \Psi$-proof of $(\bigwedge_{i=1}^n \varphi_i) \to \theta_k$. □

Proposition 5. *Optimal Hilbert-style proof systems have the weak deduction property. Similarly, polynomially bounded Hilbert-style proof systems have the efficient deduction property.*

Proof. Let P be an optimal Hilbert-style proof system and let Φ be a printable set of tautologies. Then $P \cup \Phi$ is a well defined proof system which by the optimality of P is simulated by P. Hence we have polynomial-size P-proofs of all formulas from Φ. Given a finite set Φ_0 and a $P \cup \Phi_0$-proof π of a formula ψ we can therefore first derive all formulas from Φ_0 in polynomial-size P-proofs and concatenate this with π. This results in a polynomial-size P-proof of ψ from which we easily obtain a polynomial-size P-proof of $(\bigwedge_{\varphi \in \Phi_0} \varphi) \to \psi$.

If P is polynomially bounded, then we obtain by the same argument a polynomial bound on the proof size of the formulas $(\bigwedge_{\varphi \in \Phi_0} \varphi) \to \psi$ which is independent of Φ. □

Theorem 8. *Let Ψ be a polynomial-time decidable set of tautologies. Then the following conditions are equivalent:*

1. *$EF + \Psi$ has the weak deduction property.*
2. *$EF + \Psi$ is an optimal proof system.*
3. *For all polynomial-time decidable sets $\Phi \subset TAUT$ the systems $(EF+\Psi)^- \cup \Phi$ and $(EF + \Psi) \cup \Phi$ are equivalent.*
4. *For all polynomial-time decidable sets $\Phi \subset TAUT$ the proof system $(EF + \Psi)^- \cup \Phi$ is closed under substitutions of variables.*

Proof. We will prove the implications $1 \Leftrightarrow 2$, $2 \Rightarrow 3$, $3 \Rightarrow 4$, and $4 \Rightarrow 2$.

To prove item 2 from item 1, let us assume that $EF + \Psi$ has the weak deduction property. By definition, the system $EF + \Psi$ is closed under modus ponens under substitutions by constants. In order to conclude the optimality of the proof system by Theorem 6, it remains to verify the closure of $(EF + \Psi) \cup \Phi$ under substitutions of variables for arbitrary printable sets Φ of tautologies. Going back to the proof of Theorem 6, we observe that instead of proving closure under substitutions of variables, it actually suffices to derive formulas $\varphi(\bar{q})$ from Φ in arbitrary variables $\bar{q}$. For this assume that Φ contains the formula $\varphi(\bar{p})$, and we want to derive the formula $\varphi(\bar{q})$. This can be done as follows: we introduce the extension axioms $p_1 \equiv q_1, \ldots, p_k \equiv q_k$ for all variables p_i in $\bar{p}$. By induction on the formula φ we then prove the equivalence $\varphi(\bar{p}) \equiv \varphi(\bar{q})$ with polynomial-size EF-proofs and finally use modus ponens to conclude with $\varphi(\bar{q})$.

The implication $2 \Rightarrow 1$ was proven in Proposition 5.

Clearly, item 2 implies item 3.

Now we prove the implication $3 \Rightarrow 4$. Let us denote the system $(EF+\Psi)^- \cup \Phi$ by P^- and the system $(EF+\Psi) \cup \Phi$ by P. Assuming the equivalence of P^- and P, it suffices to prove closure under substitutions of variables for the system P, as this property is preserved inside the degree of a proof system. Let π be a P-proof of the formula $\varphi(\bar{p})$, and let $\bar{q}$ be a set of variables distinct from $\bar{p}$. By the

equivalence of P and P^- we have a P^--proof π^- of φ that is only polynomially longer than π. From π^- we will devise a P-proof of $\varphi(\bar{q})$ as follows: If there is an extension axiom $q \equiv \theta$ in π^- with extension variable q from $\bar{q}$, then we rename q in the entire proof π^- to a new variable not occurring in π^-. This does not affect axioms from Φ, as P^--proofs may not use variables from Φ as extension atoms. Let us call this transformed proof π'. Now we construct the proof π of $\varphi(\bar{q})$: π starts with the extension axioms $p_1 \equiv q_1, \dots p_k \equiv q_k$, introducing the original variables $\bar{p}$ as extension atoms. This is followed by the proof π'. By induction on the formula φ we then prove the equivalence $\varphi(\bar{p}) \equiv \varphi(\bar{q})$ with polynomial-size EF-proofs and finally use modus ponens to conclude with $\varphi(\bar{q})$.

The final implication $4 \Rightarrow 2$ follows from Theorem 6 as the systems $(EF \cup \Psi)^-$ even have the classical deduction property by Theorem 3. □

Lemma 13. *Let Φ be a printable set of tautologies and let P be a proof system with the weak deduction property. Then* $(\mathrm{Ref}(P \cup \Phi), \mathrm{SAT}^*) \leq_p (\mathrm{Ref}(P), \mathrm{SAT}^*)$.

Proof. Let Φ be printable and let p be the polynomial from the weak deduction property for P and Φ. Because Φ is printable there exists a polynomial q such that for each number m the set Φ contains at most $q(m)$ tautologies of length $\leq m$. Let $\Phi_m = \Phi \cap \Sigma^{\leq m}$ be the set of these tautologies.

