Basic Research in Computer Science

BRICS RS-94-21

Count(q) does not imply Count(p)

Søren Riis

BRICS Report Series **RS-94-21**

ISSN 0909-0878 **July 1994**

See back inner page for a list of recent publications in the BRICS Report Series. Copies may be obtained by contacting:

**BRICS
Department of Computer Science
University of Aarhus
Ny Munkegade, building 540
DK - 8000 Aarhus C
Denmark**

**Telephone: +45 8942 3360
Telefax: +45 8942 3255
Internet: BRICS@daimi.aau.dk**

Count(q) does not imply Count(p)

Søren Riis*

BRICS†

August 1993
Revised May 1994

Abstract

I solve a conjecture originally studied by M. Ajtai. It states that for different primes q, p the matching principles Count(q) and Count(p) are logically independent. I prove that this indeed is the case. Actually I show that Count(q) implies Count(p) exactly when each prime factor in p also is a factor in q.

1 The logic of elementary counting

"She loves me, she loves me not, she loves me,..." The final answer does not depend on the order in which the leaves are pulled of. Every child who is familiar with the process of counting knows that. The underlying logical principle states that a set A has a well-defined cardinality modulo 2. Yet, the Count(2) principle can fail in quite strong systems of Arithmetic [2],[3]. Similarly for the counting principle modulo p (=Count(p)) where she can be in p states of mind.

This is very difficult to visualise. In 1962 Cohen invented the famous technique of forcing. He used the method to show the independence of the continuum conjecture. Inspired by these ideas Ajtai showed that the elementary pigeon-hole principle need not hold in strong systems of Arithmetic [2]. Ajtais result was a major break through. The main novelty was the mixture of forcing and powerful probabilistic techniques.

The Count(q) versus Count(p) problem has various formulations and variants. The most famous variant is from circuit complexity theory [13]. It asks (in the

*This work was initiated at Oxford University England.

†Basic Research in Computer Science, Centre of the Danish National Research Foundation.

base case) whether there exist bounded depth, polynomial size circuits which counts the number of 1's (in the input string) modulo p. This was answered (negatively) independently in [13] and [1]. Later [14] gave a near optimal exponential lower bound. The question becomes particular challenging if we also allow gates which can count modulo q. In [27] the case was settled for different prime numbers q and p. The general classification is still open. It has been conjectured that the answer is positive exactly when q contains all prime factors in p [5]. However even the simple case where $q = 6$ and $p = 5$ has now been open for more than five years (PC. Haastad, Krajicek, Pudlak).

Ajtais version of the problem is technically more involved ('presumably more difficult' to cite [16]). One formulation (given in [10]) concerns the question whether for different primes q and p there exist arithmetical models $\mathbb{M}$, which satisfies the Count(q) principle, but which does not satisfies the Count(p) principle? The method in [27] is not sufficient. Still circuit complexity (especially the method of collapsing circuits by use of random evaluations) is of major importance [16], [24].

In this paper I answer Ajtais question. Actually I give a complete classification. It agree with what has been conjectured for the circuits. I.e. the answer is positive exactly when all prime factors in p belong to q.

1.1 Non-standard Arithmetic

It it well known that there are interesting and useful geometrical structures in which the (obvious) parallel postulate fail. The models I construct in this paper (and the ones constructed in [2] and [3]) suggests that there exist a similar phenomenon in Arithmetic! As an illustration of this idea suppose that we live in some "non-Euclidean" Arithmetical world $\mathbb{M}$. Locally the universe $\mathbb{M}$ agrees with the real universe. Statements concerning concrete finite objects have unaltered truth value. However, globally i.e. when it comes to the behaviour at infinity, there can be disagreement. Thus even though each concrete ("finite") set A of numbers has a well defined cardinality this property might not be globally valid.

To illustrate the idea further suppose (as an example) that the Count(2) principle is valid in $\mathbb{M}$. What is the status of the Count(4) principle? Or slightly less general is it possible that there exists a " number" n' such that the ordered set $\{1, 2, ..., 4n' + r\}$ of "numbers" can be divided into disjoint 4-element subsets, and $r \in \{1, 2, 3\}$?

Consider the following argument: We want to show (reasoning inside $\mathbb{M}$) that a set of numbers of the form $\{1, 2, ..., n\}$ can be divided into a collection of disjoint 4 element subsets only when n is divisible by 4. Suppose that on the contrary some interval $\{1, 2, .., 4n' + r\}$, $r \in \{1, 2, 3\}$ can be divided into a collection $\mathcal{P}$ of disjoint 4

element subsets. The case where $r = 1$ or $r = 3$ can be excluded for trivial reasons. To see this sub-divide each 4-element subset into two 2-element subsets. This induces a partitioning of $\{1, 2, ..., 4n' + r\}$ into disjoint 2-element subsets violating the Count(2) principle.

The case where $r = 2$ require a more involved argument [3] . Consider all pairs of $\{1, 2, ..., 4n' + 2\}$. It only requires a quite weak part of arithmetic to prove that these pairs are in 1-1 correspondence with $\{1, 2, ..., \binom{4n'+2}{2}\}$. And even less Arithmetical assumptions to show that $\binom{4n'+2}{2}$ is an odd number. To get a contradiction (by violating the Count(2) principle) it suffices to show that the partitioning $\mathcal{P}$ induces a partitioning $\mathcal{R}$ of all pairs of $\{1, 2, ..., 4n'+2\}$ into disjoint 2 element sets. Consider the pair $\{v_1, v_2\}$. If both v_1 and v_2 belongs to the same 4-element subset $\{v_1, v_2, v_3, v_4\} \in \mathcal{P}$ let $\{\{v_1, v_2\}, \{v_3, v_4\}\} \in \mathcal{R}$. Otherwise suppose $v_1 \in \{w_1, w_2, w_3, w_4\} \in \mathcal{P}$ and $v_2 \in \{\tilde{w}_1, \tilde{w}_2, \tilde{w}_3, \tilde{w}_4\} \in \mathcal{P}$. All elements are listed after size. So there are unique $i, j \leq 4$ such that $v_1 = w_i$ and $v_2 = \tilde{w}_j$. If $i \neq j$ let $\{\{v_1, v_2\}, \{\tilde{w}_i, w_j\}\} \in \mathcal{R}$. If $i = j$ let $\{\{v_1, v_2\}, \{w_{i'}, \tilde{w}_{i'}\}\} \in \mathcal{R}$ where $1' = 2,\ 2' = 1,\ 3' = 4$ and $4' = 3$. This completes the argument.

To summarise: We considered a structure S_I constructed from $I := \{1, 2, ..., n\}$. In this concrete case the structure consisted of all pairs of $\{1, 2, .., 4n' + 2\}$. This structure S_I had the property that partial partitions of $\{1, 2, ..., 4n'+2\}$ into 4 element subsets induced (in a flexible way) pairings of the elements in S_I. And crucially the structure S_I contained an odd number of elements. One could try to modify the type of argument to the case where for example $q = 2$ and $p = 3$. At an early stage in this research J.Krajicek showed me some ingenious constructions attempting show that Count(3) was a consequence of Count(2). However as J.Krajicek pointed out careful calculations always seems to give the wrong parities. Irrespectively of the ingenuity however clever the structures S was constructed, it always seemed to end up containing an even number of elements. So it seemed that strong "forces" wanted Count(2) and Count(3) to be independent.

In retrospect this is of course a simple consequence of the general classification. It is a direct consequence of the fact that the Count(2) and the Count(3) principles are independent in powerful Arithmetical structures. The first step in showing this was obtained when I reduced the general Count(q) versus Count(p) problem to the study of "generic systems". And by introducing a certain refinement technique I was able to reduce the Count(q) versus Count(p) problem to questions concerning forests of specially labelled trees.

[3] I learned this type of argument from J. Krajicek and P.Pudlak.

1.2 A forest containing 16821302548060 trees

The first main result in the paper links the Count(q) versus Count(p) questions to a class of purely combinatorial problems.

Suppose $T_1, T_2, ..., T_u$ is a collection of specially labelled trees (i.e. a forest). Suppose that each type of branch appears 0 modulo q times. Does q divide u ? This of course depends on how the trees are labelled. I consider labels of a type which is determined by two numbers (p, n). A naive conjecture states that (besides some trivial counter examples) q indeed divides u.

It turns out that there exist "exceptional" forests which violate this naive conjecture. As an example when $q = 2$ and $p = 4$, I show that there is a forest where each type of branch appears an even number of times. However the forest contains 635 trees (which is an odd number). When $q = 3$ and $p = 9$ there are also exceptional forests. In these each type of branch appears 0 modulo 3 times, yet the number of trees is not divisible by 3. The smallest concrete example I have found contains 16821302548060 trees.

The first main result in the paper shows that the existence of such exceptional forests and the existence of (non-trivial) implications between Count(q) and Count(p) are two sides of same coin. The two examples correspond to the fact that Count(2) implies Count(4) and that Count(3) implies Count(9). It turns out that Count(q) implies Count(p) in systems of Bounded Arithmetic when all prime factors in p appears in q. According to my first main result *a priori* there must exist exceptional forest for all such q and p. Actually I follow an alternative route. I show how one can construct proofs (in systems of Bounded Arithmetic) of Count(p) from Count(q) directly based on such forests.

Early in this research the exceptional forests caused a major complication. At that stage all my attempts to collapse forests to particularly nice normal forms failed. The probabilistic arguments did not quite work. Essentially the exceptional forests was the only obstacle. First when I managed to isolate these asymptotically, I was able to complete the analysis.

At present I do not have a complete picture of all exceptional forests. However it turns out that the asymptotic classification in this paper is sufficiently strong to provide a complete solution of the Count(q) versus Count(p) problem in the base-case (i.e. when the terms in underlying language have polynomial growth rate).

1.3 Why are these problems important

The counting principles themselves are of course trivial. Or more specifically they hold in the category of finite sets. There are various reasons to examine these elementary counting principles.

First of all they play an important role in Bounded Arithmetic. As already pointed out in non-Euclidean geometry the (obvious) parallel postulate is not assumed to hold. Bounded Arithmetic resembles this phenomenon. Here the (obvious) induction axiom scheme is restricted. Which parts of number theory holds in models of Bounded Arithmetic? This question was first studied intensively by J.Paris, A.Wilkie and many of their students. Many basic number theoretical facts are provable in system of Bounded Arithmetic [7]. Other facts require new proofs. I believe that Bounded Arithmetic raises an important and serious possibility. It seems that the provability (in specific systems of Bounded Arithmetic) of elementary number theoretical statements as a rule could be intimately linked to deeper number theoretical problems/theorems. At present there are only sporadic suggestions of this. One such is that if a certain fragment (often denoted by S_2^1 [8]) proves that the set of prime numbers is in NP (this can be proved in ordinary Arithmetic), then the prime numbers must actually be polynomial time recognisable. At present this is only known conditionally by assuming the validity of the General Riemann Hypothesis [17]. A stronger fragment (often denoted S_2) are know to show the infinitude of the set of prime numbers. This fact goes hand in hand with Sylvesters prime number theorem [18]. Besides this consider the quantifier elimination phenomenon (the strength of eliminating logic!). Clearly Bounded Arithmetic does not have quantifier elimination. However, one might still be able to eliminate many of its logical-like features. It should be possible to get our hands on the underlying unifications features arising from the induction schema. So perhaps Bounded Arithmetic is tight up with the prestigious discipline of number theory (see [17] for a further discussion).

In any case the work by [18] and later [7] illustrates the central role of elementary counting principles in Bounded Arithmetic. In general the status of the elementary counting principles in models of Bounded Arithmetic seems to be a very deep problem. The paper solves this in the special case where all terms of the underlying language have polynomial growth rate, and contain at least one unspecified function or relation symbol [4] .

Second, systems of Bounded Arithmetic are linked to "low complexity reasoning". One fundamental problem is to clarify the relation between automated versus

[4] One of the major challenges is to understand the case where each function and relation are fully specified.

intelligent reasoning. It seems natural to suggest that automatic reasoning (when this implemented in praxis) is only able to give a proper representation of objects of low complexity. The elementary process of counting introduces unpleasantly high complexity. A computation involving a counting task might (asymptotically) require exponentially many steps as a function of the length of the input. In practice this very soon becomes intractable for computers. Thus in low complexity reasoning we cannot assume that we are be able to count. To verify that the cardinality of a set A is unique, we would have verify that all bijections $f : A \to \{1, 2, ..., m\}$ requires the same m. This is computationally intractable even for small sets A.

We can view Count(p) as a spark of pure intelligence. The paper shows that (mechanical) systems, more specifically systems which reason (using first or even second order logic) within finite structures in certain cases are not able to establish any link between Count(q) and Count(p).

Finally another (related) problem is to examine the efficiency of propositional proof systems. This type of problems has already been studied intensively in the literature [2], [3], [11], [16], [19], [21], [24]. In S.Cook and Recknow [11] it was shown that the efficiency of propositional proof systems is a natural way of studying the NP versus co-NP problem. Then later [19] these problems was linked to Bounded Arithmetic. And then in [2] the problems was shown also to be tight up with methods and problems from circuit complexity. Recently a fascinating 'ultra filter construction' by Razborov [22] even suggest links to higher set-theory. In any case the study of the complexity of elementary counting provides some of the strongest known results in the field of circuit complexity.

The growth rate of the terms in the underlying language L of Bounded Arithmetic is a very precise measure of the axiom systems "intelligence". Most number theory is provable when the language contains function of exponential growth rate. At this level we have real intelligence. Ideally we would like to study what happens to the relative strength of the counting principles, when the intelligence of the underlying system approaches the level of real intelligence. The paper allows us to do this in principle. However, until we have a general picture of the exceptional forests, this problem remains open.

1.4 The main results

In the following discussion let L be a countable first order language. Assume that L contains function symbols for the basic arithmetical operations '$+$' and '$\cdot$'. Also assume that the behaviour of terms and (the specified) relations are specified through

a suitable set Ψ_L of purely universal axioms. And assume that L contains at least one unspecified relation symbol.

An axiom system (= $I\Delta_0(L)$ or just $I\Delta_0$ when L is clear from the context) of Bounded Arithmetic consists of the axioms Ψ_L together with the celebrated induction axiom schema, $(\theta(0) \wedge \forall x\ (\theta(x) \Rightarrow \theta(x+1))) \rightarrow \forall z\ \theta(z)$. However, in Bounded Arithmetic (unlike in ordinary Arithmetic), we require all quantifiers in each θ to be bounded by terms in the language L. More specifically, each quantifier is required to appear in the context $\forall x(x \leq t \rightarrow ...$ or $\exists x(x \leq t \wedge ....$

The elementary pigeon-hole principle (=PHP$_p$ $p \in \mathbb{N}$) states (in one of its many formulations) that for no n do there exists a bijection from $\{1, 2, ..., n\}$ onto $\{1, 2, ..., n+p\}$. More specifically, the Δ_0-PHP$_p$ axiom schema states (for each bounded formula $\theta(x, y)$) that,

$$\forall z\ (\neg\forall x \leq z \exists! y \leq z + p\ \theta(x, y, z) \vee \neg\forall y \leq z + p \exists! x \leq z\ \theta(x, y, z)).$$

A weak form of the pigeon-hole principle is obtained by only considering monotone bijections. It is not hard to show that this form of the pigeon hole principle is equivalent to the usual induction principle.

The Count(p) principle (for a fixed number $p \in \mathbb{N}$) states that if $\{1, 2, ..., n\}$ is divided into disjoint subsets each containing exactly p elements, then p divides n. More specifically, the Δ_0-Count(p) principle is the schema,

$$\forall z\ ((\forall x_1 \leq z \exists! x_2, ..., x_p \leq z\ (x_2 < x_3 < ... < x_p \wedge \theta(x_1, x_2, ..., x_p)$$

$$\wedge \neg x_1 = x_2 \wedge .... \wedge \neg x_1 = x_p)) \rightarrow \exists y\ y \cdot p = z).$$

In the first section I show,

Theorem *Assume that $p \geq 2$. Let L be any language where all terms have sub-exponential growth rate. Then there exists a model $\mathbb{M}^*$ in which*

(1) *The* Count*(p) principle fails.*

(2) *All Δ_0-pigeon-hole principles holds.*

A similar result was proved by Ajtai in [3], but only in case where all terms was assumed to have polynomial growth rate. Later Krajicek, Pudlak and Wood [16] made a major improvement in the underlying probabilistic method. They showed the theorem (in essence) in the case where (2) is replaced by the Δ_0-induction principle (or equivalent the Δ_0-pigeon-hole principle for monotone bijections). The theorem has been shown independently by Beame and Pitassi [21]. Actually they showed a different (but essentially equivalent) result concerning the length of proofs in propositional proof systems.

In section 2, the next section I construct the model $\mathbb{M}^*$. And in the next two sections I show that $\mathbb{M}^*$ has the required properties. Actually in section 4 it is shown that,

Theorem *Besides* (1) *and* (2) *the model* $\mathbb{M}^*$ *satisfies the* Δ_0-Count(q) *principle exactly (under some weak extra assumptions) when there are no forest* $T_1, T_2, ...T_u$ *of* (p,n)*-labelled trees where all branches appear* 0 *modulo* q *times, but* $u \neq 0$ *modulo* q*.*

The precise formulation of the result link the growth rate of terms in the underlying language L *to an extra condition on the asymptotic hight of the trees.*

In section 5 I develop a general method to produce exceptional forests. It is shown that exceptional forests exist (for q and p) when all prime factors in p divides q. Furthermore the construction of such forests can be carried out inside any model of Bounded Arithmetic, so we get the following positive part of the classification.

Proposition *Let* $\mathbb{M}^*$ *be a model of Bounded Arithmetic in which the* Δ_0-Count(q) *principle holds. If all prime factors in* p *divides* q*, then* $\mathbb{M}^*$ *satisfies the* Δ_0-Count(p) *principle.*

In section 6 I return to the main problem. This is to show that Count(p) not is a logical consequence of Count(q) when p contains a prime factor not in q. This is shown (in the case all terms have polynomial growth rate) by showing

(1) For each exception q-forests $T_1, T_2, ..., T_u$ of (p,n) trees, one can construct an exceptional q-forest $T'_1, T'_2, .., T'_{u'}$ of labelled trees related to the PHP$_{q^k}$-principle. No tree in this new forests has higher hight than all trees in the old forest.

(2) Suppose that $T'_1, T'_2, ..., T'_{u'}$ is an exceptional q-forest of decision trees for the PHP$_{q^k}$-principle. Then at least one of the trees has hight $\geq k$.

Combining this we get,

Theorem *Suppose that* q *and* p *are fixed. Suppose that* p *contains a prime factor which does not divide* q*. For each* k *there exists* n_k *such that for each* $n \geq n_k$ *there are no exceptional* q*-forests of* (p,n)*-labelled trees.*

Finally in section 7 I combine this result with theorem 1.4 and proposition 1.4. This gives the full classification,

Main Theorem (formulation 1) *Let* $\mathcal{T}$ *be any system of Bounded Arithmetic over some countable language* L*. Suppose that* L *in addition to containing the language of arithmetic also contains at least one undefined relational symbol. Suppose that all terms* t *in* L *have polynomial growth rate. Then for all* $q, p \in \mathbb{N}$ *the following are equivalent:*

(a) *there exists a model* $\mathbb{M}$ *of* $\mathcal{T}$ *in which* Count(q) *holds and* Count(p) *fails.*

(b) *All prime factors in* p *divide* q*.*

The result has various essentially equivalent formulations.

Main Theorem (formulation 2) *Let* $\mathbf{ACA}^{\mathrm{top}}$ *be the following modification of the celebrated system* **ACA***. As* **ACA** *the system* $\mathbf{ACA}^{\mathrm{top}}$ *has the full arithmetical comprehension. And it is equipped with the full induction axiom for sets. The "only" difference between this system and the normal second order Arithmetic is that the basic universal axioms are modified so the that universe contains a largest (unspecified) number* c*. All basic operations are modified (e.g.* $c + 1 = c$*). Any list of purely universal axioms might also be added. Suppose that the axiomatisation is non-trivial e.g. allows an infinite model. Then the following are equivalent:*

(a) Count(p) *holds in all structures which satisfies* $\mathbf{ACA}^{\mathrm{top}}$ *and the* Count(q) *principle.*

(b) *All prime factors in* p *appear in* q*.*

Another formulation states that,

Main Theorem (formulation 3) *Let* $\mathcal{P}$ *be one of the usual textbook systems in Hilbert style propositional logic. Let* $\mathrm{Count}_{\mathrm{scheme}}(q)$ *denote the substitution axiom scheme which arrives from the canonical Booleanization of the* Count(q) *principle. Let* $\mathcal{P}' := \mathcal{P}+ \mathrm{Count}_{\mathrm{schema}}(q)$*. Then there are polynomial-size bounded depth* $\mathcal{P}'$*-proofs of* Count(p) *exactly when all prime factors in* p *divide* q*.*

In all formulations the negative part of the classification has a heuristic explanation. The analysis shows that when k becomes large, it becomes arbitrarily difficult [5] (but as it turns out never impossible) to show PHP_{q^k} from Count(q). On the other hand if p contains a prime factor not in q it is uniformly (in k) easy to show PHP_{q^k} from Count(p). So Count(p) is not a consequence (a bounded depth polynomial-size consequence in formulation 2) of Count(q) in this case.

