

# A lower bound on the size of resolution proofs of the Ramsey theorem

Pavel Pudlák\*

December 7, 2011

## Abstract

We prove an exponential lower bound on the lengths of resolution proofs of propositions expressing the finite Ramsey theorem for pairs.

Assuming that  $n \geq R(k)$ , where  $R(k)$  denotes the Ramsey number, the Ramsey theorem for pairs and two colors,  $n \rightarrow (k)_2^2$ , is presented by the following unsatisfiable set of clauses. The variables are  $x_{ij}$ , for  $1 \leq i < j \leq n$ . The clauses are  $\bigvee_{i,j \in K} x_{ij}$  and  $\bigvee_{i,j \in K} \neg x_{ij}$ , for all sets  $K \subseteq \{1, \dots, n\}$ ,  $|K| = k$ . The corresponding tautology will be denoted by  $RAM(n, k)$ .

The Ramsey theorem was proposed as a hard tautology by Krishnamurthy in [6]. He studied the tautology  $RAM(R(k), k)$  and proved a lower bound  $R(k)/2$  on the width of resolution proofs (see also [5]). This implies an exponential lower bound on the tree-like resolution proofs. Krajíček proved an exponential lower bound on this tautology by reducing the proofs of the pigeonhole principle to it, [4]. The problem with this tautology is that we do not know what is  $R(k)$ . This prevent us from proving an upper bound on the proof complexity of this tautology. Therefore researchers focused on the tautology  $RAM(n, k)$  for  $k = \lfloor \frac{1}{2} \log n \rfloor$  (all logarithms are to the base 2 in this paper). This tautology is provable in a bounded depth Frege system, see [7, 4]. For this tautology, Krajíček proved an exponential lower bound on tree-like resolution proofs with conjunctions of logarithmic size, [3]. The complexity of unrestricted resolution proofs with conjunctions of logarithmic size proofs of  $RAM(n, \lfloor \frac{1}{2} \log n \rfloor)$  is still an open problem. An exponential lower bound on such proofs would have interesting consequences in proof complexity and bounded arithmetic. In particular it would give a separation of the relativized theories  $T_2^2$  and  $T_2^3$  by a  $\forall \Sigma_1^b$  sentence (see [2, 1]). In this paper we prove an exponential lower bound on unrestricted resolution proofs.

**Theorem 1** *Resolution proofs of  $RAM(n, \lfloor \frac{1}{2} \log n \rfloor)$  have size at least  $2^{n^{\frac{1}{4}-o(1)}}$ .*

---

\*Institute of Mathematics, Academy of Sciences, Prague, and Institute of Theoretical Computer Science, Prague, e-mail: pudlak@math.cas.cz. Partially supported by Institutional Research Plan No. AV0Z10190503, project No. 1M0021620808 of MŠMT ČR and grant IAA100190902 of GA AV ČR.

*Proof.* We will use the following bound for the sum of Bernoulli variables  $X = \sum_{i=1}^r X_i$  with  $\Pr(X_i = 1) = q$ .

$$\Pr(X \geq cr) \leq q^{cr} 2^{H(c)r},$$

where  $H$  is the entropy function, which follows from  $\binom{r}{cr} \leq 2^{H(c)r}$ .

Let  $\delta > 0$ . We will prove a lower bound  $2^{\Omega(n^{\frac{1}{4}-\delta})}$ . Let  $k = \lfloor \frac{1}{2} \log n \rfloor$  and  $m = \lfloor n^{\frac{1}{4}-\delta} \rfloor$ . In the rest of the proof we will ignore rounding.  $\varepsilon$  and  $p$  will be sufficiently small constant whose values will be determined later.

Let  $\rho$  be the random restriction that sets  $x_{ij}$  to 0 with probability  $\frac{p}{2}$ ,  $x_{ij}$  to 1 with probability  $\frac{p}{2}$  and leaves  $x_{ij}$  free with probability  $1 - p$ . Let a proof  $P$  be given and let  $S$  be its size. After hitting  $P$  by  $\rho$ , some clauses become true and we delete them. The others may have reduced length, because some literals become false. We will denote by  $P_\rho$  the reduced proof. The probability that  $P_\rho$  contains a clause of length  $> \frac{m}{2}$  is less than

$$S(1 - \frac{p}{2})^{\frac{m}{2}} = S \cdot 2^{\log(1 - \frac{p}{2}) \frac{1}{2} n^{\frac{1}{4}-\delta}}.$$

Hence, if  $S < 2^{-\log(1 - \frac{p}{2}) \frac{1}{2} n^{\frac{1}{4}-\delta} - 1}$ , the probability is  $< \frac{1}{2}$ . We will assume this and show a contradiction.