Then $(\mathrm{Ref}(P \cup \Phi), \mathrm{SAT}^*)$ reduces to $(\mathrm{Ref}(P), \mathrm{SAT}^*)$ via the function

$$(\psi, 1^m) \;\mapsto\; \big(\big(\bigwedge_{\varphi \in \Phi_m} \varphi\big) \to \psi, 1^{p(mq(m)+m)}\big) \;.$$

To verify the claim assume that $(\psi, 1^m) \in \mathrm{Ref}(P \cup \Phi)$. Let π be a $P \cup \Phi$-proof of ψ of length $\leq m$. This proof π can use only formulas of length $\leq m$ from Φ of which there are only $\leq q(m)$ many. Hence the tautologies used in the proof π are contained in $\bigwedge_{\varphi \in \Phi_m} \varphi$. Therefore we know that π is also a proof for ψ in the proof system $P \cup \Phi_m$. Using the weak deduction property of P we get a P-proof of size $\leq p(mq(m) + m)$ of $(\bigwedge_{\varphi \in \Phi_m} \varphi) \to \psi$.

Now assume $(\psi, 1^m) \in \mathrm{SAT}^*$. Then $\neg\psi$ is satisfiable and therefore

$$\neg\big(\big(\bigwedge_{\varphi \in \Phi_m} \varphi\big) \to \psi\big) \;=\; \big(\bigwedge_{\varphi \in \Phi_m} \varphi\big) \wedge \neg\psi$$

is also satisfiable because $(\bigwedge_{\varphi \in \Phi_m} \varphi)$ is a tautology. □

Theorem 14. *Let P be a Hilbert-style proof system that simulates the truth-table system and fulfills the following three conditions:*

1. *P is closed under modus ponens.*
2. *For all printable sets of tautologies Φ the proof system $P \cup \Phi$ is closed under substitutions by constants.*

3. P has the weak deduction property.

Then the canonical pair of P is a complete disjoint NP*-pair.*

Proof. Let (A, B) be a disjoint NP-pair. Similarly as in Cook's proof of the NP-completeness of SAT [5], we can construct in polynomial time propositional formulas $\psi_n(\bar{x}, \bar{y})$ such that $\psi_n(\bar{a}, \bar{y})$ is satisfiable if and only if $\bar{a} \in A$. Similarly, we build such propositional formulas $\theta_n(\bar{x}, \bar{z})$ for B. We choose the variables of $\psi_n(\bar{x}, \bar{y})$ and $\theta_n(\bar{x}, \bar{z})$ in such a way that the input variables $\bar{x}$ are the common variables of ψ_n and θ_n, and the auxiliary variables $\bar{y}$ and $\bar{z}$ are distinct. We define the sequence φ_n as

$$\varphi_n = \psi_n(\bar{x}, \bar{y}) \to \neg\theta_n(\bar{x}, \bar{z}) \ .$$

Let P' denote the system $P \cup \{\varphi_n \mid n \geq 0\}$. We first claim that the reduction from (A, B) to the canonical pair of P' is given by

$$a \ \mapsto \ (\neg\theta_{|a|}(\bar{a}, \bar{z}), 1^{p(|a|)})$$

for some suitable polynomial p. To see the correctness of the reduction let first a be an element from A of length n. As ψ_n represents A there exists a witness $\bar{b}$ such that $\psi_n(\bar{a}, \bar{b})$ is a tautological formula. The P'-proof of $\neg\theta_n(\bar{a}, \bar{z})$ proceeds as follows. First we use the axiom $\psi_n(\bar{x}, \bar{y}) \to \neg\theta_n(\bar{x}, \bar{z})$ and substitute the variables $\bar{x}$ and $\bar{y}$ by $\bar{a}$ and $\bar{b}$, respectively, obtaining

$$\psi_n(\bar{a}, \bar{b}) \to \neg\theta_n(\bar{a}, \bar{z}) \ .$$

As $\psi_n(\bar{a}, \bar{b})$ is a true propositional formula without variables we can provide a polynomial-size P-proof for it. This is possible as by assumption P simulates the truth-table system. An application of modus ponens gives a P-proof of $\neg\theta_n(\bar{a}, \bar{z})$, as desired.

Assume now $a \in B$. Then $\neg\neg\theta_{|a|}(\bar{a}, \bar{z}) = \theta_{|a|}(\bar{a}, \bar{z})$ is satisfiable and hence $(\neg\theta_{|a|}(\bar{a}, \bar{z}), 1^{p(|a|)}) \in \mathrm{SAT}^*$.

By Lemma 13 the canonical pair of P' reduces to the canonical pair of P, hence (A, B) is $\leq_p$-reducible to $(\mathrm{Ref}(P), \mathrm{SAT}^*)$. □

Corollary 16. *Assume that for all printable sequences Φ of tautologies the proof systems $F \cup \Phi$ and $EF \cup \Phi$ are equivalent. Then the canonical pair of the Frege proof system is complete for the class of all disjoint* NP*-pairs.*

Proof. To apply Theorem 14 we need to show that $F \cup \Phi$ is closed under substitutions by constants for all printable sets of tautologies Φ. By assumption $F \cup \Phi$ is equivalent to $EF \cup \Phi$, hence it suffices to show this closure property for $EF \cup \Phi$. Given a formula $\varphi(\bar{p}, \bar{q})$ and an $EF \cup \Phi$-proof for it, we construct an $EF \cup \Phi$-proof of an instance $\varphi(\bar{p}, \bar{a})$ with constants $a_1, \ldots, a_k$ substituted for the variables $q_1, \ldots, q_k$ as follows: First we use the extension axioms $q_1 \equiv a_1, \ldots, q_k \equiv a_k$, then we repeat the proof of $\varphi(\bar{p}, \bar{q})$, and finally we show the equivalence $\varphi(\bar{p}, \bar{q}) \equiv \varphi(\bar{p}, \bar{a})$ by induction on the formula φ and using the formulas $q_i \equiv a_i$. This yields the $EF \cup \Phi$-proof of $\varphi(\bar{p}, \bar{a})$. □