Finally I mention the recent and independent developments in [4] and [6].

2 Constructing the model

2.1 Translating formulas into circuits

Let $\mathbb{M}$ be a countable non-standard model of Th$(\mathbb{N})$ over a countable first order language L (which extends the language of arithmetic). Let $p \in \omega$, $p \geq 2$ and let

[5] Measured by the hight of the corresponding forest.

$I := \{1, 2, ..., n\} \subseteq \mathbb{M}$, $n \in \mathbb{M} \setminus \omega$ be fixed. Here ω denote the set of standard integer in $\mathbb{M}$. As is common a set $A \subseteq \mathbb{M}$ is said to be $\mathbb{M}$-definable if there exists $m \in \mathbb{M}$ such that $a \in A$ if and only if a belong to the sequence coded by m.

Definition 2.1.1 For each $A \subseteq I$ with $\mid A \mid = p$ we introduce a variable p_A. The set of all such variables is denoted by $\mathrm{VAR}_{I,p}$. ♣

Definition 2.1.2 A *(Boolean) circuit* π (with input variables in X) of *size* $s(\pi)$ and *depth* $d(\pi)$ is defined inductively as follows:

(a) The constants '0' and '1' are circuits with $s(\text{'1'}) = s(\text{'0'}) = d(\text{'1'}) = d(\text{'0'}) = 1$.

(b) Each $p \in X$ is a circuit with $s(p) = d(p) = 1$.

(c) If π is a circuit, then $\neg\pi$ is a circuit with $s(\neg\pi) = s(\pi)+1$ and $d(\neg\pi) = d(\pi)+1$.

(d) If $\pi_1, \pi_2, ..., \pi_r$ are circuits, then $\wedge_j \pi_j$ and $\vee_j \pi_j$ are circuits with $s(\wedge_j \pi_j) = s(\vee_j \pi_j) = 1 + \Sigma_j\ s(\pi_j)$ and $d(\wedge_j \pi_j) = d(\vee_j \pi_j) = 1 + \max_j d(\pi_j)$. ♣

Definition 2.1.3 Let $B_d(X)$ denote the (Boolean) circuits π with input variables X of depth $d(\pi) \leq d$. Let $B_{<\omega}(X) := \cup_{d \in \omega}\ B_d(X)$. ♣

Definition 2.1.4 For $\psi \in B_{<\omega}(\mathrm{VAR}_{I,p})$ and $\rho : \mathrm{VAR}_{I,p} \to \{0,1\}$ (not required to be $\mathbb{M}$-definable), we define the *truth-table evaluation* ψ^ρ inductively as follows:

(a) $\text{'0'}^\rho = 0, \quad \text{'1'}^\rho = 1$.

(b) $p_A^\rho = 1$ iff $\rho(p_A) = 1$.

(c) $(\neg\pi)^\rho = 1$ iff $\pi^\rho = 0$.

(d) $(\wedge_j \pi_j)^\rho = 1$ iff $\pi_j^\rho = 1$ for all j.

(e) $(\vee_j \pi_j)^\rho = 1$ iff $\pi_j^\rho = 1$ for some j. ♣

Let $L_{\mathbb{M}}$ be L extended by a constant c_a for each $a \in \mathbb{M}$. Let $L_{\mathbb{M}}(P)$ be $L_{\mathbb{M}}$ extended with an p-ary relation symbol. There exists a canonical translation of Bounded $L_{\mathbb{M}}(P)$-sentences into circuits in $B_{<\omega}(\mathrm{VAR}_{I,p})$.

Definition 2.1.5 For each sentence $\psi \in L_{\mathbb{M}}(P)$ we define $\epsilon_\psi \in B_{<\omega}(\mathrm{VAR}_{I,p})$ inductively as follows:

(a) For any k-ary relation symbol ($\neq P$): $\epsilon_{R(a_1,...,a_k)}$:='1' if $\mathbb{M} \models R(a_1, .., a_k)$, '0' otherwise.

(b) $\epsilon_{P(a_1,...,a_p)} := p_A$ if $A = \{a_1, ..., a_p\} \subseteq I$ and $\mid A \mid = p$, '0' otherwise.

(c) $\epsilon_{\neg\pi} := \neg\epsilon_\pi$.

(d) $\epsilon_{\pi \vee \pi'} := \epsilon_\pi \vee \epsilon_{\pi'}$

(e) $\epsilon_{\pi \wedge \pi'} := \epsilon_\pi \wedge \epsilon_{\pi'}$

(f) $\epsilon_{\exists x(x \leq u \wedge \theta(x,u))} := \vee_{a \leq u}\ \epsilon_{\theta(a,u)}$.

(g) $\epsilon_{\forall x(x \leq u \to \theta(x,u))} := \wedge_{a \leq u}\ \epsilon_{\theta(a,u)}$. ♣

Notice that if $\psi \in L_{\mathbb{M}}(P)$ has $\leq d$ quantifiers, all bounded by $t \in \mathbb{M}$, and ψ contains k logical connectives, then $s(\epsilon_\psi) \leq kt^d$ and $d(\epsilon_\psi) \leq d + k$.

Lemma 2.1.6 *Suppose that P is a partitioning of $\{1, 2, ..., n\}$ into disjoint classes each containing exactly p elements. Let $\rho_P : \mathrm{VAR}_{I,p} \to \{0, 1\}$ be defined by $A \in P \leftrightarrow \rho_P(p_A) = 1$. Then for $\psi \in L_{\mathbb{M}}(P)$ the following statements are equivalent:*

(a) $(\mathbb{M}, P) \models \psi$.

(b) $(\epsilon_\psi)^{\rho_P} = 1$.

Proof: Induction on the number of logical constants in ψ. □

2.2 The forcing set up

As above let $\mathbb{M}$ be a countable non-standard model of $\mathrm{Th}(\mathbb{N})$ over a countable first order language L which extends the language of arithmetic. We have fixed $p \geq 2$ and $I := \{1, 2, ..., n\} \subseteq \mathbb{M}$, $n \in \mathbb{M} \setminus \omega$. Let $L_{\mathbb{M}}$ and $L_{\mathbb{M}}(P)$ be defined as above.

Definition 2.2.1 We say that ρ is a *partial p-partitioning* iff

(a) $\forall A \in \rho \ A \subseteq I$.

(b) $\forall A \in \rho \ \mid A \mid = p$.

(c) $\forall A, B \in \rho \ A \neq B \to A \cap B = \emptyset$.

Let $\mathrm{Set}(\rho) := \cup_{A \in \rho} A \subseteq I$. ♣

Definition 2.2.2 For $k \in \mathbb{N}$ let

$$\mathcal{P}_k := \{\rho : \ \rho \text{ is a partial } p\text{ −partitioning and } (n - \mid \mathrm{Set}(\rho) \mid)^k \geq n\}.$$

We define $\mathcal{P} := \cup_{k \in \mathbb{N}} \mathcal{P}_k$. The elements in $\mathcal{P}$ are ordered under inclusion. An element $\rho \in \mathcal{P}$ is called a *(forcing) condition.* We use letters $\mathcal{A}, \mathcal{B}, \mathcal{C}, ...$ to denote subsets of $\mathcal{P}$. ♣

Notice that $\mathcal{P}_1 \subseteq \mathcal{P}_2 \subseteq .... \subseteq \mathcal{P}_r \subseteq .... \subseteq \mathcal{P}$, for each $r \in \omega$. The idea is to use $(\mathcal{P}, \subseteq)$ as the set of forcing conditions. As in [23]:

Definition 2.2.3 We say that $\mathcal{D} \subseteq \mathcal{P}$ is *dense* iff $\forall g \in \mathcal{P} \exists h \in \mathcal{D} \ h \supseteq g$.

We say that $\mathcal{D}$ is *quasi-definable* iff there exists a formula $\theta(x) \in L_{\mathbb{M}}(R_\omega)$ such that $\mathcal{D} := \{m \in \mathbb{M} : \ \mathbb{M} \models \theta(m)\}$ (the relation R_ω is defined by $R_\omega(a) \leftrightarrow a \in \omega$). ♣

Example 2.2.4 *$\mathcal{P}$ is dense and quasi-definable. $\mathcal{P}$ is not $L_{\mathbb{M}}$-definable.*

Definition 2.2.5 We say that $\rho_G \subseteq \mathcal{P}$ is a *generic filter* iff

(i) $\forall \alpha \in \rho_G \forall \beta \in \mathcal{P}\ \beta \subseteq \alpha \rightarrow \beta \in \rho_G$.

(ii) $\forall \alpha, \beta \in \rho_G \exists \gamma \in \rho_G\ \gamma \supseteq \alpha \wedge \gamma \supseteq \beta$.

(iii) For $\mathcal{D} \subseteq \mathcal{P}$ dense and quasi-definable $\rho_G \cap \mathcal{D} \neq \emptyset$.

We use the abbreviation $\tilde{\rho}_G := \cup_{\alpha \in \rho_G} \alpha$. ♣

2.3 Generic objects

Lemma 2.3.1 *If $\rho_G \subseteq \mathcal{P}$ is a generic filter, then $\tilde{\rho}_G$ defines a partition of $\{1, 2, ..., n\}$ into disjoint p-subsets.*

Proof: The only problem is to show $\mathrm{Set}(\rho_G) = I$. For an arbitrary $u \in I$ let $\mathcal{D}_u := \{\alpha \in \mathcal{P} :\ u \in \mathrm{set}(\alpha)\}$. It is straightforward to show that $\mathcal{D}_u$ is dense and quasi-definable so $\mathcal{D}_u \cap \rho_G \neq \emptyset$. Thus for each $u \in I$ there exists $\alpha_u \in \mathcal{D}_u \cap \rho_G$, and thus each $u \in \mathrm{Set}(\tilde{\rho}_G)$. □

Lemma 2.3.2 *For each $\rho_0 \in \mathcal{P}$ there exists a generic filter $\rho_G \subseteq \mathcal{P}$ such that $\rho_0 \in \rho_G$.*

Proof: Recall that both $\mathbb{M}$ and L are assumed to be countable, so there are only countably many quasi-definable dense sets. Let these be $\mathcal{D}_1, \mathcal{D}_2, ...$ According to the definition of denseness there exists a sequence of conditions $\rho_1 \subseteq \rho_2 \subseteq .... \in \mathcal{P}$ with $\rho_j \in \mathcal{D}_j,\ j = 1, 2, ...$ and $\rho_1 \supseteq \rho_0$. Clearly $\rho_0 \in \rho_G := \{\rho :\ \rho \subseteq \rho_k \text{ for some } k \in \omega\}$ is a generic filter. □

Definition 2.3.3 For a sentence $\psi \in L_{\mathbb{M}}(P)$ we define the *forcing relation* $|\!\vdash$ by letting

$\rho\ |\!\vdash \psi$ iff $(\mathbb{M}, \tilde{\rho}_G) \models \psi$ for all generic filters $\rho_G \ni \rho$. ♣

Lemma 2.3.4 *If $(\mathbb{M}, \tilde{\rho}_G) \models \psi$ for a generic filter ρ_G, then there exists $\rho_0 \in \mathcal{P}$ such that $\rho_0\ |\!\vdash \psi$.*

Proof: By use of induction on the logical complexity of a general formula $\psi(\vec{x})$, it is not hard to show that $\{(\vec{a}, \rho) \in \mathbb{M}^r \times \mathcal{P} :\ \rho\ |\!\vdash \psi(c_{\vec{a}})\}$ is quasi-definable. Continuing this argument for each $L_{\mathbb{M}}(P)$-sentence ψ, $\mathcal{D} := \{\rho \in \mathcal{P} :\ \rho\ |\!\vdash \psi \vee \rho\ |\!\vdash \neg\psi\}$ is both quasi-definable and dense. For the required ρ_0 take any $\rho_G \cap \mathcal{D}$. □

Definition 2.3.5 For $\theta, \psi \in B_{<\omega}(\mathrm{VAR}_{I,p})$ and $\rho \in \mathcal{P}$, $\theta \equiv_\rho \psi$ if $\theta^{\tilde{\rho}_G} = \psi^{\tilde{\rho}_G}$ for each generic filter $\rho_G \ni \rho$.

For $\theta \in B_{<\omega}(\mathrm{VAR}_{I,p})$ and $\rho \in \mathcal{P}$ we say that ρ *forces* $\theta^P = 1$ ($\theta^P = 0$) if for all generic $\rho_G \ni \rho$, $\theta^{\tilde{\rho}_G} = 1$ ($\theta^{\tilde{\rho}_G} = 0$). This is written $\rho\ |\!\vdash \theta^P = 1$ ($\rho\ |\!\vdash \theta^P = 0$). ♣

The next lemma shows how each appearance of $\neg$ can be eliminated.

Lemma 2.3.6 *Suppose that* $i_0 \in A, A \subseteq I, \mid A \mid = p$.

Suppose that $\pi_1 := \neg p_A$ *and* $\pi_2 := \vee_B \, p_B$, *where* B *runs through all* $B \subseteq I$ *with* $\mid B \mid = p, A \neq B$ *and* $i_0 \in B$. *Then* $\pi_1 \equiv_\rho \pi_2$ for all $\rho \in \mathcal{P}$.

Proof: Direct verification. □

Lemma 2.3.7 *For any Boolean circuit* $\theta \in B_d(\mathrm{VAR}_{I,p})$, *there exists a negation-free circuit* $\tilde{\theta} \in B_d(\mathrm{VAR}_{I,p})$ *such that* $\theta \equiv_\rho \tilde{\theta}$ *for any* $\rho \in \mathcal{P}$. *Furthermore,* $s(\tilde{\theta}) \leq s(\theta) \cdot \binom{n-1}{p-1}$.

Proof: First notice that $\neg \vee_i \pi_i \equiv_\emptyset \wedge_i \neg \pi_i$, and that $\neg \wedge_i \pi_i \equiv_\emptyset \vee_i \neg \pi_i$. So without loss of generality we can assume that negations appear only in front of the input variables. For each input variable p_A pick $i_0 \in A$ and replace each appearance of $\neg p_A$ with $\vee_{B:\ i_0 \in B \wedge B \neq A} \, p_B$. According to lemma 2.3.6 $\theta \equiv_\emptyset \tilde{\theta}$. This new circuit $\tilde{\theta}$, still has depth d. Furthermore, $s(\tilde{\theta}) \leq s(\theta) \cdot \max_{i_0}(s(\vee_{B:\ i_0 \in B, B \neq A} \, p_B)) = s(\theta) \cdot \binom{n-1}{p-1}$. □

Lemma 2.3.8 *For each bounded* $\psi \in L_{\mathbb{M}}(P)$, $\rho \mid\!\vdash \psi$ iff $\rho \mid\!\vdash (\epsilon_\psi)^P = 1$.

Proof: Induction on the number of logical constants in ψ. □

Definition 2.3.9 Two conditions α and β are *incompatible* ($\alpha \perp \beta$) if
$\exists A \in \alpha \exists B \in \beta \; A \neq B \wedge A \cap B \neq \emptyset$.

Two conditions α and β are *compatible* ($\alpha \parallel \beta$) if
$\forall A \in \alpha \forall B \in \beta \; A \neq B \rightarrow A \cap B = \emptyset$. ♣

Definition 2.3.10 $\mathcal{B} \subseteq \mathcal{P}$ is a *basis* for $\mathcal{P}$ iff
(a) $\forall \alpha, \beta \in \mathcal{B} \quad \alpha \neq \beta \rightarrow \alpha \perp \beta$.
(b) $\forall \rho \in \mathcal{P} \exists \alpha \in \mathcal{B} \; \rho \parallel \alpha$. ♣

Definition 2.3.11 $\parallel \mathcal{B} \parallel := \max_{\beta \in \mathcal{B}}(\mid \mathrm{Set}(\beta) \mid)$. ♣

Lemma 2.3.12 *Suppose that* $\parallel \mathcal{B} \parallel^k < n$ *for all* $k \in \omega$ *(or in short-hand notation* $\parallel \mathcal{B} \parallel < n^{\frac{1}{\omega}}$*). If* $\rho \in \mathcal{P}$ *and* $\rho \parallel \beta$, *then* $\rho \cup \beta \in \mathcal{P}$.

Proof: Assume that $\rho \in \mathcal{P}$. Thus there are $k_0 \in \omega$ such that $(n - \mid \mathrm{Set}(\rho) \mid)^{k_0} \geq n$. Also assume that $\beta \in \mathcal{B}$, where $\parallel \mathcal{B} \parallel \leq n^{\frac{1}{\omega}}$. Clearly $\mid \mathrm{Set}(\beta) \mid^{2k_0} \leq n$. Suppose $\rho \parallel \beta$. We have to show $\rho \cup \beta \in \mathcal{P}$. To show this, it suffices to show that

$$(n - \mid \mathrm{set}(\rho \cup \beta) \mid)^{2k_0} \geq (n - \mid \mathrm{set}(\rho) \mid - \mid \mathrm{Set}(\beta) \mid)^{2k_0}$$

$$\geq (n^{\frac{1}{k_0}} - n^{\frac{1}{2k_0}})^{2k_0} \geq n.$$

To do this notice that $2^k n^{\frac{1}{2}} \leq n$ for any $k \in \omega$. □

The next lemma shows an important technical point in Ajtai's choice of $\mathcal{P}$. It allows us to assume that $\emptyset \mid\vdash \psi$ in cases where $\rho_0 \mid\vdash \psi$ for some $\rho_0 \in \mathcal{P}$. To see this replace $I := \{1, 2, ..., n\}$ by $I' := \{1, 2, ..., n'\}$ where $n' := n - \mid \mathrm{Set}(\rho_0) \mid$. The lemma shows that if $\mathcal{P}$' is defined as $\mathcal{P}$ but with the underlying set I replaced by I', then $\mathcal{P}$' can be identified with the set of conditions in $\mathcal{P}$ which extends ρ_0.

Lemma 2.3.13 *Fix $\rho \in \mathcal{P}$. Define*

$\mathcal{P}^\rho := \{\tilde{\rho} : \tilde{\rho}$ *is a partial p-partition of* $I \setminus \mathrm{set}(\rho)$ *and* $\tilde{\rho} \cup \rho \in \mathcal{P}\}$,

$\mathcal{P}_k(J) := \{\tilde{\rho} : \tilde{\rho}$ *is a partial p-partition of* J *and* $(n' - \mid \mathrm{Set}(\tilde{\rho}) \mid)^k \geq n'\}$ *where* $J \subseteq I$ *and* $n' := \mid J \mid$.

Let $\mathcal{P}(J) := \cup_{k \in \omega} \mathcal{P}_k(J)$.

If $J = \mathrm{Set}(\rho)$ *for* $\rho \in \mathcal{P}$, *then* $\mathcal{P}^\rho = \mathcal{P}(\mathrm{Set}(\rho))$.

Proof: First we show the inclusion $\mathcal{P}^\rho \subseteq \mathcal{P}(\mathrm{Set}(\rho))$. Suppose that $\tilde{\rho} \in \mathcal{P}^\rho$. By definition for some $k_0 \in \omega$, such that $n' \leq n \leq (n - \mid \mathrm{Set}(\tilde{\rho} \cup \rho) \mid)^{k_0}$
$= (n - \mid \mathrm{Set}(\tilde{\rho}) \mid - \mid \mathrm{Set}(\rho) \mid)^{k_0} = (n' - \mid \mathrm{Set}(\tilde{\rho}) \mid)^{k_0}$, so $\tilde{\rho} \in \mathcal{P}(\mathrm{Set}(\rho))$.