Consider an initial clause. The probability that  $\rho$  sets at least  $\varepsilon \binom{k}{2}$  literals of the clause is at most

$$\left(\frac{p}{2}\right)^\varepsilon \binom{k}{2} 2^{H(\varepsilon) \binom{k}{2}} = 2^{(\log \frac{p}{2} \cdot \varepsilon + H(\varepsilon)) \binom{k}{2}} \leq 2^{\frac{1}{8} (\log \frac{p}{2} \cdot \varepsilon + H(\varepsilon) + o(1)) (\log n)^2}.$$

Hence the probability that this happens for at least one initial clause is at most

$$\begin{aligned} & 2^{\frac{1}{8} (\log \frac{p}{2} \cdot \varepsilon + H(\varepsilon) + o(1)) (\log n)^2} \cdot 2^{\binom{n}{k}} \leq \\ & 2^{\frac{1}{8} (\log \frac{p}{2} \cdot \varepsilon + H(\varepsilon) + o(1)) (\log n)^2} n^{\frac{1}{2} \log n} = \\ & 2^{\frac{1}{8} (\log \frac{p}{2} \cdot \varepsilon + H(\varepsilon) + \frac{1}{2} + o(1)) (\log n)^2}. \end{aligned}$$

If  $p$  is sufficiently small w.r.t.  $\varepsilon$ , then the term  $\log \frac{p}{2} \cdot \varepsilon + H(\varepsilon) + \frac{1}{2} + o(1)$  is negative for large  $n$ . Hence, for such a  $p$  and large  $n$  the probability is  $< \frac{1}{2}$ .

Thus there exists  $\rho$  such that in the proof  $P_\rho$

1. every clause has length at most  $m/2$ ;
2. every initial clause has at least  $(1 - \varepsilon) \binom{k}{2}$  variables.

Following an idea of Krajíček [3], we will use a random graph  $G$  on  $m$  vertices to show that such a proof does not exist. While Krajíček only needed that  $G$  does not have a homogeneous set of size  $k$ , we will need more: the number of edges on every subset of size  $k$  is strictly between  $\varepsilon \binom{k}{2}$  and  $(1 - \varepsilon) \binom{k}{2}$ . (This is why we need  $m$  larger than  $n^{\frac{1}{4}}$ .) The probability that this condition fails for one fixed set of size  $k$  is at most

$$2 \cdot 2^{-(1-\varepsilon) \binom{k}{2}} \binom{\binom{k}{2}}{\varepsilon \binom{k}{2}} \leq 2 \cdot 2^{(-1+\varepsilon+H(\varepsilon)) \binom{k}{2}} = 2^{(-1+\varepsilon+H(\varepsilon)+o(1)) \frac{(\log n)^2}{8}}.$$

The probability that there exists a set of size  $k$  for which it fails is at most

$$\begin{aligned} & 2^{(-1+\varepsilon+H(\varepsilon)+o(1))\frac{(\log n)^2}{8}} \cdot \binom{m}{k} \leq \\ & 2^{(-1+\varepsilon+H(\varepsilon)+o(1))\frac{(\log n)^2}{8}} \cdot m^k \leq \\ & 2^{\frac{-1+\varepsilon+H(\varepsilon)+o(1)}{8}(\log n)^2 + \frac{1}{2} \log n \cdot (\frac{1}{4}-\delta) \log n} = \\ & 2^{\frac{\varepsilon+H(\varepsilon)-4\delta+o(1)}{8}(\log n)^2}. \end{aligned}$$

Hence if we choose  $\varepsilon > 0$  so that  $\varepsilon + H(\varepsilon) < 4\delta$ , the exponent will be negative for sufficiently large  $n$ . Thus we obtain the auxiliary graphs.

Now, as in Krajíček's proof, construct a path in  $P_\rho$  from the empty clause to an initial clause such that for every clause  $C$  on the path the following condition is satisfied. There exists a bijection between the indices of the variables of  $C$  and vertices of the graph  $G$  such that if  $x_{ij}$  (or  $\neg x_{ij}$ ) is a literal in  $C$  and  $u, v$  are vertices corresponding to  $i, j$ , then  $(u, v)$  is not an edge (respectively is an edge in  $G$ ). We can construct this path, because every clause has at most  $m/2$  literals. However the latter condition cannot be satisfied by the initial clauses, because each initial clause has at least  $(1 - \varepsilon)\binom{k}{2}$  literals of the same kind, but in  $G$  there are less than  $(1 - \varepsilon)\binom{k}{2}$  pairs  $(u, v)$  of the same kind (edges or non-edges) on every  $k$ -element subset. This contradiction finishes the proof. ■

## References

- [1] K. Aehlig, A. Beckmann, A remark on the induction needed to prove the Ramsey principle. 2006, unpublished.
- [2] M. Chiari and J. Krajíček, Lifting independence results in bounded arithmetic, *Archive for Mathematical Logic*, 38(2), (1999), pp. 123-138.
- [3] J. Krajíček, On the weak pigeonhole principle, *Fundamenta Mathematicae*, 170(1-3), (2001), pp. 123-140.
- [4] J. Krajíček, A note on propositional proof complexity of some Ramsey-type statements, *Archive for Mathematical Logic*, 50(1-2), (2011), pp. 245-255.
- [5] B. Krishnamurthy and R.N. Moll, Examples of hard tautologies in the propositional calculus. In: *Proc. 13th ACM Symposium on Theory of Computing* (1981), pp. 28-37.
- [6] B. Krishnamurthy, Examples of hard tautologies and worst-case complexity results. Ph.D. Thesis, Univ. of Massachusetts, 1981.
- [7] P. Pudlák: Ramsey's Theorem in Bounded Arithmetic. In: *Proc. Computer Science Logic'90*, eds. E. Borger, H. Kleine Buning, M.M. Richter, W. Schonfeld, LNCS 533, Springer-Verlag, 1991, pp. 308-317.