Second, we show that the inclusion $\mathcal{P}(\mathrm{Set}(\rho)) \subseteq \mathcal{P}^\rho$. According to the assumption that $\rho \in \mathcal{P}$ there exists $l_0 \in \omega$ such that $(n - \mid \mathrm{Set}(\rho) \mid)^{l_0} \geq n$. According to the assumption that $\tilde{\rho} \in \mathcal{P}^\rho$, there exists $l_1 \in \omega$ such that $(n - \mid \mathrm{Set}(\rho) \mid - \mid \mathrm{Set}(\tilde{\rho}) \mid)^{l_1} \geq n - \mid \mathrm{Set}(\rho) \mid$. Now $(n - \mid \mathrm{Set}(\rho \cup \tilde{\rho}) \mid)^{l_0 l_1} = (n - \mid \mathrm{Set}(\rho) \mid - \mid \mathrm{Set}(\tilde{\rho}) \mid)^{l_0 l_1}$
$\geq (n - \mid \mathrm{Set}(\rho) \mid)^{l_0} \geq n$, so $\rho \cup \tilde{\rho} \in \mathcal{P}$. □

Lemma 2.3.14 *Suppose that $\mathcal{B}$ is a basis for $\mathcal{P}$ and $\mathcal{H} \subseteq \mathcal{B}$. Suppose also that $\mid\mid \mathcal{B} \mid\mid < n^{\frac{1}{\omega}}$. Then*

(a) $\rho \mid\vdash (\vee_{h \in \mathcal{H}} h)^P = 1$ *iff ρ is incompatible with all conditions $h' \in \mathcal{B} \setminus \mathcal{H}$.*

(b) $\rho \mid\vdash (\neg \vee_{h \in \mathcal{H}} h)^P = 1$ *iff ρ is incompatible with all conditions $h' \in \mathcal{H}$.*

Proof: $(a) \Rightarrow$: Suppose that $\rho \mid\vdash (\vee_{h \in \mathcal{H}} h)^P = 1$, but ρ is compatible with $h' \in \mathcal{B} \setminus \mathcal{H}$. By use of lemma 2.3.12 $\rho' := \rho \cup h' \in \mathcal{P}$. By using property (a) of a basis (definition 2.3.10) h' is incompatible with all conditions in $\mathcal{H}$. Clearly $\rho' \supseteq h$ so ρ' is incompatible with all conditions in $\mathcal{H}$. But then $(\vee_{h \in \mathcal{H}} h)^{\tilde{\rho}_G} = 0$ for each generic filter $\rho_G \ni \rho'$ (which exists by lemma 2.3.2). This contradicts $\rho \mid\vdash (\vee_{h \in H} h)^P = 1$.

$(a) \Leftarrow$: Assume that ρ is incompatible with all $h' \in \mathcal{B} \setminus \mathcal{H}$, and let
$\mathcal{D} := \{\rho' \in \mathcal{P} : (\rho'$ is compatible with some $h' \in \mathcal{H})$ or $(\rho'$ is incompatible with $\rho)\}$. By definition 2.3.10, $\mathcal{D} \subseteq \mathcal{P}$ is dense. Also $\mathcal{D}$ is quasi-definable. So according to

lemma 2.3.2, there exists a generic filter $\rho_G \ni \rho$. By definition 2.2.5 (iii) there exists $\alpha \in \mathcal{D} \cap \rho_G$, so there exists $h \in \mathcal{H}$ with $h \subseteq \alpha \subseteq \tilde{\rho}_G$.

$(b) \Rightarrow$ / $(b) \Leftarrow$ are proved as $(a) \Rightarrow$ / $(a) \Leftarrow$. □

Lemma 2.3.15 *Let $\epsilon_1, \epsilon_2, ..., \epsilon_u \quad u \in \mathbb{M}$, be an $\mathbb{M}$-definable sequence of Boolean circuits, each of the form $\epsilon_j :\equiv \vee_{h \in \mathcal{H}_j} h$. Let $\mathcal{B}_1, ..., \mathcal{B}_u$ be an $\mathbb{M}$-definable sequence and suppose that $t < n^{\frac{1}{\omega}}$ such that:*

(a) *for each $j = 1, 2, ...u \quad \mathcal{B}_j \subseteq \mathcal{P}$, is a basis for $\mathcal{P}$,*

(b) *for each $j = 1, 2, .., u \quad || \mathcal{B}_j || < t$,*

(c) *for each $j = 1, 2, ..., u, \ \mathcal{H}_j \subseteq \mathcal{B}_j$.*

Then for every generic filter ρ_G either

(a) *for all $j \in \{1, 2, .., u\}, \ \epsilon_j^{\tilde{\rho}_G} = 0$, or*

(b) *there exists $j_0 \leq u$ such that $\epsilon_{j_0}^{\tilde{\rho}_G} = 1$ and $\epsilon_j^{\tilde{\rho}_G} = 0$ for each $j < j_0$.*

Proof: Let

$$\mathcal{D} := \{\rho \in \mathcal{P} : \ (\exists j_0 \exists \beta \in \mathcal{H}_{j_0} \ \beta \mid\mid \rho \wedge \forall \gamma \in \cup_{j<j_0} \mathcal{H}_j \ \rho \perp \gamma) \text{ or } (\forall \gamma \in \cup_{j \leq u} \mathcal{H}_j \ \rho \perp \gamma)\}.$$

Clearly $\mathcal{D}$ is quasi-definable. For each $\rho_0 \in \mathcal{P}$, if ρ_0 is compatible with some $\beta \in \cup_j \mathcal{H}_j$, then there must be a smallest j_0 such that ρ_0 is compatible with some $\beta \in \mathcal{H}_{j_0}$. Here we uses that the least number principle is valid in $\mathbb{M}$. Now $\rho := h \cup \rho_0 \in \mathcal{P}$ (by lemma 2.3.12), and thus $\rho \in \mathcal{D}$. So $\mathcal{D}$ is dense. By definition 2.2.5 (iii) there exists $\rho \in \rho_G \cap \mathcal{D}$. This condition ρ is incompatible with all $h \in \mathcal{H}_j, \ j < j_0$. As $\tilde{\rho}_G \supseteq \rho \supseteq h \in \mathcal{H}_{j_0}$ clearly $(\vee_{h \in \mathcal{H}_{j_0}} h)^{\tilde{\rho}_G} = 1$. □

2.4 The key lemma

Recall that $\mathbb{M}$ is a countable non-standard model of $\mathrm{Th}(\mathbb{N})$ over a countable first order language L. As above we have fixed $p \in \omega \setminus \{1\}$, and $I := \{1, 2, ..., n\} \subseteq \mathbb{M}, \ n \in \mathbb{M} \setminus \omega$. As above the set $\mathcal{P}$ of forcing conditions consists of partial p-partitions ρ of I with $|\ \mathrm{Set}(\rho)\ | \leq n - n^{\frac{1}{\omega}}$ for some $k \in \omega$.

Lemma 2.4.1 (key lemma) *Let $\theta_1, \theta_2, ..., \theta_u$ be an $\mathbb{M}$-definable sequence of depth $\leq d \in \omega$ circuits with $\Sigma_{j=1}^u \ s(\theta_j) \leq n^t$ for some $t < n^{\frac{1}{\omega}}$ (i.e. $t^k < n$ for all $k \in \omega$).*

Let $\rho_0 \in \mathcal{P}$. There exists $\rho \supseteq \rho_0, \ \rho \in \mathcal{P}$ and an $\mathbb{M}$-definable sequence $\epsilon_1, \epsilon_2, ..., \epsilon_u$ of circuits together with an $\mathbb{M}$-definable sequence $\mathcal{B}_1, \mathcal{B}_2, .., \mathcal{B}_u$ such that

(a) *for $j = 1, 2, .., u$ each $\mathcal{B}_j$, is a basis for $\mathcal{P}$,*

(b) *for $j = 1, 2, ...u$ each ϵ_j is of the form $\vee_{h \in \mathcal{H}_j} h$ for some $\mathcal{H}_j \subseteq \mathcal{B}_j$,*

(c) *for each $j = 1, 2, ..., u, \quad \theta_j \equiv_\rho \epsilon_j$,*

(d) *for some $s \leq t \cdot \log(t)$ (actually for some $s \leq \omega \cdot t$), $\quad || \beta || \leq s$.*

If we combine the key lemma with lemma 2.3.15 we get:

Corollary 2.4.2 *If $\theta_1, \theta_2, ...\theta_u$ is an $\mathbb{M}$-definable sequence of depth $d \in \omega$ circuits with $\Sigma_{j=1}^{n}\ s(\theta_j) \leq n^t$ for some $t < n^{\frac{1}{\omega}}$, then for any generic filter $\rho_G \subseteq \mathcal{P}$ either*

(a) *for all $j \leq u$ $\theta_j^{\tilde{\rho}_G} = 1$, or*

(b) *there exists $j_0 \leq u$, such that $\theta_{j_0}^{\tilde{\rho}_G} = 1$ and $\tilde{\theta}_j^{\tilde{\rho}_G} = 0$ for all $j < j_0$.*

Before we prove the key lemma, we need to do some preparatory work.

2.5 Random conditions

My aim is to add a suitable probability distribution μ on the space $\mathcal{P}$ of forcing conditions.

Lemma 2.5.1 *For $k \geq 2p+1$, $k \in \mathbb{N}$, and fix $m < n$ such that $(n-m)^{k+1} > n \geq (n-m)^k$. Let μ_{sym} be the symmetrical probability distribution (perceived from inside $\mathbb{M}$) on the set $\{\rho \in \mathcal{P} :|\ \text{Set}(\rho)\ |= m\}$. For each $h \in \mathcal{P}$ with $|\ h\ |< n^{\frac{1}{\omega}}$, $\Pr(h \subseteq \rho) > n^{\frac{1}{2}}\ \Pr(h\ ||\ \rho \wedge \neg(h \subseteq \rho))$.*

Proof: Notice that for fixed $J \subseteq I$ with $|\ J\ |= m$ the number $\eta(m,p)$ of partial p-partitions ρ with $\text{Set}(\rho) = J$ is

$$\eta(m,p) = \frac{m!}{(p!)^{\frac{m}{p}}\ (\frac{m}{p})!}$$

when m is divisible by p and 0 otherwise. The set $\{\rho \in \mathcal{P} :|\ \text{Set}(\rho)\ |= m\}$ contains $\binom{n}{m} \cdot \eta(m,p)$ elements. If $h \in \mathcal{P}$, $|\ \text{Set}(h)\ |= up$ and $J \subseteq I \setminus \text{Set}(h)$ with $|\ J\ |= b$, then

$$\Pr(h \subseteq \rho \wedge J \cap \text{Set}(\rho) = \emptyset) = \frac{\binom{n-up-b}{m-up}\ \eta(m-up,p)}{\binom{n}{m}\ \eta(m,p)} = \frac{(n-up-b)!(n-m)!(p!)^u(\frac{m}{p})!}{n!(n-m-b)!(\frac{m}{p}-u)!}$$

Now suppose $n - n^{\frac{1}{k}} \leq m < n - n^{\frac{1}{k+1}}$, and $b, u < n^{\frac{1}{\omega}}$. There exists a suitable real (in the sense of $\mathbb{M}$) $c \in [0,1]$ such that $\Pr(h \subseteq \rho \wedge J \cap \text{Set}(\rho) = \emptyset)$
$= (\frac{1}{n})^{u(p-1)+b(1-\frac{1}{k+c})}$. Here we use the fact that a sufficiently strong part of real analysis can be developed inside $\mathbb{M}$. Now

$$\Pr(h\ ||\ \rho \wedge \neg(h \subseteq \rho)) = \Sigma_{j=0}^{u-1}\Sigma_{h' \subseteq h, |h'|=j}\ \Pr(h' \subseteq \rho \wedge (\text{Set}(h) \setminus \text{Set}(h') \cap \text{Set}(\rho) = \emptyset))$$

$$= \Sigma_{j=0}^{u-1}\Sigma_{h' \subseteq h, |h'|=j}\ (\frac{1}{n})^{(\frac{p}{k+c}-1)j+pu(1-\frac{1}{k+c})}$$

$$= \Sigma_{j=0}^{u-1} \binom{u}{j} (\frac{1}{n})^{(\frac{p}{k+c}-1)j+pu(1-\frac{1}{k+c})}$$

In general $\Sigma_{j=0}^{u-1}\binom{u}{j}a^j = (a+1)^u - a^u$. Let $a := n^{1-\frac{p}{k+c}}$, and notice that $\Sigma_{j=0}^{u-1}\binom{u}{j}a^j \leq 2ua^{u-1}$. Thus

$$\Sigma_{j=0}^{u-1}\binom{u}{j}(\frac{1}{n})^{(\frac{p}{k+c}-1)j+pu(1-\frac{1}{k+c})} = 2u(\frac{1}{n})^{pu(1-\frac{1}{k+c})} \cdot (\frac{1}{n})^{(\frac{p}{k+c}-1)(u-1)}$$

$$= 2u \cdot (\frac{1}{n})^{(p-1))u+(1-\frac{p}{k+c})}$$

$$= \Pr(h \subseteq \rho) \cdot 2u(\frac{1}{n})^{(1-\frac{p}{k+c})} \leq \Pr(h \subseteq \rho) \cdot (\frac{1}{n})^{\frac{1}{2}}$$

when $k \geq 2p+1$. In all estimates c is chosen as a suitable real constant in $[0,1]$. □

Lemma 2.5.2 *Fix $k \in \mathbb{N}$. Also fix $t < n^{\frac{1}{\omega}}$. Then there exists a (global) probability distribution $\mu_{\rm glo}$ on the $\mathbb{M}$-definable set consisting of all partial p-partitions, such that for each $h \in \mathcal{P}$ with $\mid h \mid \leq t$*

(i) *If $C(\rho)$ is a monotone property (i.e. $C(\rho) \wedge \rho \subseteq \rho' \rightarrow C(\rho')$), then*

$$\Pr(C(\rho) \mid \ h \mid\mid \rho) \geq \frac{1}{8} \cdot \Pr(C(\rho) \mid \ h \mid\mid \rho \wedge \neg(h \subseteq \rho)),$$

(ii) *there is $s \in \mathbb{M} \setminus \omega$ such that $\Pr(\rho \notin \mathcal{P}_{4k} \vee (\rho \notin \mathcal{P} \setminus \mathcal{P}_k)) \geq 1 - \exp(n^s)$.*

Proof: Notice that in general $\Pr(C \mid B_1 \vee B_2 \vee ... \vee B_k) \leq \max_j \Pr(C \mid B_j)$, so if $h := \cup_{i \in F} \{A_i\}$, it suffices to construct a suitable $\mu_{\rm glo}$ which besides (ii) has $\Pr(C \mid \wedge_{i \in F} A_i \in \rho) \geq \Pr(C \mid (\wedge_{i \in G} A_i \in \rho) \wedge (\wedge_{j \in F \setminus G} \mathrm{Set}(\rho) \cap A_j = \emptyset))$ for any $G \subseteq F$. Let $A :\equiv \wedge_{i \in F}(A_i \in \rho)$, $B :\equiv \wedge_{i \in G}(A_i \in \rho) \wedge (\wedge_{i \in F \setminus G}(A_i \cap \mathrm{Set}(\rho) = \emptyset))$, $C \equiv C(\rho)$, and for $l = 0, 1, 2, ...$ let $D_l :\equiv \mid \rho \mid = l$. Let $P_l :\equiv \Pr(D_l)$, and let $g :=\mid G \mid$. We choose $\mu_{\rm glo}$ symmetric on each each set $\{\rho : D_l\}$. We define $\mu_{\rm glo}$ by choosing suitable numbers $p_0, p_1, ...p_u$ with $\Sigma_j p_j = 1$. Notice that any $\mu_{\rm glo}$ defined this way, for any $l = g, g+1, ..., u$ has $\Pr(C \mid A \wedge D_l) \geq \Pr(C \mid B \wedge D_{l-g})$. To see this notice that any monotone property C can be written as a disjunction $\vee_\delta (\delta \subseteq \rho)$. Thus for $l = g, g+1, ..., u$ we have

$$(*) \qquad \frac{\Pr(C \wedge A \wedge D_l)}{\Pr(A \wedge D_l)} \geq \frac{\Pr(C \wedge B \wedge D_{l-g})}{\Pr(B \wedge D_{l-g})}$$

We have to show that $\frac{\Pr(C \wedge A)}{\Pr(A)} \geq \frac{1}{8} \cdot \frac{\Pr(C \wedge B)}{\Pr(B)}$. Now

$$\frac{\Pr(C \wedge A)}{\Pr(A)} = \frac{\Sigma_l \Pr(C \wedge A \wedge D_l) \cdot p_l}{\Sigma_l \Pr(A \wedge D_l) \cdot p_l} \geq (1 + \Sigma_{l=0}^{g-1})^{-1} \cdot \frac{\Sigma_{l=g}^{u} \Pr(C \wedge A \wedge D_l) \cdot p_l}{\Sigma_{l=g}^{u} \Pr(A \wedge D_l) \cdot p_l}$$

The inequality holds because A is a monotone property and $\Sigma_{l=0}^{g-1}\ \Pr(A \wedge D_l)p_l \leq \Sigma_{l=0}^{g-l}\Pr(A \wedge D_g)p_l \leq \Sigma_{l=g}^{2g}\ \Pr(A \cap D_l)p_l \leq \Sigma_{l=g}^{u}\Pr(A \wedge D_l)p_l$. This holds because $2g < n^{\frac{1}{\omega}} < l_0$ as long as $p_{l+g} \geq p_l$ for $l \leq g$. But by $(*)$

$$\frac{\Sigma_{l=g}^{u}\ \Pr(C \wedge A \wedge D_l)\cdot p_l}{\Sigma_{l=g}^{u}\ \Pr(A \wedge D_l)\cdot p_l} \geq \frac{\Sigma_{0}^{u-g}\ \Pr(C \wedge B \wedge D_l)\cdot p_{l+g}}{\Sigma_{l=0}^{u-g}\ \Pr(B \wedge D_l)\cdot p_{l+g}}$$

$$\geq \frac{\Sigma_l\ \Pr(C \wedge B \wedge D_l)\cdot p_l \cdot \min_l(\frac{p_{l+g}}{p_l})}{\Sigma_l\ \Pr(B \wedge D_l)\cdot p_l \cdot \max_l(\frac{p_{l+g}}{p_l})} = \min_l(\frac{p_{l+g}}{p_l})\cdot\min_l(\frac{p_l}{p_{l+g}})\cdot\frac{\Pr(C \wedge B)}{\Pr(B)}.$$

Now we are ready to define $p_0, p_1, ...$ Recall that $g \leq |\ h\ | \leq t$.

For $l < l_0$ let $p_{l+1} := 2^{\frac{1}{t}}\cdot p_l$, and for $l_0 \leq l$ let $p_{l+1} := 2^{-\frac{1}{t}}p_l$. Now $\min_l(\frac{p_{l+t}}{p_l}) = \min_l(\frac{p_l}{p_{l+t}}) = \frac{1}{2}$ and $(1+\Sigma_{l=0}^{t-1}p_j)^{-1} \leq \frac{1}{2}$ so

$$\frac{\Pr(C \wedge A)}{\Pr(A)} \geq \frac{1}{8}\cdot\frac{\Pr(C \wedge B)}{\Pr(B)}.$$

Thus (i) holds. Furthermore notice that if $(n - pl_0)^k \geq n > (n - pl_0)^{k-1}$ then the probabilities are sufficiently concentrated around l_0 to ensure that (ii) is satisfied. □

The factor $\frac{1}{8}$ can be replaced by any standard rational $q < 1$. Also notice that there are many other choices of the distribution $p_0, ..., p_u$. One can for instance choose the binomial distribution with mean l_0. The point is that $\min_l(\frac{p_{l+t}}{p_l})\cdot\min_l(\frac{p_l}{p_{l+t}})$ is not too small, while at the same time the probability distribution tails off sufficiently fast.

Notice that a phenomenon reminiscent of the complementary principle, is involved. If μ_{glo} is focussed on some $\mathcal{P}_k$ then (i) cannot hold. On the other hand if μ_{glo} is unfocussed and global (ii) cannot hold. As an example of the first claim consider the property $C(\rho) :\equiv \exists\delta\ (|\ \delta\ | = l_0 - |\ h\ | + 1 \wedge \text{Set}(\delta) \cap \text{Set}(h) = \emptyset \wedge \delta \subseteq \rho$. If $p_{l_0} = 1$ then $\Pr(C \mid\ h \subseteq \rho) = 0$ while $\Pr(C \mid\ h\ ||\ \rho \wedge \neg(h \subseteq \rho)) = 1$. By lemma 2.5.1 this is a violation of condition (i).

Corollary 2.5.3 *For $k \geq 2p+1$, $k \in \mathbb{N}$ and $t < n^{\frac{1}{\omega}}$ there exists a $\mathbb{M}$-definable probability distribution μ on P_{4k}, such that*

(1) *for each $h \in \mathcal{P}$ with $|\ h\ | < t$, $Pr(h \subseteq \rho) > n^{\frac{1}{2}}\cdot\Pr(h\ ||\ \rho \wedge \neg(h \subseteq \rho))$,*

(2) *for each monotone property $C(\rho)$*
$\Pr(C(\rho) \vee \rho \notin \mathcal{P}_{2k} \mid\ h\ ||\ \rho) \geq \frac{1}{8}\cdot\Pr(C(\rho) \vee \rho \notin \mathcal{P}_{2k} \mid\ h\ ||\ \rho \wedge \neg(h \subseteq \rho))$.

Proof: Let μ be the (normalised) probability distribution obtained by restricting μ_{glo} to P_{4k}. □

We need the following elementary fact:

Lemma 2.5.4 *For each number w,*

$$\Pr(A \mid\ B \wedge C) \geq w \cdot \Pr(A \mid\ B) \text{ iff } \Pr(C \mid\ A \wedge B) \geq w \cdot \Pr(C \mid\ B).$$

Proof: Both sides holds iff $\Pr(A \wedge B \wedge C)\Pr(B) \geq w \cdot \Pr(A \wedge B)\Pr(B \wedge C)$. □

Corollary 2.5.5 *If μ is chosen on $\mathcal{P}_{4k}$ such that condition (1) and (2) in corollary 2.5.3 hold, then*

$$\Pr(h^j \subseteq \rho \vee \rho \notin \mathcal{P}_{2k} \mid\ h^j \mid\mid \rho \wedge (\rho \perp h^1) \wedge .... \wedge (\rho \perp h^{j-1})) \geq \frac{1}{8} \cdot \Pr(h^j \subseteq \rho \mid\ h^j \mid\mid \rho).$$

Proof: Let $A :\equiv (\rho^j \subseteq \rho), B :\equiv h^j \mid\mid \rho$ and let $C \equiv (\rho \perp h^1) \wedge .... \wedge (\rho \perp h^{j-1}) \vee \neg\rho \in \mathcal{P}_{2k}$. The lemma now follows by use of condition (2) in corollary 2.5.3 and lemma 2.5.4. □

2.6 Collapse of circuits

In order to prove the key lemma we prove that:

Lemma 2.6.1 *Suppose that $\theta :\equiv \neg(\vee_{h \in \mathcal{H}}\ h)$ where $s(\theta) \leq n^t$ for some $t < n^{\frac{1}{\omega}}$. Let $s < n^{\frac{1}{\omega}}$. Suppose that μ is a probability distribution satisfying (1) and (2) in corollary 2.5.3 on $\mathcal{P}_{4k}$ for some $k \geq 2p+1$, $k \in \omega$. Then there exists an $\mathbb{M}$-definable set $\mathcal{C} \subseteq \mathcal{P}_k$ such that $\mathcal{C} \subseteq \{\rho \in \mathcal{P}_{4k} :\ \exists \tilde{\mathcal{H}} \subseteq \mathcal{P},\ \mid\mid \tilde{\mathcal{H}} \mid\mid \leq ps$ such that $\theta \equiv_\rho \vee_{h \in \tilde{\mathcal{H}}}\ h$ with $s(\vee_{h \in \tilde{\mathcal{H}}}\ h) \leq n^t\}$ and such that*

$$\mu(\mathcal{C}) \geq 1 - (\frac{1}{n})^{\frac{ks-2ps}{2k}}$$

Corollary 2.6.2 *Let $\epsilon_1, \epsilon_2, ..., \epsilon_u$ be an $\mathbb{M}$-definable sequence of depth $\leq d$ circuits with $\Sigma_j\ s(\epsilon_j) \leq n^t$ for some $t < n^{\frac{1}{\omega}}$. Let $\rho_0 \in \mathcal{P}$. There exists $\rho \supseteq \rho_0$, $\rho \in P$ and an $\mathbb{M}$-definable sequence $\epsilon'_1, ..., \epsilon'_u$ of depth $\leq d-1$ (when $d \geq 3$) circuits, with $\Sigma_j\ s(\epsilon'_j) \leq n^t$ such that $s(\epsilon'_j) \leq s(\epsilon_j)$ $j = 0, 1, 2, ...$ When $d = 2$ there exists an $\mathbb{M}$-definable sequence $\epsilon'_1, \epsilon'_2, ..., \epsilon'_u$ of depth ≤ 2 circuits of the form $\epsilon'_j :\equiv \vee_{h \in H_j}\ h$. Furthermore, each set $\mathcal{H}_j \subseteq \mathcal{P}$ contains conditions h which have all $\mid \mathrm{Set}(h) \mid \leq ps$ for some s with $s > \omega \cdot t$, and $s < n^{\frac{1}{\omega}}$.*

Proof: (lemma 2.6.1 $\Rightarrow$ corollary 2.6.2). By use of lemma 2.3.7 we can assume $\epsilon_1, \epsilon_2, ..., \epsilon_u$ are all negation-free. There is an $\mathbb{M}$-definable sequence $\tilde{\epsilon}_1, ..., \tilde{\epsilon}_u$ of depth $\leq d-1$ circuits, where all "input nodes" are depth ≤ 2 circuits. Let $\pi_1, \pi_2, ..., \pi_r$ be the $\mathbb{M}$-definable sequence of these. Clearly $r \leq n^t$. Without loss of generality, each π_j is either a "disjunction of conjunctions" or is a "conjunction of disjunctions". Notice $\Sigma_j\ s(\pi_j) < n^t$. By use of lemma 2.6.1 (which according to lemma 2.3.13 also holds

when the underlying set I is replaced by $I' := I \setminus \mathrm{Set}(\rho_0))$ for some fixed $k \geq 2p+1$, for each $j \leq r$ there exists an $\mathbb{M}$-definable sequence $\mathcal{C}_1, ..., \mathcal{C}_r \subseteq \mathcal{P}_k$ each with

$$\mu(\mathcal{C}_j) \geq 1 - (\frac{1}{n - \mid \mathrm{Set}(\rho) \mid})^{\frac{ks-2ps}{2k}}$$

such that for $j = 1, 2, .., r$ and all $\rho \in \mathcal{C}_j \quad \pi_j \equiv_\rho \vee_{h \in H'} h$ if π_j is a disjunction of conjunctions, and $\pi_j \equiv_\rho \neg(\vee_{h \in H'} h)$ if π_j is a conjunction of disjunctions.

Now $\mu(\mathcal{C}_1 \cap \mathcal{C}_2 \cap .... \cap \mathcal{C}_r) \geq 1 - r \cdot (\frac{1}{n})^{\frac{ks-2ps}{2k}} > 0$ (when $s \geq 2kt$). So there exists $\rho \in \mathcal{C}_1 \cap .... \cap \mathcal{C}_r$ with $\rho \supseteq \rho_0$. Replace each depth ≤ 2 "input" circuits with a suitable depth ≤ 2 circuit. □

Repeated use of this corollary (applied at most d times) reduces problem of proving the key lemma to that of proving lemma 2.6.1.

2.7 The switching lemma

Definition 2.7.1 For $i \in I$ let

$$\mathcal{E}_i := \{h' : h' \supseteq h \wedge \mathrm{Set}(h') \supseteq \mathrm{Set}(h) \cup \{i\} \wedge \mid h' \mid \leq \mid h \mid +1\}.$$

♣

Definition 2.7.2 We say that $\mathcal{H}_2 \subseteq \mathcal{P}$ is an *atomic tree-like refinement* of $\mathcal{H}_1 \subseteq \mathcal{P}$, $(\mathcal{H}_1 \rightarrow_{\mathrm{ATR}} \mathcal{H}_2)$ if $\exists h \in \mathcal{H}_1 \exists i \in I \; \mathcal{H}_2 = (\mathcal{H}_1 \cup \mathcal{E}_i(h)) \setminus \{h\}$.

We say that $\tilde{\mathcal{H}}$ is a *tree-like refinement* of $\mathcal{H}$, $(\mathcal{H} \rightarrow_{TR} \tilde{\mathcal{H}})$ if there exists an $\mathbb{M}$-definable sequence $(\mathcal{H}_0, \mathcal{H}_1, ..., \mathcal{H}_r)$ such that $\mathcal{H}_0 = \mathcal{H}$, $\mathcal{H}_r = \tilde{\mathcal{H}}$ and for each $j < r \quad \mathcal{H}_j \rightarrow_{ATR} \mathcal{H}_{j+1}$. ♣

Definition 2.7.3 $\mathcal{B} \subseteq \mathcal{P}$ is a *tree-like basis* if $\{\emptyset\} \rightarrow_{TR} \mathcal{B}$. ♣

Lemma 2.7.4 *Suppose that $\mathcal{B} \subseteq \mathcal{P}$ is $\mathbb{M}$-definable and that $\| \mathcal{B} \| < n^t$ for some fixed $t < \frac{1}{\omega}$. If $\mathcal{B}$ is a tree-like basis, then $\mathcal{B}$ is a basis for $\mathcal{P}$.*

Proof: First notice that each $\alpha, \beta \in \mathcal{B}$ $(\alpha \neq \beta)$ are incompatible. Suppose that $\rho_0 \in \mathcal{P}$ is incompatible with all $\beta \in \mathcal{B}$. Let $(\mathcal{B}_0, ..., \mathcal{B}_r)$ be an $\mathbb{M}$-definable sequence with $\mathcal{B}_0 := \{\emptyset\}$, $\mathcal{B}_r := \mathcal{B}$ and where $\mathcal{B}_j \rightarrow_{ATR} \mathcal{B}_{j+1}$ for $j = 0, 1, ..., r-1$. $\mathbb{M}$ satisfies the least number principle so there must be a smallest j_0 such that ρ_0 is incompatible with all $\beta \in \mathcal{B}_{j_0}$. There exists $\beta' \in \mathcal{B}_{j_0 - 1}$ compatible with ρ_0. As $\mid \mathrm{set}(\beta') \mid \leq n^{\frac{1}{\omega}}$ by lemma 2.3.12 $\rho = \rho_0 \cup \beta' \in \mathcal{P}$. Let $i_0 \in I$ such that $\mathcal{B}_{j_0} = (\mathcal{B}_{j_0 - 1} \cup \mathcal{E}_{i_0}(\beta'')) \setminus \{\beta''\}$.

Now as ρ (like ρ_0) is incompatible with all conditions in $\mathcal{B}_{j_0}$, $\beta' \notin \mathcal{B}_{j_0}$ and thus $\beta' = \beta''$. We get the required contradiction by noticing that $\rho \in \mathcal{P}$ must be compatible with some $\rho' \in \mathcal{E}_{i_0}(\beta')$ □

Definition 2.7.5 For $\mathcal{H} \subseteq \mathcal{P}$ and for $\rho \in \mathcal{P}$ let

$$\mathcal{H}^\rho := \{h' \in \mathcal{P} : \ h' = h \setminus \rho \text{ for some } h \in \mathcal{H} \text{ with } h \mid\mid \rho\}.$$

We say that $\mathcal{B}$ *refines* $\mathcal{H}$ if for each $\beta \in \mathcal{B}$ and each $h \in \mathcal{H}$ if $\beta \mid\mid h$, there is $h' \in \mathcal{H}$ such that $h' \subseteq \beta$. ♣

Lemma 2.7.6 *If $\mathcal{B}$ is a basis which refines $\mathcal{H}$, and $\mathcal{H}_\mathcal{B} := \{\beta \in \mathcal{B} : \ \exists h \in \mathcal{H} \ \beta \supseteq h\}$ then $\vee_{h\in\mathcal{H}} \ h \equiv_\rho \ \vee_{h'\in\mathcal{H}_\mathcal{B}} \ h'$ for all $\rho \in \mathcal{P}$.*

Proof: Straightforward. □

Lemma 2.7.7 *Let $\mathcal{H} \subseteq \mathcal{P}$ be a collection of conditions with $\forall h \in \mathcal{H}$, $\mid \mathrm{Set}(h) \mid \leq n^t$ for some fixed $t < n^{\frac{1}{\omega}}$. Let $\rho_0 \in \mathcal{P}_k$, $k \in \omega$, and let μ be a probability distribution on $\mathcal{P}_{4l}(I \setminus \mathrm{Set}(\rho_0))$ $(l \geq 2p+1)$ which satisfies the conditions in corollary 2.5.3.*

If $\rho \in \mathcal{P}_{4l}(I \setminus \mathrm{Set}(\rho_0))$ is chosen randomly according to the probability distribution μ, then for each $s < n^{\frac{1}{\omega}}$, with probability $\geq 1 - (\frac{1}{n-|\mathrm{Set}(\rho_0)|})^{\frac{ks-2ps}{2k}}$ there exists a tree-like basis $\mathcal{B}$ which refines $\mathcal{H}^\rho$ such that $\mid\mid \mathcal{B} \mid\mid \leq ps$.

This lemma immediately implies lemma 2.6.1. To see this let

$$\mathcal{C} := \{\rho : \ \exists \mathcal{B} \text{ a tree} - \text{like basis which refines } \mathcal{H} \text{ and } \ \mid\mid \mathcal{B} \mid\mid \leq ps\}.$$

Notice $\mathcal{C}$ is $\mathbb{M}$-definable. According to lemma 2.7.6, the θ in lemma 2.6.1, has $\theta \equiv_\rho \vee_{h\in\mathcal{H}_\mathcal{B}} \ h$ where $\mathcal{H}_\mathcal{B} := \{\beta \in \mathcal{B} : \exists h \in \mathcal{H} \ \beta \supseteq h\}$. Thus to show the key lemma it suffices to show lemma 2.7.7.

2.8 Some games involving forcing

As above assume $\mathbb{M}$ to be a countable non-standard model. Assume also that $p \in \mathbb{N} \setminus \{1\}$ and $I := \{1, 2, ..., n\} \subseteq \mathbb{M}$, with $n \in \mathbb{M} \setminus \omega$ be fixed. Let $\mathcal{P}_1 \subseteq \mathcal{P}_2 \subseteq .... \subseteq \mathcal{P}_r \subseteq .... \subseteq \mathcal{P}$, $r \in \omega$, be the stratification of $\mathcal{P}$ defined above. Our aim is to show lemma 2.7.7.

Definition 2.8.1 Suppose that $t \leq s$ where $s < n^{\frac{1}{\omega}}$ (e.g. t, s are small) and let $< h^1, h^2, ..., h^v > \in \mathbb{M}$, $v \in \mathbb{M}$ be a sequence of conditions with $\mid \mathrm{Set}(h^j) \mid \leq t$, $j \leq v$. Suppose also that $\{h^1, h^2, ..., h^v\}$ is complete for $\mathcal{P}$ (i.e. $\forall \rho \in \mathcal{P} \ \exists j \leq v \ \rho \mid\mid h^j$). The game $G(n, k, t, s, < h^1, h^2, ..., h^v >)$ is played by two players I and II as follows:

Round 0: Player I selects a condition $\rho \in \mathcal{P}_k$.

Round 1: Consider the first $i \leq v$ where $h_1 := h^i$ is compatible with ρ (which exists because the collection $\{h^1, .., h^v\}$ is assumed to be complete for $\mathcal{P}$).

If $\rho \supseteq h_1$ player I wins and the *output* of the game is $\emptyset$.

If $\mathrm{Set}(h_1) \setminus \mathrm{Set}(\rho) \neq \emptyset$ let $a_1 := \min(\mathrm{Set}(h_1) \setminus \mathrm{Set}(\rho)) \in I$. Player II selects an p-element set $A \subseteq I$ such that:

(1) $\{A\}$ is compatible with ρ $\quad (\{A\} \mid\mid \rho)$.

(2) $\{A\}$ is incompatible with h_1 $\quad (\{A\} \perp h_1)$.

(3) $a_1 \in A$.

Let $\delta_1 := \{A\}$ and proceed to the next round.

Round j+1: Consider the next condition $h_{j+1} := h^{i_{j+1}}$, $i_{j+1} > i_j$ compatible with $\rho \cup \delta_j$ (according to lemma 2.3.12 such exists because $\rho \cup \delta_j \in \mathcal{P}$ when $j \leq s$, and $\{h^1, ..., h^v\}$ is complete for $\mathcal{P}$).

If $\rho \cup \delta_j \supseteq h_{j+1}$ player I wins and the *output* of the game is δ_j.

If $\mathrm{Set}(h_{j+1}) \setminus \mathrm{Set}(\rho) \neq \emptyset$ let $a_{j+1} := \min(\mathrm{Set}(h_{j+1}) \setminus \mathrm{Set}(\rho)) \in I$. Player II selects a p-element set $A \subseteq I$ such that:

(1) $A \cap \mathrm{Set}(\rho) = \emptyset$.

(2) $\{A\}$ is incompatible with h_{j+1}.

(3) $a_{j+1} \in A$.

Let $\delta_{j+1} := \delta_j \cup \{A\}$, and proceed to the next round.

Round s+1: If this round is reached, player II wins and the game is terminated. ♣

Notice that player I does influence the game after the choice of ρ. The strategies of player I can thus be identified with the conditions in $\mathcal{P}_k$.

Definition 2.8.2 We call $\rho \in \mathcal{P}_k$ *a winning strategy* for player I, if player wins irrespectively of what player II chooses. ♣

Lemma 2.8.3 *Suppose that $\mathcal{H} = \mathcal{H}_0 \cup \mathcal{H}_1$ is complete for $\mathcal{P}$ (i.e. $\forall \rho \in \mathcal{P}\ \exists h \in \mathcal{H}\ \rho \mid\mid h$). Suppose that $\mathcal{H}_0 := \{h^1, ..., h^u\}$ and $\mathcal{H}_1 := \{h^{u+1}, ..., h^v\}$ $u \leq v \in \mathbb{M}$. Consider the game $G(n, k, t, s, < h^1, h^2, ..., h^v >)$, and suppose that $\rho \in \mathcal{P}_k$ is a winning strategy for player I. Let $\mathcal{B}$ be the set of possible outputs (when player II varies his/her possible plays). Then $\mathcal{B}$ is a tree-like basis relative to $I \setminus \mathrm{Set}(\rho)$. Furthermore, $\mathcal{B}$ refines $\mathcal{H}_0^\rho$ and has $\mid\mid \mathcal{B} \mid\mid \leq ps$.*

Proof: We are given a winning strategy ρ for player I. We have to show that $\mathcal{B}$ is a tree-like basis. We view each δ constructed at a certain stage in an actually played game, as a (uniquely defined) "situation". Let $S(\delta)$ denote the situations which can be reached from δ. We want to construct $\mathcal{B}$ as a sequence

$$\{\emptyset\} \rightarrow_{ATR} \mathcal{B}_1 \rightarrow_{ATR} .... \rightarrow_{ATR} \mathcal{B}_j \rightarrow_{ATR} \mathcal{B}_{j+1} \rightarrow .... \rightarrow \mathcal{B}.$$

Suppose that $\mathcal{B}_j$ has been constructed. Pick any situation δ' which has not been reached so far, but which can be reached from a situation corresponding to a $\delta \in \mathcal{B}_j$ which has already been considered. Let $\mathcal{B}_{j+1} := (\mathcal{B}_j \cup \mathcal{E}_a(\delta)) \setminus \{\delta\}$ where $a := \min(\mathrm{Set}(h) \setminus \mathrm{Set}(\rho))$. Here h denotes the next h^i compatible with ρ in the situation corresponding to δ. As ρ was assumed to be a winning strategy for player I, this procedure terminates, and all $\beta \in \mathcal{B}$ get $\mid \mathrm{Set}(\beta) \mid \leq ps$.

Finally we show that $\mathcal{B}$ refines $\mathcal{H}_0^\rho$. We have to show that if $h \mid\mid \beta$ for some $h \in \mathcal{H}_0^\rho$ and $\beta \in \mathcal{B}$, then there exists $h' \in \mathcal{H}_0^\rho$ such that $h' \subseteq \beta$. So suppose β is compatible with $(h^j)^\rho \in \{(h^1)^\rho, (h^2)^\rho, ..., (h^u)^\rho\}$. If $\beta \supseteq (h^j)^\rho$ we are done. If $\neg(\beta \supseteq (h^j)^\rho$ the game which produced β, must have terminated before h^j so there must be $j' < j$ such that $(h^{j'})^\rho \subseteq \beta$. As the sequence $h^1, h^2, ..., h^u, h^{u+1}, ..., h^v$ had all the elements from $\mathcal{H}_0$ listed in the beginning, $(h^{j'})^\rho \in \mathcal{H}_0^\rho$. □

The next theorem shows that "almost all" (in the sense of μ) strategies ρ are winning strategies for player I. More specifically:

Theorem 2.8.4 *Consider the game $G(n, k, t, s, < h^1, h^2, ..., h^v >)$. Let $\mathcal{W}_I \subseteq \mathcal{P}_k$ be the set of winning strategies for player I (we only consider $\mathbb{M}$-definable strategies). If μ is a probability distribution on $\mathcal{P}_{4k}$ which satisfies condition (1) and (2) in corollary 2.5.3, then*

$$\mu(\mathcal{W}_I) \geq 1 - (\frac{1}{n})^{\frac{ks-2ps}{2k}}.$$

Notice that t does not enter the estimate as long as $t < n^{\frac{1}{\omega}}$.

We show theorem 2.8.4 by comparing the game $G(n, k, t, s, < h^1, h^2, .., h^v >)$ with another game $G'(n, k, t, s)$.

Definition 2.8.5 The game $G'(n, k, t, s)$ is played by two players I and II as follows (all sets etc. are $\mathbb{M}$-definable).

Player II selects $J \subseteq I$, with $\mid J \mid \leq ps$, and selects a sequence $h^1, h^2, .., h^l$ of conditions each with $\mathrm{Set}(h^i) \subseteq I \setminus J$, and $\mid \mathrm{Set}(h^i) \mid \leq t$.

Player I then selects a condition $\rho \in \mathcal{P}_k$. Consider the first condition $h := h^j$ compatible with ρ (if there is no such player I wins). If $h \subseteq \rho$ player I wins, otherwise player II wins. ♣

In this game player II makes the choices before player I. Clearly player I always has a winning reply (just choose $\rho \supseteq h^1$). We claim almost all player I's replies are winning:

Theorem 2.8.6 *Let $\tilde{\mathcal{W}}_I(\tau)$ be the set ($\mathbb{M}$-definable) of replies ρ which ensure a win for player I after player II made a choice τ. Then if μ is a probability distribution which satisfies condition (1) and (2) in corollary 2.5.3,*

$$\mu(\tilde{\mathcal{W}}_I(\tau)) \geq 1 - (\frac{1}{n})^{\frac{1}{2}}.$$

Proof: First notice

$$\mu(\tilde{\mathcal{W}}_I(\tau)) \geq \min_j \Pr(h^j \subseteq \rho \mid \quad h^j \mid\mid \rho \wedge \rho \perp h^1 \wedge ... \wedge \rho \perp h^{j-1}).$$

According to lemma 2.5.4 and condition (2) in corollary 2.5.3 for any j

$$\Pr(h^j \subseteq \rho \mid \quad h^j \mid\mid \rho \wedge (\rho \perp h^1) \wedge .... \wedge (\rho \perp h^{j-1})) \geq \Pr(h^j \subseteq \rho \mid \ h^j \mid\mid \rho)$$

(let $A \equiv (h^j \subseteq \rho)$, $B \equiv h^j \mid\mid \rho$ and $C \equiv \rho \perp h^1 \wedge .... \wedge \rho \perp h^{j-1}$). But by condition (1) in corollary 2.5.3

$$\Pr(h^j \subseteq \rho \mid \ h^j \mid\mid \rho) = \frac{\Pr(h^j \subseteq \rho)}{\Pr(h^j \subseteq \rho) + \Pr(h^j \mid\mid \rho \ \wedge \neg(h^j \subseteq \rho))} \geq 1 - (\frac{1}{n})^{\frac{1}{2}}. \qquad \square$$

Lemma 2.8.7 *Suppose that μ satisfies condition (1) and (2) in corollary 2.5.3. Let $w = \max_\tau(\mu(\tilde{\mathcal{W}}_I(\tau))$. Then for each strategy v of player II in the first game*

$$\mu(\tilde{\mathcal{W}}_I(v)) \geq 1 - (1-w)^s \geq 1 - (\frac{1}{n})^{\frac{s}{2}}.$$

Proof: The task for player II to survive round 1 of the game $G'(n,k,t,s)$ (if player I selects the reply ρ randomly) is "easier" than the task of surviving any specific round j of the game $G(n,k,t,s,< h^1, h^2, ..., h^v >)$. More formally the probability Pr(survives round $k \mid$ history of the game) is

$$\geq \min_j Pr(h^j \subseteq \rho \mid \ h^j \mid\mid \rho \wedge (\rho \perp h^1) \wedge ... \wedge (\rho \perp h^{j-1})) \geq \min_j \Pr(h^j \subseteq \rho \mid \quad h^j \mid\mid \rho).$$

$\square$

Lemma 2.8.8

$$\mu(\mathcal{W}_I) \geq 1 - (\Sigma_v(\mu(\tilde{\mathcal{W}}_I(v)))) \geq 1 - (\frac{1}{n})^{\frac{s}{2}} \cdot n^{\frac{ps}{k}} \geq 1 - (\frac{1}{n})^{\frac{ks-2ps}{2k}} > 0$$

when $k \geq 2p+1$.

Proof: The number of strategies for player II in the first game is $\leq n^{\frac{ps}{k}}$. $\square$
This completes the proof of theorem 2.8.4. Now lemma 2.7.7 follows by combining lemma 2.8.3 and theorem 2.8.4.

2.9 Some consequences

Suppose that $\mathbb{M}$ is a countable non-standard model of $\mathrm{Th}(\mathbb{N})$ in some countable first order language L. Suppose L extends the language of Arithmetic. Let $p \geq 2$, $p \in \omega$ and let $n \in \mathbb{M} \setminus \omega$. Assume that n not is divisible by p. Let

$$\mathbb{M}_n^* := \{m \in \mathbb{M} : \ t(n) > m \text{ for some term } t \in L\}.$$

So far we are able to prove:

Theorem 2.9.1 (weak version) *If all terms $t \in L$ have sub-exponential growth rate, then for each generic filter ρ_G $(\mathbb{M}_n^*, \tilde{\rho}_G) \models \neg\mathrm{Count}(p)$. On the other hand $(\mathbb{M}_n^*, \tilde{\rho}_G)$ satisfies induction for bounded L_P-formulas. As above $\tilde{\rho}_G := \cup_{\alpha \in \rho_G} \alpha$.*

In the next two sections I strengthen this result. I show that the the model $(\mathbb{M}_n^*, \tilde{\rho}_G)$ satisfies the $\mathrm{Count}(q)$ principle exactly when certain exceptional forests do not exists.
Proof: The argument is very similar to the argument in [2]. So I only outline the argument.

It suffices to show that the least number principle is valid for bounded L_P-formulas with parameters in $\mathbb{M}_n^*$. Now translate each instance of the least number principle into a Boolean circuit of the form $\mathrm{LNP}_u(\pi_1, \pi_2, .., \pi_u) := \pi_u \vee (\vee_{j \leq u}(\neg\pi_j \wedge (\wedge_{k<j} \pi_k)))$. According to the general collapsing result from section 1, each π_j can be replaced (and this can be done simultaneously) by disjunction of small positive conjunctions (or by negations of disjunctions of small positive conjunctions). According to the key lemma (lemma 2.4.1) for <u>any</u> generic filter ρ_G if $(\pi_u)^{\tilde{\rho}_G} = 0$ there exists $j_0 \leq u$ with $(\pi_{j_0})^{\tilde{\rho}_G} = 0$ and with $(\pi_j)^{\tilde{\rho}_G} = 1$ for all $j < j_0$. A simple argument shows that $\mathrm{LNP}_u(\pi_1, ..., \pi_u)^{\tilde{\rho}_G} = 1$. By lemma 2.1.6 $(\mathbb{M}_n^*, \tilde{\rho}_G)$ satisfies induction for bounded L_P-formulas with parameters in $\mathbb{M}_n^*$. □

3 Forests of decision trees

The specially labelled trees we are going to consider can also be viewed as decision trees [6] . In our case the decisions concern a (hypothetical) partitioning of a finite set $I := \{1, 2, ..., n\}$ into disjoint p element subsets. To avoid trivial counter examples we always assume n is much larger than both p and the hight of the trees. All trees are rooted and finite (in later parts of the argument "finite" in the sense of a non-standard model of first order Arithmetic). When we follow a branch from the root towards the leafs we make successive decisions building up (parts of) some mathematical object. In this case a partial partitioning of I into disjoint p-element subsets. At each vertex v, except at the leafs, there is assigned a "question" $i_v \in I$. At the vertex v we are asked to decide which p element subset $A \subseteq I$ the element i_v belongs to. All possible choices which define the partitioning at i_v have to be represented. There is a one to one correspondence between possible choices (at i_v) and the sons from i_v. The label α of a branch is identified with the final object (here a partial partitioning) which has been constructed.

Suppose that we are given a forest $T_1, T_2, ..., T_u$ of decision trees. If each object (label

[6] I think that this view is due to P.Beame and T.Pitassi

on branch) appears 0 modulo q times, does q divide u? If there exists a global object (in this case when p divides n) the answer is always positive.

This type of question has not previously been considered in the literature. For almost any mathematical structure, it is possible to define such decision trees. They specify the local diagrams. In section 6 our analysis naturally leads us to consider another type of decision trees. Now let us focus on (p,n)-labelled trees. Notice first that each (p,n)-labelled tree is a graphical representation of a tree-like basis. Because of this, the concepts from section 2 (like *conditions* and *restrictions*) will keep their obvious meaning.

My aim is show that we have the following characterisation.

Theorem *Let $q, p \geq 2$ and $h \in \mathbb{N}$. Suppose that $h > q$. Then the following statements always hold simultaneously.*

(a) *All prime factors in p divide q.*

(b) *There exists n_0 such that for all $n > n_0$ which are not divisible by p there is a (p,n)-labelled forest $T_1, T_2, ..., T_u$ such that:*

(i) *All trees have hight $\leq h$.*

(ii) *Each type of branch appears 0 modulo q times.*

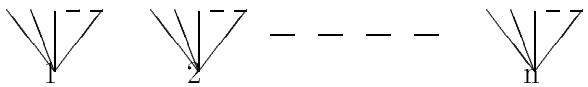
(iii) *$u \neq 0$ modulo q.*

Later I also discuss the general case where there are less restrictions on the (asymptotic) hight of the trees.

3.1 Some easy results

First let me illustrate the definition with a few trivial observations.

Example 3.1.1 *Suppose that p divides q and that p does <u>not</u> divide n. Consider the forest*



It contains n trees ($\neq 0$ modulo q) trees. Each branch appears exactly q (=0 modulo q) times.

This type of forests are so simple that we don't consider them as exceptional. They correspond to the fact that in the special case where p divides q, Count(p) is a (trivial) consequence of Count(q).

Observation 3.1.2 *Suppose that $\mathcal{F}$ is a forest of (p,n)-labelled trees. Suppose that each branch appears 0 modulo q times. Then there exists a forest $\mathcal{F}'$ of $(p,n+p)$-labelled trees such that*

(i) *The forest $\mathcal{F}'$ contains the same number of trees as $\mathcal{F}$.*

(ii) *Each branch in $\mathcal{F}'$ appears 0 modulo q times.*

(iii) *The hight of the highest trees in $\mathcal{F}'$ is at most 1 higher than the highest tree in $\mathcal{F}$.*

This immediately shows that if there exists an exceptional forest $\mathcal{F}$ for some n, this will also be the case for all larger n' as long as $n' = n$ modulo p. Here are two easy negative results,

Theorem 3.1.3 *Suppose that p divides n. Suppose that $T_1, T_2, ..., T_u$ is a (p,n)-labelled forest where each branch appears 0 modulo q times. Then $u = 0$ modulo q.*

Proof: According to the assumption p divides n so there exists a partitioning ρ_{global} of $\{1,2,...,n\}$ into disjoint sets $A_1, A_2, .., A_{\frac{n}{p}} \subseteq \{1,2,..,n\}$ each containing p elements. The partition ρ_{global} extends exactly one branch from each tree. Clearly, ρ_{global} allows us to define a partitioning of the trees $T_1, T_2, ..., T_u$ into disjoint classes each containing exactly q trees. □

Using a similar idea we notice

Theorem 3.1.4 *Suppose that $T_1, T_2, ..., T_u$ is a forest of (p,n)-labelled trees. Suppose that the sum of the heights of all trees is smaller than $\frac{n}{p}$. If all branches appear 0 modulo q times, then $u = 0$ modulo q.*

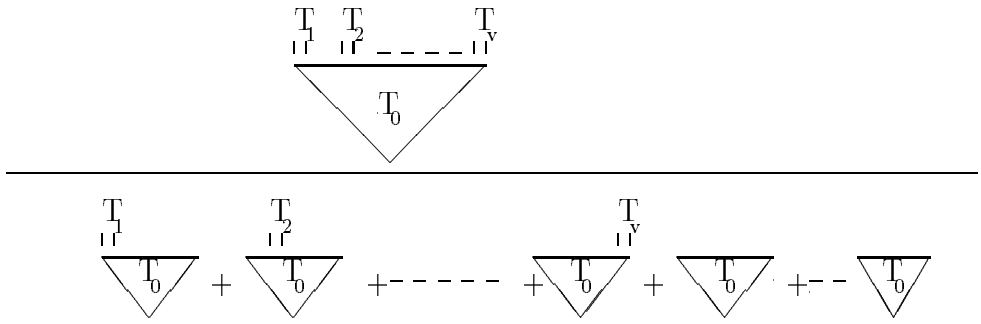
Proof: Select a branch $\beta_1 := \alpha_1$ from the tree T_1. The branch must be compatible with at least one branch $\alpha_2 \in T_2$. Let $\beta_2 := \beta_1 \cup \alpha_2$. This branch (=condition) must be compatible to at least one branch $\alpha_3 \in T_3$. Eventually we construct a condition ρ which extends exactly one branch in each tree. □

One can try to elaborate on this type of argument. A (very naive) strategy is to try to choose short branches from each tree. It is not hard to see that this method breaks down when $u > n$.

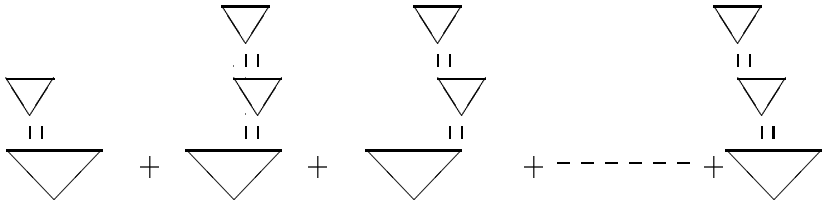
In [24] I presented a graph theoretical argument. It used a generalisation of a well-known theorem from graph theory. This theorem states that if in a graph G all vertex have degree at least as large as $\frac{1}{2} \cdot \mid G_{\text{vertex}} \mid$, then G contains a Hamiltonian circuit. This type of argument breaks down even when u is significantly smaller than n^2. Very early in this work it was clear that results relevant for Bounded Arithmetic all would require techniques which at least would be able to deal (when n tends to infinity) with the case where $u > n^k$ for arbitrarily fixed k.

3.2 Breaking down trees

Let T be a (p,n)-labelled tree. Consider the following "move".



Suppose that $\mathcal{F} := \{T_1, ..., T_u\}$ is any forest. Repeated application of this allows us to break down the trees in $\mathcal{F}$. Eventually all trees can be brought on the following normal form.



Let us call such trees *perfectly unbalanced* (=**PU**). Clearly we have,

Lemma 3.2.1 *Fix $q \geq 2$, $q \in \mathbb{N}$. Let $\mathcal{F} := \{T_1, ..., T_u\}$ be any forest. There exists a forest $\mathcal{F}' := \{T'_1, T'_2, ..., T'_{u'}\}$ in which each (type of) branch counted modulo q appears the same number of times as in $\mathcal{F}$. Furthermore (also counted modulo q) the number u' of trees in $\mathcal{F}'$ equals the number u of trees in the forest $\mathcal{F}$.*

Notice that the **PU**-trees have a very simple representation. The **PU**-tree represented by,

$$(u_{1,1}, \{u_{1,2}, ..., u_{1,p}\})(u_{2,1}, \{u_{2,2}, ..., u_{2,p}\})...(u_{h-1,1}, \{u_{h-1,2}, ..., u_{h-1,p}\})(u_h),$$

where $u_{i,j} \in I$.

3.3 Bringing the forest on a special normal form

Now here are some nice operations on **PU**-trees.

Example 3.3.1 *Consider the* **PU***-trees*

$$T := (4, \{5, 6\})(1, \{2, 3\})(7)$$

and

$$T' := (1, \{2,3\})(4, \{5,6\})(7).$$

Notice that T and T' contain the same branches of length 3. The branch $\beta := \{\{1,2,3\}, \{4,5,6\}\}$ does not appear in T and T'. Except for the branch β the tree T contains the same branches of length ≤ 2 as the tree $(4, \{5,6\})(1)$. Also, except for the branch β the tree T' contains the same branches of length ≤ 2 at the tree $(1, \{2,3\})(4)$.

This can be expressed by the equation,

$$(4, \{5,6\})(1, \{2,3\})(7) = (1, \{2,3\})(4, \{5,6\})(7) - (4, \{5,6\})(1) + (1, \{2,3\})(4).$$

The equation expresses the fact that both sides of the identity contain 1 tree (counted with signs). And it expresses the fact that both sides contain exactly the same set of branches.

Here is another operation.

Example 3.3.2 *We have the identity,*

$$(2, \{1,3\})(4) = (1, \{2,3\})(4) - (1) + (2).$$

The identities from the examples can be expressed generally.

Lemma 3.3.3 *We have the following identities.*

(1) $$(w_1, W_1)(w_2, W_2)...(b, B)(a, A)...(w_h) =$$

$$(w_1, W_1)(w_2, W_2)...(a, A)(b, B)...(w_h) - (w_1, W_1)(w_2, W_2)...(a) + (w_1, W_1)(w_2, W_2)...(b).$$

(2) $$(w_1, W_1)(w_2, W_2)...(a_2, \{a_1, a_3, .., a_p\})...(w_h) =$$

$$(w_1, W_1)...(a_1, \{a_2, a_3, .., a_p\})...(w_h) - (w_1, W_1)...(a_2) + (w_1, W_1)...(a_1).$$

It follows immediately from these principles that,

Lemma 3.3.4 *Let $\mathcal{F}' := \{T'_1, .., T'_{u'}\}$ be a forest where all trees are* **PU***-trees. Then there exist a forest $\mathcal{F}'' := \{T''_1, ...T''_{u''}\}$ where all trees are of the form*

$$(u_{1,1}, \{u_{1,2}, .., , u_{1,p}\})...(u_{i,1}, \{u_{i,2}, ..., u_{i,p}\})...(u_h)$$

where $u_{1,1} < u_{2,1} < ... < u_{h-1,1}$, and where $u_{i,1} < u_{i,2} < ... < u_{i,p}$ for $i = 1, 2, .., h-1$. Furthermore,

(i) *The forests $\mathcal{F}'$ and $\mathcal{F}''$ contain the same number of trees (modulo q).*

(ii) *Each (type of) branch appears the same number (modulo q) of times in the forests $\mathcal{F}'$ and $\mathcal{F}''$.*

We will come back to this normal form.

4 The first main result

Suppose that $\mathbb{M}$ is a countable non-standard model of Th($\mathbb{N}$) over a countable first order language L, which extends the language of Arithmetic. Suppose that $p \geq 2$ and $I := \{1, 2, ..., n\} \subseteq \mathbb{M}$ for some $n \in \mathbb{M} \setminus \omega$. Assume that n is not divisible by p. As above, let $\mathbb{M}_n^* := \{m \in \mathbb{M} : \ t(n) > m, \text{ for some term } t \in L\}$.

Theorem 4.0.5 (Main result) *Suppose that all terms $t \in L$ have sub-exponential growth rate. Then for each generic filter ρ_G (see definition 2.2.5, page 12),*

(a) $(\mathbb{M})_n^*, \tilde{\rho}_G) \models \neg$ Count(p).

(b) $(\mathbb{M}_n^*, \tilde{\rho}_G)$ *satisfies induction for bounded L_P-formulas.*

(c) $(\mathbb{M}_n^*, \tilde{\rho}_G)$ *satisfies (all versions of) the pigeon-hole principle for bounded L_P-formulas.*

Furthermore, there exists a sequence $s_k(x), k = 1, 2, ...$ of (arithmetical) functions (which depend on the exact growth rate of the terms in L), such that (under the harmless extra assumption that the underlying language L might need an extension) the following are equivalent:

(i) $(\mathbb{M}_n^*, \tilde{\rho}_G)$ *satisfies the Count(q) principle.*

(ii) *Each forest $T_1, T_2, .., T_u$ of (p, n)-labelled trees in which all trees have hight $\leq s_k(n)$ where each branch appears 0 modulo q times, has $u = 0$ modulo q.*

(iii) *As (ii) but for (p, n)-labelled* **PU***-trees.*

Suppose that all terms in L have (at most) polynomial growth rate. Then $s_k(x) := k$ gives the required characterisation.

In general $s_k(x)$ can be chosen such that $(s_k(n))^l < n$ for all $l \in \mathbb{N}$.

Our overall question is when systems of Bounded Arithmetic extended by an axiom scheme for the Count(q) principle, are able to prove Count(p). The first main result, links this to an understanding of the structure of exceptional forests. Furthermore, it shows that the asymptotic hight of the trees in the minimal exceptional forests is directly linked to the strength of the underlying axiom system.

I have already proved (a) and (b). I have also showed (iii) $\Rightarrow$ (ii). To show (i) $\Rightarrow$ (iii) assume that there is a forest $\mathcal{F}$ which violates (iii). Assume that the language contains a suitable relation symbol which allows us to define the forest by a Bounded formula (this is the harmless extra assumption). I claim that the Count(q) principle fails in $(\mathbb{M}_n^*, \tilde{\rho}_G)$. To see this, notice that there is a Bounded L_P-formula with parameters in $\mathbb{M}_n^*$ which defines (by use of $\tilde{\rho}_G$ a partitioning of the trees in the

forest $\mathcal{F}$. And this in such a way that each class contains exactly q trees. But by assumption $\mathcal{F}$ contains trees $T_1, T_2, ..., T_u$ for some u with $u \neq 0$ modulo q.

The difficult implication is (iii) $\Rightarrow$ (i). As it turns out our proof of (c) also provides a first step in showing the implication (i) $\Rightarrow$ (iii).

Lemma 4.0.6 *Suppose that for some $a \in \mathbb{M}_n^*$, some bounded $L_{\mathbb{M}}(P)$-formula $\theta(\cdot,\cdot)$ (with parameters in $\mathbb{M}_n^*$) defines a bijection from a to b. Let $\rho_0 \in \rho_G$ be given. Then there exists an $\mathbb{M}$-definable sequence $\mathcal{H}_{i,j}$ and $\mathcal{B}_{i,j}$, $(i,j) \in a \times b$ such that for some $\rho \supseteq \rho_0$, $\rho \in \mathcal{P}$:*

(i) *for all $(i,j) \in a \times b$, $\mathcal{B}_{i,j}$ is a tree-like basis on $I \setminus \mathrm{Set}(\rho)$,*

(ii) *for all $(i,j) \in a \times b$, $\| \mathcal{B}_{i,j} \| \leq t$ for some fixed $t < n^{\frac{1}{\omega}}$,*

(iii) *for all $(i,j) \in a \times b$, $\mathcal{H}_{i,j} \subseteq \mathcal{B}_{i,j}$,*

(iv) *For each $i_0 \leq a$: $\mathcal{B}_{i_0}^{\rightarrow} := \cup_{j \leq b} \mathcal{H}_{i_0,j}$ is a basis for $\mathcal{P}$,*

(v) *For each $j_0 \leq b$: $\mathcal{B}_{j_0}^{\downarrow} := \cup_{i \leq a} \mathcal{H}_{i,j_0}$ is a basis for $\mathcal{P}$.*

Proof: Suppose that some Bounded $L_{\mathbb{M}}(P)$-formula $\theta(\cdot,\cdot)$ defines a bijection from $h : \{1,2,..,a\}$ onto $\{1,2,..,b\}$ for $a \neq b$. According to lemma 2.1.5 there exists $d \in \omega$ and $t \in \mathbb{M}_n^*$ and a $\mathbb{M}$-definable sequence of circuits $\theta_{i,j}$ $(i,j) \in a \times b$, such that each $(\theta_{i,j})^{\tilde{\rho}_G}$ holds exactly when $(\mathbb{M}_n^*, \tilde{\rho}_G) \models \theta(i,j)$. Now according to the key lemma (lemma 2.4.1) there exists $\rho \supseteq \rho_0$, $\rho \in \mathcal{P}$, an $\mathbb{M}$-definable sequence $\mathcal{B}_{i,j}$ $(i,j) \in a \times b$ where each $\mathcal{B}_{i,j}$ is a tree-like basis, and an $\mathbb{M}$-definable sequence $\mathcal{H}_{i,j} \subseteq \mathcal{B}_{i,j}$ $(i,j) \in a \times b$ such that for each $(i,j) \in a \times b$:

$$\theta_{i,j} \equiv_{\rho} \vee_{h \in H_{i,j}} h.$$

We claim that the sequences $\mathcal{B}_{i,j}$ and $\mathcal{H}_{i,j}$ satisfy (i)-(v). By use of the fact that h is an injective function it is straightforward to show that the conditions in $\mathcal{B}_i^{\rightarrow}$ must be pairwise incompatible. The fact that h is a (mono-valued) function ensures that conditions in $\mathcal{B}_j^{\downarrow}$ are pairwise incompatible.

The only problem is to show that each $\mathcal{B}_i^{\rightarrow}$ and each $\mathcal{B}_j^{\downarrow}$ are complete for $\mathcal{P}(I \setminus \mathrm{Set}(\rho))$ (i.e. satisfies condition (2) in definition 2.5.3). We can simplify the notation by assuming that $\rho = \emptyset$. This simplification is possible by lemma 2.3.13 because the lemma allows us to replace I by $I \setminus \mathrm{Set}(\rho)$.

Suppose that $\rho' \in \mathcal{P}$ is incompatible with all conditions in $\mathcal{B}_i^{\rightarrow}$ for some fixed $i \leq a$. Let ρ_G be a generic filter (without the simplification we assume $\rho_G \ni \rho$). Now for each $j \leq b$, $\tilde{\rho}_G$ is incompatible with all conditions in $\mathcal{H}_{i,j}$, so by use of lemma 2.3.14 $\theta_{i,j}^{\tilde{\rho}_G} = 0$ for all $j \leq b$. This is in contradiction with lemma 2.3.4 which ensures that $\emptyset$ (or ρ in the un-simplified case) forces h to take a value $j \leq b$.

The completeness of the conditions in each $\mathcal{B}_j^{\downarrow}$ follows by use of the assumption that h was forced onto. □

4.1 Using a combinatorial phenomenon

My aim here is to show that (i)-(v) in lemma 4.0.6 can be only satisfied when $a = b$. First we show

Lemma 4.1.1 *Suppose that for some $a, b \in \mathbb{M}_n^*$ there exist $\mathbb{M}$ definable sequences $\mathcal{B}_{i,j}$ and $\mathcal{H}_{i,j}$ $(i,j) \in a \times b$. If they satisfy condition (i)-(iii) in lemma 4.0.6, together with:*

(iv)′ *For each $i_0 \leq a$: $\mathcal{B}_{i_0}^{\rightarrow} := \cup_{j \leq b} \mathcal{H}_{i_0,j}$ is a tree-like basis.*

(v)′ *For each $j_0 \leq b$: $\mathcal{B}_{j_0}^{\downarrow} := \cup_{i \leq a} \mathcal{H}_{i,j_0}$ is a tree-like basis.*

Then $a = b$.

Proof: First, notice that we can assume that all conditions $h, h' \in \mathcal{H}_{i,j}$ have $\mid h \mid = \mid h' \mid$. Otherwise make suitable tree-like refinements. Second, notice that $\mathcal{P}$ (the set of forcing conditions), has the property that the number $N(n,p,c)$ of conditions in a tree-like basis where all conditions h have $\mid h \mid = c$, only depends on n, p and c. Now $ac = \Sigma_{i \leq a} \mid \mathcal{B}_i^{\rightarrow} \mid = \Sigma_{i \leq a, j \leq b} \mid \mathcal{H}_{i,j} \mid = \Sigma_{j \leq b} \mid \mathcal{B}_j^{\downarrow} \mid = bc$, so $a = b$. □

Suppose that we could replace "is a basis for $\mathcal{P}$" with "is a tree-like basis" in lemma 4.0.6. Then according to lemma 4.1.1 this would ensure that the pigeon-hole principle could never be forced false. So if a basis $\mathcal{B}$ for $\mathcal{P}$ in general would be tree-like, we would be done. Unfortunately, the reality is more complex.

Example 4.1.2 *The converse of lemma 2.7.4 does not hold in general. The following example ($p = 2$) is due to Krajicek (personal communication). The collection*

$$\mathcal{B} := \{\{\{1,2\}\}, \{\{1,3\}\}, \{\{2,3\}\}, \{\{1,i\}, \{2,j\}, \{3,k\}\}_{i,j,k \geq 4 \wedge \mid \{i,j,k\} \mid = 3}\}$$

is a basis for $\mathcal{P}$. However $\mathcal{B}$ is <u>not</u> a tree-like basis (there is no $i_0 \in I$ such that all $\beta \in \mathcal{B}$ has $i_0 \in \mathrm{Set}(\rho)$).

Observation 4.1.3 *Consider example 4.1.2. Let $\mathcal{B}' := (\mathcal{B} \cup \mathcal{E}_1(\{\{2,3\}\})) \setminus \{\{\{2,3\}\}\}$, so $\mathcal{B} \rightarrow_{\mathrm{TR}} \mathcal{B}'$. Notice that $\mathcal{B}'$ is a tree-like basis. To see this, notice that $\mathcal{B}'$ can be obtained from $\{\emptyset\}$ by the atomic tree-like refinements:*

$$\mathcal{E}_1(\emptyset),$$

$$\mathcal{E}_2(\{1,4\}), \mathcal{E}_2(\{1,5\}), ..., \mathcal{E}_2(\{1,n\}),$$

$$\mathcal{E}_3(\{\{1,4\},\{2,5\}\}), \mathcal{E}_3(\{\{1,4\},\{2,6\}), ..., \mathcal{E}_3(\{\{1,4\},\{2,n\}\}),$$

$$\mathcal{E}_3(\{\{1,5\},\{2,4\}\}), \mathcal{E}_3(\{\{1,5\},\{2,6\}\}), ..., \mathcal{E}_3(\{\{1,5\},\{2,n\}\}),$$

$$....$$

$$\mathcal{E}_3(\{\{1,n\},\{2,4\}\}), \mathcal{E}_3(\{\{1,n\},\{2,5\}\}), ..., \mathcal{E}_3(\{\{1,n\},\{2,n-1\}\}).$$

This observation is part of a general phenomenon. It turns out (and this was the combinatorial discovery which made my general approach possible), that any basis $\mathcal{B}$ for $\mathcal{P}$ has a tree-like refinement to a tree-like basis.

Lemma 4.1.4 *Assume that $\mathcal{B}$ is a basis for $\mathcal{P}$, and that $u \in I$. Then there exists a tree-like refinement $\mathcal{B}'$ of $\mathcal{B}$ such that for all $\beta' \in \mathcal{B}'$ $u \in \mathrm{Set}(\beta')$.*

Proof: Let $\mathcal{B}' := \cup_{\beta \in B} \mathcal{E}_u(\beta)$. Notice that this is actually a tree-like refinement of $\mathcal{B}$, and that $\mathcal{B}'$ has the required properties. □

Definition 4.1.5 For $U \subseteq I$, we let $\mathcal{C}_U$ denote the tree-like basis $\{\alpha : \forall A \in \alpha\ \exists u \in U\ u \in A \wedge \mathrm{Set}(\alpha) \supseteq U\}$. We say $\mathcal{B}$ is a *tree-like basis on $U \subseteq I$* if for each $\alpha \in \mathcal{C}_U$, there exists $\beta \in \mathcal{B}$ with $\beta \supseteq \alpha$. ♣

Lemma 4.1.6 *Suppose that $\mathcal{B}$ is a basis for $\mathcal{P}$, and $U \subseteq I$ with $\mid U \mid < n^{\frac{1}{\omega}}$. There is a tree-like refinement $\mathcal{B}'$ of $\mathcal{B}$, such that $\mathcal{B}'$ is a tree-like basis on U.*

Proof: Let $U = \{u_1, u_2, ..., u_r\}$. According to lemma 4.1.4 there exists a sequence

$$\mathcal{B} = \mathcal{B}_0 \rightarrow_{\mathrm{TR}} \mathcal{B}_1 \rightarrow_{\mathrm{TR}} .... \rightarrow_{\mathrm{TR}} \mathcal{B}_r,$$

such that for all $\beta \in \mathcal{B}_j$ $u_j \in \mathrm{Set}(\beta)$. Let $\mathcal{B}' := \mathcal{B}_r$. We have to show that for each $\alpha \in \mathcal{C}_U$ there exists $\beta \in \mathcal{B}$', $\beta \supseteq \alpha$. Now by use of a calculation similar to the one in the proof of lemma 2.3.12, $\mathcal{B}$' is a basis for $\mathcal{P}$, so each $\alpha \in \mathcal{C}_U$ is compatible with some $\beta \in \mathcal{B}$'. Now as $\mathrm{Set}(\beta) \supseteq U$ actually $\beta \supseteq \alpha$. □

Lemma 4.1.7 *Suppose that $\mid\mid \mathcal{B} \mid\mid \leq t$ for some $t < n^{\frac{1}{\omega}}$. Also suppose that the conditions in $\mathcal{B}$ are pairwise incompatible. Then $\mathcal{B}$ is a basis for $\mathcal{P}$ iff each condition $\rho \in \mathcal{P}_1$ is compatible with some $\beta \in \mathcal{B}$.*

Proof: Repeated application of lemma 4.1.6. □

Lemma 4.1.8 *If $\mathcal{B}$ is a basis for $\mathcal{P}$, and $\mid\mid \mathcal{B} \mid\mid \leq t$ for some $t < n^{\frac{1}{\omega}}$, then there exists a tree-like basis $\tilde{\mathcal{B}}$ such that $\mid\mid \tilde{\mathcal{B}} \mid\mid < pt(t+1)$ and such that $\mathcal{B} \rightarrow_{\mathrm{TR}} \tilde{\mathcal{B}}$.*

Proof: First we construct $\tilde{\mathcal{B}}$'. Pick a set $V \subseteq I$ such that $V := \mathrm{Set}(\beta)$ for some $\beta \in \mathcal{B}$. According to lemma 4.1.6 there exists a tree-like refinement $\mathcal{B}_1$ of $\mathcal{B}$ such that $\mathcal{B}_1$ is tree-like on V. Now fix $\gamma \in \mathcal{C}_V$ and consider $\mathcal{B}_1^\gamma \subseteq \mathcal{P}(I \setminus V)$. It is not hard to show $\mathcal{B}_1^\gamma$ is a basis for $\mathcal{P}(I \setminus V)$. Now by use of lemma 4.1.7 we notice that we can prove the lemma by use of induction after t inside $\mathbb{M}$. Let $\mathcal{B}_1(\gamma) := \{\beta : \beta' \cap \gamma = \beta,\ \beta' \in \mathcal{B}_1^\gamma\}$. Notice that $\mathcal{B}_1(\gamma)$ is a tree-like refinement of γ. Finally let $\tilde{\mathcal{B}} := \cup_\gamma \mathcal{B}_1(\gamma)$. By induction after t we have $\| \tilde{\mathcal{B}} \| \leq |\mathrm{Set}(\gamma)| + p(t-1)t$. Now $|\mathrm{Set}(\gamma)| \leq pt$, from which the required inequality follows. □

We need a two-dimensional version of lemma 4.1.8.

Lemma 4.1.9 *Suppose that there exists an $\mathbb{M}$-definable "generic system". That is a sequence $\mathcal{H}_{i,j},\ (i,j) \in a \times b$ such that:*

(i) *For each $i \leq a$ $\mathcal{B}_i^{\rightarrow} := \cup_{j \leq b} \mathcal{H}_{i,j}$ is a basis for $\mathcal{P}$.*

(ii) *For each $j \leq b$ $\mathcal{B}_j^{\downarrow} := \cup_{i \leq a} \mathcal{H}_{i,j}$ is a basis for $\mathcal{P}$.*

(iii) *$max_{(i,j) \in a \times b} \| \mathcal{H}_{i,j} \| \leq t$ for some $t < n^{\frac{1}{\omega}}$.*

Then there exists an $\mathbb{M}$-definable "tree-like generic system". That is a sequence $\tilde{\mathcal{H}}_{i,j}$ $(i,j) \in a \times b$ such that:

(i)' *For each $i \leq a$ $\tilde{\mathcal{B}}_i^{\rightarrow} := \cup_{j \leq b} \tilde{\mathcal{H}}_{i,j}$ is a tree-like basis.*

(ii)' *For each $j \leq b$ $\tilde{\mathcal{B}}_j^{\downarrow} := \cup_{i \leq a} \tilde{\mathcal{H}}_{i,j}$ is a tree-like basis.*

(iii)' *$max_{(i,j) \in a \times b} \| \tilde{\mathcal{H}}_{i,j} \| \leq p^3(t+1)^4$.*

Proof: Fix $i \leq a$. According to lemma 4.1.8 there exists a tree-like refinement $\mathcal{B}_i'^{\rightarrow}$ of $\mathcal{B}_i^{\rightarrow}$, which is a tree-like basis. For each $j \leq b$ this procedure induces a tree-like refinement $\mathcal{H}'_{i,j}$ of $\mathcal{H}_{i,j}$. This way we get an $\mathbb{M}$-definable sequence $\mathcal{H}'_{i,j},\ (i,j) \in a \times b$, so (i)', (ii) and $\| \mathcal{H}'_{i,j} \| \leq pt(t+1)$.

Now fix $j \leq b$. Again according to lemma 4.1.8 there exists a tree-like refinement $\tilde{\mathcal{B}}_j^{\downarrow}$ of $\mathcal{B}_j'^{\downarrow}$, which is a tree-like basis. For each $i \leq a$ this procedure induces a tree-like refinement $\tilde{\mathcal{H}}_{i,j}$ of $\mathcal{H}'_{i,j}$. Now notice that $\tilde{\mathcal{B}}_i^{\rightarrow}$ remains tree-like basis, and thus the $\mathbb{M}$-definable sequence $\tilde{\mathcal{H}}_{i,j},\ (i,j) \in a \times b$ satisfies (i)', (ii)'. Clearly also (iii)' holds because $\| \tilde{\mathcal{H}}_{i,j} \| \leq p(pt(t+1)+1)(pt(t+1)) \leq p^3(t+1)^4$. □

This immediately shows (c) in theorem 4.0.5, in the case of the bijective pigeon-hole principle. The other versions of the pigeon-hole principle are treated with minor changes.

4.2 Reducing the Count(q) versus Count(p) problem

The implication (i) $\Rightarrow$ (iii) follows by the same type of argument.

Lemma 4.2.1 *Suppose that $\theta(x_1, x_2, ..., x_q)$ is a bounded $L(P)$-formula with q free variables, and all its parameters in $(\mathbb{M}_n^*, \tilde{\rho}_G)$. If θ defines a partition of $I_a := \{1, 2, .., a\}$, $a \in \mathbb{M}_n^*$, then there exists an $\mathbb{M}_n^*$-definable map $A \to \mathcal{H}_A$, which to each q-subset A of I_a assigns a collection of conditions $\mathcal{H}_A \subseteq \mathcal{P}$ such that for some $t < n^{\frac{1}{\omega}}$, $max_A(|| \mathcal{H}_A ||) \leq t$. Furthermore, for each $v \in I_a, \mathcal{B}_v := \cup_{A \subseteq I_a, |A|=q, v \in A} \mathcal{H}_A$ is a basis for $\mathcal{P}$.*

Proof: Suppose that some bounded $L(P)$-formula $\theta(x_1, x_2.., x_q)$ defines a partition of $\{1, 2, ..., a\}$ into disjoint q-subsets, and q does not divide a. According to 2.1.5 there exists $d \in \omega, t < n^{\frac{1}{\omega}}$ and an $\mathbb{M}$-definable sequence of circuits $\theta_{v_1,...,v_q}$, $v_1, ..., v_q \in I_a$, such that $\theta^{\tilde{\rho}_G}_{i_1,...,i_q}$ exactly when $(\mathbb{M}_n^*, \tilde{\rho}_G) \models \theta(i_1, ..., i_q)$. Now according to the key lemma (lemma 2.4.1), there exists $\rho \supseteq \rho_0$ (for any given ρ_0), and an $\mathbb{M}$-definable sequence $\mathcal{B}_{v_1,..,v_q}$ $v_1, .., v_q \in I_a$ where each $\mathcal{B}_{v_1,..,v_q} \subseteq \mathcal{P}$ is a tree-like basis with $|| \mathcal{B}_{v_1,...,v_q} || \leq t$. Furthermore, there exists a $\mathbb{M}$-definable sequence $\mathcal{H}_{v_1,v_2,...,v_q} \subseteq \mathcal{B}_{v_1,..,v_q}$ such that for each $v_1, ..., v_q \in I_a$,

$$\theta_{v_1,v_2...,v_q} \equiv_\rho \vee_{h \in \mathcal{H}_{v_1,v_2,...,v_q}} h.$$

Fix $v \in I_a$ and consider $\mathcal{B}_v := \cup_{A \subseteq I_a, v \in A} \mathcal{H}_A$. For $\alpha, \beta \in \mathcal{B}_v$, $\alpha \neq \beta$ we claim $\alpha \perp \beta$. To see this notice that otherwise there would exist $\rho \supseteq \alpha \cup \beta$, and $\rho \in \mathcal{P}$ would force both $\theta_{v_1,...,v_q}$ and $\theta_{v'_1,...,v'_q}$ true. Now $v \in \{v_1, ..., v_q\} \cap \{v'_1, ..., v'_q\}$ so this is only possible when $\{v_1, ..., v_q\} = \{v'_1, ..., v'_q\}$. Thus both α and β belong to $\mathcal{H}_{v_1,..,v_q} \subseteq \mathcal{B}_{v_1,..,v_q}$. As $\mathcal{B}_{v_1,...,v_q}$ is a (tree-like) basis, $\alpha \perp \beta$.

It remains to show that $\mathcal{B}_v$, $v \in I_a$ is complete for $\mathcal{P}(\{1, 2, .., n\} \setminus \mathrm{Set}(\rho_0))$. Assume for the simplicity of the notation that $\rho_0 = \emptyset$. According to lemma 2.3.13, this assumption is harmless. We have to show that no $\rho \in \mathcal{P}$ is incompatible with all the conditions $h \in \mathcal{B}_v$. Now using lemma 2.3.14 each generic filter ρ_G contains some $h \in \mathcal{B}_{v_1,...,v_q}$ for each $v_1, v_2, .., v_q \in I_a$. But this contradicts the assumption that ρ_0 (in our case $\emptyset$) forces θ to define a total partition of I_a into disjoint q subsets. □

We conjecture that Count(q) is always forced true (when p and q are different primes. To show that Count(q) is never forced false, it suffices to show that if $\mathcal{H}_A$ is an $\mathbb{M}$-definable assignment as in lemma 4.2.1, then q must divide a.

Example 4.2.2 *Suppose that $a << n$. Consider $I_a := \{1, 2, ..., a\}$. Pick $\rho_1 \subseteq \rho_2 \subseteq .... \subseteq \rho_a$ such that for each $v \leq a$ there is $\alpha_v \in \mathcal{B}_v$ such that $\alpha_v \subseteq \rho_v$. This is possible whenever $\rho_v \in \mathcal{P}$, $v = 1, 2, ..., a$ (which is the case when $a << n$). Notice that ρ_a*

induces an $\mathbb{M}$-definable partition of I_a into disjoint q-subsets. As $\mathbb{M}$ shares its first order properties with $\mathbb{N}$, this is only possible when q divides a.

As a major step in solving the Count(q) versus Count(p) problem, I show that we can strengthen the conclusion by replacing 'each $\mathcal{B}_v$ is a basis for $\mathcal{P}$', with 'each $\mathcal{B}_v$ to be a tree-like basis'.

Lemma 4.2.3 *Let $t < n^{\frac{1}{\omega}}, a \in \mathbb{M}$. Let $\mathcal{P}$ be the set of forcing conditions defined as on page 11. Suppose that $A \to \mathcal{H}_A$ is an $\mathbb{M}$-definable map which assigns a collection of conditions $\mathcal{H}_A \subseteq \mathcal{P}$, to each q-subset A of $I_a = \{1, 2, .., a\}$ such that*

$$(i) \qquad \max_A(|| \mathcal{H}_A ||) \leq t$$

$$(ii) \qquad \mathcal{B}_v := \cup_{A \subseteq I_a, |A|=q, v \in A} \mathcal{H}_A \text{ is a basis for } \mathcal{P} \quad (v = 1, 2, ..., a).$$

Then there exists a $\mathbb{M}$-definable map $A \to \tilde{\mathcal{H}}_A$ which assigns a tree-like refinement $\tilde{\mathcal{H}}_A$ of $\mathcal{H}_A$, to each q-subset A of $I_a := \{1, 2, .., a\}$ such that

$$(i)' \qquad \max_A(|| \mathcal{H}_A ||) \leq qpt(t+1)$$

$$(ii)' \qquad \tilde{\mathcal{B}}_v := \cup_{A \subseteq I_a, |A|=q, v \in A} \tilde{\mathcal{H}}_A \text{ is a tree} - \text{like basis } (v = 1, 2, ..., a).$$

As a first attempt of a proof consider the following argument. According to lemma 4.1.8 there exists a tree-like basis $\mathcal{B}_1^{(1)}$ which is a tree-like refinement of $\mathcal{B}_1 := \cup_{1 \in A} \mathcal{H}_A$. This refinement induces tree-like refinements $\mathcal{H}_A \to_{\mathrm{TR}} \mathcal{H}_A^{(1)}$ for each $A \subseteq I_a, | A | = q$ (when $1 \notin A$, $\mathcal{H}_A^{(1)} = \mathcal{H}_A$). For each $v \in I_a$ let $\mathcal{B}_v^{(1)} := \cup_{v \in A} \mathcal{H}_A^{(1)}$.

Again by lemma 4.1.8 there exists a tree-like basis $\mathcal{B}_2^{(2)}$ which is a tree-like refinement of $\mathcal{B}_2^{(1)}$. This refinement induce a tree-like refinement $\mathcal{H}_A^{(1)} \to_{\mathrm{TR}} \mathcal{H}_A^{(2)}$ for each $A \subseteq I_a, | A | = q$ (when $2 \notin A$, $\mathcal{H}_A^{(2)} = \mathcal{H}_A^{(1)}$). For each $v \in I_a$ let $\mathcal{B}_v^{(2)} := \cup_{v \in A} \mathcal{H}_A^{(2)}$.

Eventually (again using lemma 4.1.8) there exists a tree-like basis $\mathcal{B}_2^{(a)}$ which is a tree-like refinement of $\mathcal{B}_a^{(a-1)}$. This refinement induces a tree-like refinement $\mathcal{H}_A^{(a-1)} \to_{\mathrm{TR}} \mathcal{H}_A^{(a)}$ for each $A \subseteq I_a, | A | = q$ (when $a \notin A$, $\mathcal{H}_A^{(a)} = \mathcal{H}_A^{(a-1)}$). For each $v \in I_a$ let $\mathcal{B}_v^{(a)} := \cup_{a \in A} \mathcal{H}_A^{(a)}$.

Let $\tilde{\mathcal{H}}_A := \mathcal{H}_A^{(a)}$. We claim that each $\tilde{\mathcal{B}}_v := \cup_{v \in A} \tilde{\mathcal{H}}(a)_A$ is a tree-like basis. To see this notice $\tilde{\mathcal{B}}_v = B_v^{(a)}$. By construction each $\mathcal{B}_v^{(v)}$ is a tree-like basis. Now

$$\mathcal{B}_v^{(v)} \to_{\mathrm{TR}} \mathcal{B}_v^{(v+1)} \to_{\mathrm{TR}} ... \to_{\mathrm{TR}} \mathcal{B}_v^{(a)}$$

so $\mathcal{B}_v^{(a)}$ is a tree-like basis.

This argument has to be adjusted. We have to ensure that all conditions h are small throughout the construction. To this end we need some more lemmas.

Definition 4.2.4 For $\mathcal{H}, \mathcal{H}' \subseteq \mathcal{P}$ let $\mathcal{H} \times \mathcal{H}' := \{h \cup h' : \ h \in \mathcal{H}, h' \in \mathcal{H}'\}$. ♣

Lemma 4.2.5 *Let $A \to \mathcal{H}_A$ and $A \to \mathcal{H}'_A$ be two $\mathbb{M}$-definable maps. Suppose that*

(i) *for each condition in $\mathcal{H}_A$ is compatible with some condition in $\mathcal{H}'_A$ and vice versa.*

(ii) *for A, B with $A \neq B$ and $A \cap B \neq \emptyset$, all conditions in $\mathcal{H}_A$ are incompatible with all conditions in $\mathcal{H}'_A$ and vice versa.*

Suppose that both the maps $A \to \mathcal{H}_A$ and $A \to \mathcal{H}'_A$ satisfy conditions (i) and (ii) in lemma 4.2.3. Then the $\mathbb{M}$-definable map $A \to \mathcal{H}_A \times \mathcal{H}'_A$ ensures that (i) and (ii) remain valid with t replaced by $2t$.

Proof: Direct verification. □

Lemma 4.2.6 *Suppose that $\mathcal{B}_0$ is a basis for $\mathcal{P}$. If both B_1 and B_2 are tree-like refinements of $\mathcal{B}_0$, then $\mathcal{B}_1 \times \mathcal{B}_2$ is a tree-like refinement of both $\mathcal{B}_1$ and $\mathcal{B}_2$.*

Proof: Proved by induction on the number of atomic tree-like refinements needed to get from $\mathcal{B}_0$ to $\mathcal{B}_1$ added to the number of atomic tree-like refinements needed to get from $\mathcal{B}_0$ to $\mathcal{B}_2$. □

The following proof simplifies an argument in an earlier and preliminary version of this paper.

Proof of lemma 4.2.3: For each $v \in I_a$ let $\mathcal{H}^v_A$ be the tree-like refinement of $\mathcal{H}_A$ induced when $B_v := \cup_{A \ni v} \mathcal{H}_A$ is refined to a tree-like basis $\mathcal{B}^v_v$. Consider the $\mathbb{M}$-definable map

$$A \to \tilde{\mathcal{H}}_A := \mathcal{H}^{a_1}_A \times \mathcal{H}^{a_2}_A \times .... \times \mathcal{H}^{a_q}_A,$$

where $A = \{a_1, a_2, ..., a_q\}$. Now $\mathcal{B}^v_v := \cup_{A \in v} \mathcal{H}^v_A$ is a tree-like basis. Furthermore, $\tilde{\mathcal{H}} \equiv \mathcal{H}^v_A \times (\mathcal{H}^{a_2}_A \times ... \times \mathcal{H}^{a_q}_A)$ where $A = \{v, a_2, a_3, ..., a_q\}$. To see this notice that according to lemma 4.2.6 for fixed $a_2, .., a_q$, $(\cup_{A \ni v} \mathcal{H}^v_A) \times (\mathcal{B}^{a_2} \times \mathcal{B}^{a_3} \times ... \times \mathcal{B}^{a_q})$ is a tree-like refinement of $\cup_{A \ni v} \mathcal{H}^v_A$. This tree-like refinement induces a tree-like refinement

$\mathcal{H}^v_A \to_{\mathrm{TR}} \mathcal{H}^v_A \times (\mathcal{H}^{a_2}_A \times ... \times \mathcal{H}^{a_q}_A)$.

Thus for each $v \in I_a$, $\tilde{\mathcal{B}}_v := \cup_{A \ni v} \tilde{\mathcal{H}}^v_A$ is a tree-like refinement of $\cup_{A \ni v} \mathcal{H}^v_A$, which was constructed as a tree-like basis. Thus each $\tilde{\mathcal{B}}_v$ is a tree-like basis. □

Combining these results it is not hard to show that (i) ⇒ (iii).

5 The positive part

In this section I develop a method of constructing exceptional forests. And the existence of these forests immediately gives us the positive part of the classification.

Definition 5.0.7 By $[i_1, ..., i_l]_p$ we denote the (p,n)-labelled trees which contain all the branches α of the form $\alpha = \{A_1, ..., A_r\}$ where $A_j \subseteq I$, $\mid A_j \mid = p$, $j = 1, 2, .., r$ and where $A_j \cap A_k = \emptyset$ for $j \neq k$. Besides that we require that,

(a) $A_j \cap \{i_1, i_2, ..., i_l\} \neq \emptyset$ for $j = 1, 2, ..., r$,

(b) $\forall k \leq l \; \exists j \leq r \; i_k \in A_j$. ♣.

Definition 5.0.8 Let $p, q \geq 2$. The forest $\mathcal{F}^{\mathrm{Hom}}_{p,r,n}$ (Hom for homogeneous) consists of all the trees $[i_1, i_2, ..., i_r]_p$ where $i_1 < i_2 < ... < i_r \leq n$. ♣

Definition 5.0.9 By $A_{p,l,r}$ we denote the number of ways it is possible to select r elements from the sets $\{1, 2, ..., p\}, \{p+1, p+2, ..., 2p\}, ..., \{pl-p+1, pl-p+2, ..., pl\}$, such that at least one element is chosen from each of the p elements sets. ♣

Lemma 5.0.10 *The forest $\mathcal{F}^{\mathrm{Hom}}_{p,r,n}$ of (p,n)-labelled trees, contains $\binom{n}{r}$ trees. Each branch α with $\mid \alpha \mid = l$ appears in $A_{p,l,r}$ trees.*

Proof: Clearly $\mid \mathcal{F}^{\mathrm{Hom}}_{p,r,n} \mid = \binom{n}{r}$. Suppose that
$\alpha = \{i^1_1, i^1_2, ..., i^1_p\}, \{i^2_1, ...i^2_p\}, ..., \{i^l_1, i^l_2, ..., i^l_p\}\}$ where $i^1_2 < i^2_1 < .... < i^l_1$ and where $i^j_1 < i^j_2 < ... < i^j_p$ for $j = 1, 2, ..., l$. Now there is a one to one correspondence between the r element subset of $\cup\alpha$, which contains at least one element from each member in α, and the trees in $\mathcal{F}^{\mathrm{Hom}}_{p,r,n}$ which contain α. □

Lemma 5.0.11 *Let q be a fixed prime number. Let s be any fixed number. Then for each $v \geq 1$, Count($q^v \cdot s$) $\vdash$ Count($q^{v+1} \cdot s$).*

Proof: Without loss of generality we can assume that q is not a prime factor in s. Consider the forest $\mathcal{F}$ which contains q^{v-1} copies of the forest $\mathcal{F}^{\mathrm{Hom}}_{p,r,n}$ where $p := q^{v+1} \cdot s$, $r := q^v$. The critical cases (the only non-trivial cases) are when $n = q^{v+1}n' + r \cdot q^v$, $r = 1, 2, ..., q-1$. The forest contains $k \cdot q^{v-1}$ modulo q^v trees (for $k \in \{1, 2, .., p-1\}$). The most critical case is branches of length 1. They appears $q^{v-1} \cdot \binom{q^{v+1} \cdot s}{q^v} = 0$ modulo q^v times. Longer branches also appears 0 modulo q^v times. □
This gives the positive part of the classification:

Corollary 5.0.12 *If $p, q \geq$ and all prime factors in p appears in q then, Count(q) $\vdash$ Count(p).*

5.1 Some examples

Before I show the negative part of the classification I will examine the structure of the exceptional forests when these are on the **PU**-form. Each (irreducible) exceptional forest $\mathcal{F}$ of **PU**-trees I have found can be derived from the forests given in definition 5.0.8.

Example 5.1.1 *Let $q = 2$, $p = 4$ and $n = 4n' + 2$, $n' \geq 2$. Consider the forest $\mathcal{F}_{2,4,n}$ of $(4,n)$-labelled* **PU***-trees which contains:*

All **PU***-trees of the form $(i_1, W_1)(i_2)$ where $i_1 < i_2 \leq n$ and where $W_1 \subseteq \{1, 2, ..., n\}$ has 3 elements.*

The **PU***-trees $(1), (3), (5), ..., (4n' + 1)$.*

Each branch of length 2 appears 16 times. A branch $\{j_1, j_2, j_3, j_4\}$ of length 1, appears in $mod(j_1, 2) + mod(j_2, 2) + mod(j_3, 2) + mod(j_4, 2)$ trees of hight 1. And the branch appears in $\binom{n-2}{3} - 4n + 6 + j_1 + j_2 + j_3 + j_4$ trees of hight 2.

The forests $\mathcal{F}_{2,4,n}$ contain $\binom{n}{2}\binom{n-2}{3} + \frac{n}{2}$ trees. This is always an odd number. When $n' = 2$ the forests contain 2525 trees.

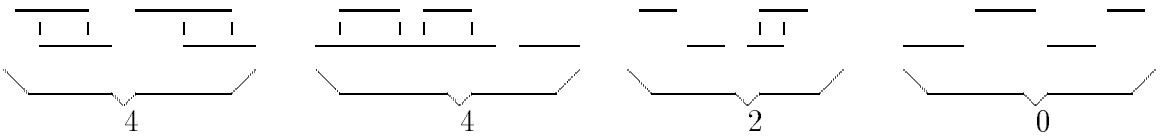
Example 5.1.2 *Let $n = 4n' + 2$. For each $n' \geq 2$ there exists a 2-exceptional forest of $(4,n)$-labelled trees. The forest contains all trees:*

$(j_1, \{j_2, j_3, j_4\})(j_5)$ where $j_1 < j_2 < j_3 < j_4$ and where $j_1 < j_5 < j_2$ or $j_3 < j_5 < j_4$.

$(1), (3), ..., (4n' + 1)$.

Each branch of length 2 appears an even number of times.

The branch $\{\{i_1, i_2, i_3, i_4\}, \{i_5, i_6, i_7, i_8\}\}$ appears a number of times depending on the number of vertical lines in the following kind of figure:



The branch $\{\{i_1, i_2, i_3, i_4\}\}$ appears in all trees (of hight 2) except $(i_2 - i_1) + (i_4 - i_3)$ which counted modulo 2 is $i_1 + i_2 + i_3 + i_4$. This is the same number (modulo 2) it appears in trees of hight 1.

When $n' = 2$ this is a forest of 635 trees. I conjecture that for $q = 2$ this is the smallest exceptional forest of **PU***-trees.*

The next example is derived from lemma 5.0.10

Example 5.1.3 *Let $q = 3$, $p = 9$ and $n = 9n' + 3, \quad n' \geq 3$. Consider the forests $\mathcal{F}_{3,9,n}$ which contain the $(9,n)$-labelled* **PU***-trees:*

(1) *The trees of the form $(i_1, W_1)(i_2, W_2)(i_3)$ where $i_1 < i_2 < i_3 \leq n$ and where W_1, W_2 are two disjoint 8-element subsets of $\{1, 2, ..., n\} \setminus \{i_1, i_2, i_3\}$.*

(2) *Two copies of each tree of the form $(i_1, W_1)(i_2)$ where $i_2 = 1$ modulo 3 and $i_1 < i_2 \leq n$.*

(3) *Each tree of the form $(i_1, W_1)(i_2)$ where $i_2 = 2$ modulo 3 and $i_1 < i_2$.*

(4) *Each tree of the form (i_1) where $\binom{n-i_1}{2} = 1$ modulo 3.*

(5) *Two copies of each tree of the form (i_1) where $\binom{n-i_1}{2} = 2$ modulo 3.*

A careful checking shows that each branch appears 0 modulo 3 times. The forests contains 1 modulo 3 trees. In the smallest case (i.e. when n'=3) the forest contains

$$\binom{30}{3}\binom{27}{8}\binom{19}{8} + \Sigma_i \Sigma j > i, j =_3 1\ 2 + \Sigma_i \Sigma_{j>i, j=_3 2}\ 1 + \Sigma_i \bmod(\binom{n-i}{2}, 3)$$

trees. This is a forest of 681259986982585 *trees. It is not the smallest exceptional forest for $q = 3$.*

There are smaller exceptional forests

Example 5.1.4 *Consider the forest $\mathcal{F}$ which contains:*

t copies of each tree $(i_1, \{i_2, i_3, ..., i_9\})(j_1, \{j_2, ...j_9\})(k)$ where $i_1 < i_2 < ... < i_9$, $i_1 < j_1$, $j_1 < j_2 < ... < j_9$ and if $i_r < k < i_{r+1}$ and $j_s < k < j_{s+1}$ then $r + s = t$ modulo 3.

t copies of each tree $(i_1, \{i_2, ...i_9\})(j)$ where $i_1 < i_2 < ... < i_9$ and if $i_r < j < i_{r+1}$ then $t = r$ modulo 3.

t copies of each tree (i) where $t = i$ modulo 3.

I claim (without proof) that each branch appears 0 modulo 3 times. However for each $n = 9n' + 3 \quad n' \geq 3 \mid \mathcal{F} \mid \neq 0$ modulo 3. More specifically $\mid \mathcal{F} \mid = \frac{1}{3} \cdot \binom{n}{9}\binom{n-9}{9} \cdot (n -$ 18). In the case when $n' = 3$ $\mathcal{F}$ only contains 16821302548060 **PU***-trees. I conjecture that this is the smallest exceptional forest for $q = 3$.*

From the examples we notice a general feature. The trees of maximal hight h are very homogeneously organised and easy to describe. The Trees of hight $h-1$ are still quite regular but each such tree's frequency $0 \leq v < q$ is slightly more complicated to describe. The collection of trees of hight 1 have the frequencies which are the most complicated to calculate.

In the next section I show that all exceptional forests asymptotically (when $n \to \infty$ and the hight of all trees is bound by a constant) can be assumed to have the same feature.

6 The negative part

The negative part of the classification states that Count(q) does not imply Count(p) when p contains a prime factor which is not in q. We consider the case where all terms in the underlying language L have (at most) polynomial growth-rate. By the first main result theorem 4.0.5 it suffice to show that for each $h \in \mathbb{N}$ asymptotically (when $n \to \infty$, and $h(\mathcal{F}) \leq h$) there are no q-exceptional forests $T_1, T_2, ..., T_u$ of (p,n)-labelled trees.

This is shown by considering forests $T'_1, T'_2, ..., T'_u$ of specially labelled trees corresponding to the Count(q) versus PHP_{q^k}

Definition 6.0.5 A (D,R)-labelled tree T is a decision tree for constructing a partial bijection $f : D \to R$. We always assume that $D \cap R = \emptyset$. Each vertex $v \in T$ corresponds to a certain stage f_v in the construction of f. At the root v_{root} we have $f_{v_{\mathrm{root}}} = \emptyset$.

At each vertex v (except the top node) there is a assigned a "question", i.e. an element $u \in D \cup R \setminus (\mathrm{dom}(f_v) \cup \mathrm{ran}(f_v))$. Each "answer" corresponds to the sons of v. If $u \in D$, there is an edge to a son, for each $r \in R \setminus \mathrm{ran}(f_v)$. Each of these edges lead to a vertex v' in which $f_{v'} \supseteq f_v$, and $f_{v'}(u) = r$ (and $\mid f_{v'} \mid = \mid f_v \mid +1$). Similarly if $u \in R$. In this case there is an edge for each $d \in D \setminus \mathrm{dom}(f_v)$. Each of these edges lead to a vertex v' in which $f_{v'} \supseteq f_v$, and $f_{v'}(d) = u$ (and $\mid f_{v'} \mid = \mid f_v \mid +1$).

The type of a branch through T is identified with the partial map f_v constructed at the leafs v. ♣

This labelling is more manageable than the (p,n)-labelling.

Definition 6.0.6 The hight $h(\mathcal{F})$ of the forest $\mathcal{F}$ denotes the maximal hight of a tree $T \in \mathcal{F}$. ♣

Definition 6.0.7 A (D,R)-labelled tree T is a **PU**-labelled tree (= on **PU**-form) if at each level all but possible one vertex is a top node. □

Observation 6.0.8 *A (D,R)-labelled tree T on* **PU***-form can be written of the form:*

$$(u_1^1, u_2^1)(u_1^2, u_2^2)...(u_1^{l-1}, u_2^{l-1})(u^l)$$

If $u_1^i \in D$ then $u_2^i \in R$ and if $u_1^i \in R$ then $u_2^i \in D$ ($i = 1, 2, ..., l-1$). The element u^l belongs to either D or R. In the first case we say T is of D-type, while we say that T is on R-type.

As an example consider the following (obvious) proposition:

Proposition 6.0.9 *Suppose that $\mathcal{F}$ solely consists of (D,R)-labelled trees of D-type. Suppose also (as usual) that $\mid D \mid \leq \mid R \mid$. Suppose that each branch α in $\mathcal{F}$ appears 0 modulo q times. Then the forest $\mathcal{F}$ contains 0 modulo q trees.*

Proof: Let $\rho : D \rightarrow R$ be an (total) injection. Each tree $T \in \mathcal{F}$ contains exactly one branch α with $\alpha \subseteq \rho$. Thus ρ induces a partitioning of the trees in $\mathcal{F}$ into disjoint classes which each contains 0 modulo q trees. □

If the forest $\mathcal{F}$ contains trees of both D-type and R-type the situation becomes more complicated.

Example 6.0.10 *Let $q \in \mathbb{N}$. Consider the following forests $\mathcal{F}$ of (D,R)-labelled trees. For each $d \in D$ it contain $(q-1)$ copies of (d), and for each $r \in R$ it contain the tree (r).*

This forests $\mathcal{F}$ contains $\mid R \mid - \mid D \mid$ modulo q trees. Each type of branch appears 0 modulo q times. So trivially if $\mid R \mid \neq \mid D \mid$ modulo q, there exists a forest $\mathcal{F}$ in which all branches appears 0 modulo q times, but $\mid \mathcal{F} \mid \neq 0$ modulo q.

This type of forest is trivial. It corresponds to the obvious fact that Count(q) implies PHP_p when $p \neq 0$ modulo q. This type of forests are so simple that we will not consider them as exceptional.

Definition 6.0.11 A forest $\mathcal{F}$ of (D,R)-labelled trees is called (q,l)-exceptional if

(i) Each type branch appears 0 modulo q times.

(ii) $\mid R \mid = \mid D \mid + q^l$.

(iii) The number of trees in $\mathcal{F}'$ is not divisible by q. ♣

Example 6.0.12 *Suppose $\mid R \mid - \mid D \mid = 4p' + 2$ for some $p' \in \mathbb{N}$. Assume that $\mid R \mid$ is an odd number. Let $\mathcal{F}$ denote the (D,R)-labelled forest which contains the following* **PU***-trees:*

(1) *All trees of the form $(d,r_1)(r_2)$ where $d \in D$ and $r_1 > r_2$ when $\mid r_1 - r_2 \mid$ is odd, and $r_1 < r_2$ when $\mid r_1 - r_2 \mid$ is even.*

(2) *All trees of the form $(d_1,r)(d_2)$ where $r \in R$ and $d_1 < d_2$.*

Each branch appears an even number in $\mathcal{F}$. However, the forest $\mathcal{F}$ contains

$$\mid D \mid \binom{\mid R \mid}{2} + \mid R \mid \binom{\mid D \mid}{2}$$

trees which is always an odd number. The smallest example of this form is when $\mid D \mid = 5$ and $\mid R \mid = 7$. In this case $\mathcal{F}$ contains 175 trees. I claim without proof that this is the smallest $(2,1)$-exceptional forest. The forest resemble the fact that Count(2) implies PHP_2.

There are (q,l)-exceptional forests for each $q,l \in \mathbb{N},\ q \neq 1$ (we do not need this fact). This follows by combining:

(1) We have a version of theorem 4.0.5 for the Count(q) versus PHP_{q^l} problem.

(2) Trivially PHP_1 follows from Count(q)

(3) By [24] PHP_{q^l} follows from PHP_1.

It is interesting to notice that this proof is non-constructive. It only shows that the exceptional forests exists. It does not shows how to constructs them. It turns out that they can be constructed along the same lines as the constructions in section 4.

6.1 Projecting forests

Let T be a (p,n)-labelled tree. Suppose $n = pn' + q^l$ for $q \in \mathbb{N} \setminus \{1\}$ and $l \in \mathbb{N}$. Then we can transform it to a (D,R)-labelled tree by the following procedure.

First divide $I := \{1,2,...,n\}$ into p disjoint sets $D_1, D_2, ..., D_{p-1}$ and R such that $\mid D_1 \mid = \mid D_2 \mid = ... = \mid D_{p-1} \mid = n'$ and $\mid R \mid = n' + q^l$. Let $D := D_1$. For $j = 1,2,...,p-1$ chose bijections $y_j : D_j \to D$. Let us call a subset $\{i_1, i_2, .., i_p\} \subseteq I$ for *regular* if $i_p \in R,\ i_j \in D_j,\ j = 1,2,..,p-1$ and $y_1(i_1) = y_2(i_2) = ... = y_{p-1}(i_{p-1})$. A branch $\{A_1, .., A_r\}$ is *regular* if each $A_j,\ j = 1,2,..,r$ is regular. By use of this definition it is straight forward to show that,

Lemma 6.1.1 *Let T be a (p,n)-labelled tree. Suppose that $n = pn' + q^l$ and let $D_1, .., D_{p-1}$ and R be given as above. Then the set of regular branches in T form a new tree T' which is (D,R)-labelled. Furthermore $h(T') \leq h(T)$.*

If T is on **PU***-form, then T' will also be on* **PU***-form.*

Instead of projecting a single tree we can project forests. The important point is that the projection of an q-exceptional forest of (p,n)-labelled trees produces an (q,l)-exceptional forest of (D,R)-labelled trees.

Lemma 6.1.2 *Let $\mathcal{F} := \{T_1, T_2, .., T_u\}$ be a forest of (p,n)-labelled trees. Suppose $\mathcal{F}$ is an q-exceptional forest. Or more specifically that p does not divide q and each branch in $\mathcal{F}$ appears 0 modulo q times, but $u \neq 0$ modulo q. Suppose also that $n = pn' + q^l$. The projection of the trees $T_1, ..., T_u$ gives an (q,l)-exceptional (D,R)-labelled forest (with $\mid D \mid = n'$ and $\mid R \mid = n' + q^l$). Furthermore $h(\mathcal{F}') \leq h(\mathcal{F})$.*

Proof: This is left to the reader to verify. □

The condition that $n = pn' + q^l$ might not in general be satisfied for a given n. However, usually we do not lose any generality by assuming n is of this form. To see this consider the following procedure:

Definition 6.1.3 Suppose that ρ is a partial partition of $I := \{1, 2, ..., n\}$ into disjoint p-element subsets. Consider a (p, n)-labelled tree T. For each branch (=condition) α through T consider the following procedure. If α is incompatible with ρ remove it. Otherwise replace it by $\beta := \alpha \setminus \rho$.

Suppose that ρ' is a partial bijection from D to R. Consider a (D, R)-labelled tree T. For each branch (=condition) α' through T consider the following procedure. If α' is incompatible with ρ' remove it. Otherwise replace it by $\beta' := \alpha' \setminus \rho'$. ♣

Lemma 6.1.4 (Stability) *Suppose that T is a (p, n)-labelled tree. Let ρ be a partial partition of $I = \{1, 2, ..., n\}$ into disjoint p-element subsets. Suppose that $p \cdot (h(T)+ \mid \rho \mid) < n$. Then the collection of all branches β which are produced from some $\alpha \in T$ (as described in definition 6.1.3) can be organised into a (p, n')-labelled tree T^ρ where $n' = n - p \mid \rho \mid$.*

Suppose that T' is a (D, R)-labelled tree. Let ρ' be a partial bijection from D to R. Suppose that $h(T)+ \mid \rho' \mid < n$. Then the collection of all branches β' which are produced from some $\alpha' \in T$ (as described in the second part of definition 6.1.3) can be organised into a (D', R')-labelled tree $T'^{\rho'}$ where $D' := D \setminus \mathrm{dom}(\rho')$ and $R' := R \setminus \mathrm{ran}(\rho')$.

If T (T') is a **PU***-tree then T^ρ ($T'^{\rho'}$) is a* **PU***-tree.*

Proof: It suffice to show the lemma when $\mid \rho \mid = 1$. Suppose that $\rho := \{\{i_1, i_2, ..., i_p\}\}$. Let V be the set of vertex in T which have assigned $v \in \{i_1, .., i_p\}$. Let $E_\perp$ denote the set of edges in T which has assigned p-subset $A \subseteq I$ with non-trivial intersection with $\{i_1, .., i_p\}$ (i.e. $\neq \emptyset$ and $\neq A$). Let $E_{||}$ be the set of edges which have assigned $A = \{i_1, i_2, .., i_p\}$.

For each vertex (=question) in V all edges (but exactly one) edge (=answer) belongs to $E_\perp$. Remove all these edges (and the sub-tree above this). Then contract the edge in $E_{||}$. Finally, after having exhausted this procedure, remove all edges in $E_\perp$ (and the sub-tree above this). The condition that $p(h(T)+ \mid \rho \mid) < n$ is exactly what in general is required to ensure that T^ρ actually becomes a properly labelled tree. The second part of the lemma is showed similarly. The last claim is also straight forward to check. □

The lemma is one of many stability results which are important for the overall argument. In short it shows that trees (**PU**-trees) remains on this form when they are "hit" by a restriction ρ.

The main lemma gives us an understanding of the asymptotic behaviour of exceptional (D, R)-labelled trees.

Lemma 6.1.5 (Main lemma) *Let q be a prime number. Let $k, l \in \mathbb{N}$. There exists $d_0 \in \mathbb{N}$ such that for any (q^k, l)-exceptional forest of (D, R)-labelled trees, where $d_0 \leq \mid D \mid \leq \mid R \mid$, we have $h(\mathcal{F}') \geq q^{l-k}$.*

Corollary 6.1.6 *Let p be any prime number which does not appear in $q = q_1^{\alpha_1} q_2^{\alpha_2} ... q_r^{\alpha_r}$. Fix $h \in \mathbb{N}$. There exists n_0 such that for each $n \geq n_0$ each forest $\mathcal{F} := \{T_1, T_2, ..., T_u\}$ of (p, n)-labelled trees never simultaneously satisfies:*

(1) *All branches appears 0 modulo q times.*

(2) *$h(\mathcal{F}) \leq h$.*

(3) *$u \neq 0$ modulo q.*

Proof: Suppose that (1) and (3) hold. Choose $j \in \{1, 2, .., r\}$ such that $u \neq 0$ modulo $q_j^{\alpha_j}$. According to the assumptions $\mathcal{F}$ is an $q_j^{\alpha_j}$-exceptional forest of (p, n)-labelled trees. By lemma 6.1.4 we can assume that $n = pn' + q_j^l$ for any l given in advance (of course l has to be reasonable i.e. $q^l << n$ etc). Choose l such that $q_j^{l-\alpha_j} > h$. By lemma 6.1.2 the projected forest $\mathcal{F}'$ is $(q_j^{\alpha_j}, l)$-exceptional and $h(\mathcal{F}') \leq h(\mathcal{F})$. According to lemma 6.1.5 $h(\mathcal{F}') \geq q^{l-\alpha_j} > h$. Now $h(\mathcal{F}') \leq h(\mathcal{F})$ so this contradicts (2). □

6.2 Creating order among trees of maximal hight h

Lemma 6.2.1 *Fix $q \in \mathbb{N} \setminus \{1\}$, and fix $l, h \in \mathbb{N}$. For each $d_0 \in \mathbb{N}$ with $d_0 \geq h$, there exists (a very large) $d_1 \in \mathbb{N}$ such that for each forest $\mathcal{F} = \{T_1, T_2, ..., T_u\}$ of (D, R)-labelled trees with $\mid R \mid = \mid D \mid + q^l$ and $\mid D \mid \geq d_1$ the following is true:*

There exists a partial bijection $\rho : D \to R$, such that the forest $\mathcal{F}^\rho := \{T_1^\rho, T_2^\rho, .., T_u^\rho\}$ of (D', R')-labelled trees, with $D' = D \setminus \mathrm{dom}(\rho)$ and $R' = R \setminus \mathrm{ran}(\rho)$, satisfies:

(1) *For each $h-1$ element subset $\{d_1, d_2, .., d_{h-1}\} \subseteq D'$ with $d_1 < d_2 < ... < d_{h-1}$ and for each permutation $\pi : \{1, 2, .., h\} \to \{1, 2, .., h\}$ the number (modulo q) of trees (in the forest $\mathcal{F}^\rho$) of the form*

$$(d_1, r_{\pi(1)})(d_2, r_{\pi(2)})....(d_{h-1}, r_{\pi(h-1)})(r_{\pi(h)})$$

does <u>not</u> depend on the elements $r_1 < r_2 < ... < r_h$ in R'.

(2) *For each $(h-1)$-element subset $\{d_1, d_2, ..., d_{h-1}\} \subseteq D'$ (where $d_1 < d_2 < ... < d_{h-1}$), for each element $d_h \in D' \setminus \{d_1, d_2, .., d_{h-1}\}$, and for each permutation $\pi : \{1, 2, ..., h-1\} \to \{1, 2, ..., h-1\}$ the number (modulo q) of trees of the form*

$$(d_1, r_{\pi(1)})(d_2, r_{\pi(2)})....(d_{h-1}, r_{\pi(h-1)})(d_h)$$

does <u>not</u> depend on the elements $r_1 < r_2 < ... < r_{h-1}$ in R'.

(3) *$\mid D' \mid \geq d_0$.*

Proof: Let $D_1 \subseteq D$ be a subset with $\mid D_1 \mid \geq d_0'$ for some number much larger than d_0. The size of d_0' can be expressed in terms of the estimates arising from the second part of the argument (where we ensure the validity of property (3)). For each $h-1$ element subset $\{d_1, d_2, ..., d_{h-1}\} \subseteq D_1$ with $d_1 < d_2 < ... < d_{h-1}$ and for each permutation $\pi : \{1, 2, .., h\} \to \{1, 2, ..., h\}$ we define a map $F(d_1, d_2, ..., d_{h-1}; \pi)$ which maps h-element subsets of R to the set $\{0, 1, 2, .., q-1\}$. It is defined by letting $F(d_1, d_2, ..., d_{h-1}; \pi)(\{r_1, r_2, ..., r_h\})$ (where $r_1 < r_2 < ... < r_h$) denote the number (modulo q) of the **PU**-trees

$$(*) \qquad (d_1, r_{\pi(1)})(d_2, r_{\pi(2)})....(d_{h-1}, r_{\pi(h-1)})(r_{\pi(h)})$$

These maps induce as a map $\tilde{F}$ which to each h element subset $\{r_1, r_2, ..., r_h\} \subseteq R$ takes one of $q^{\binom{|d_0|}{h-1}h!}$-values. This value expresses uniquely for each $h-1$ element subset of D_1 and each permutation $\pi : \{1, 2, .., h\} \to \{1, 2, ..., h\}$, the number (modulo q) of **PU**-trees of the form $(*)$.

Now by Ramseys theorem [7] if d_1 is sufficiently large (not depending on $\mathcal{F}$) there must be a set $R_1 \subseteq R$ which is homogeneous for the "collaring" $\tilde{F}$. By possible making R_1 slightly smaller we can ensure that $\mid D \setminus D_1 \mid = \mid R \setminus R_1 \mid$. Choose a partial bijection $\rho : D \to R$ such that $\mathrm{dom}(\rho) = D \setminus D_1$ and $\mathrm{ran}(\rho) = R \setminus R_1$. This ensure that the new restricted forest satisfies property (2).

This procedure is now repeated (with d_0' replaced by d_0) such that property (3) are satisfied. Notice that application of a new ρ' does not destroy property (2). □

Definition 6.2.2 Two tuples $< r_1, r_2, .., r_h >$ and $< r_1', r_2', ..., r_h' >$ have the same *order structure* if for the same permutation π we have that $r_{\pi(1)} < r_{\pi(2)} < ... < r_{\pi(h)}$ and $r'_{\pi(1)} < r'_{\pi(2)} < ... < r'_{\pi(h)}$. ♣

Lemma 6.2.3 (Stability) *Suppose that $\mathcal{F} := \{T_1, T_2, ..., T_u\}$ is a forest of (D, R)-labelled* **PU***-trees. Suppose all trees of maximal hight h satisfies (1) and (2) in lemma 6.2.1. For any partial bijection $\rho : D \to R$, with $h+ \mid \rho \mid < \mid D \mid$, the forest $\mathcal{F}' := \{T_1^\rho, T_2^\rho, ..., T_u^\rho\}$ of (D', R')-labelled trees ($D' := D \setminus \mathrm{dom}(\rho), R' := R \setminus \mathrm{ran}(\rho)$) also satisfies (1) and (2).*

Proof: By the second part of lemma 6.1.4, we only have to check that (1) and (2) in lemma 6.2.1 will be satisfied. To show (1) we have to prove that the **PU**-trees

$$(d_1, r_1)(d_2, r_2)...(d_{h-1}, r_{h-1})(r_h)$$

[7] The application of Ramseys theorem seems to play a similar role in [6].

and

$$(d'_1, r'_1)(d'_2, r'_2)...(d'_{h-1}, r'_{h-1})(r'_h)$$

appears the same number of times (modulo q) when $d_1 < d_2 < ... < d_{h-1}$, when $d'_1 < d'_2 < ... < d'_{h-1}$, and when the order type of $< r_1, r_2, .., r_h >$ and $< r'_1, r'_2, ..., r'_h >$ are the same. This follows from the fact that none of the representations can have been altered by ρ. □

6.3 Creating order among trees of hight $h' < h$

Let $\mathcal{F}$ be a forest of (D, R)-labelled trees of **PU**-form. Suppose that all trees satisfies condition (1) and (2) in lemma 6.2.1. Wright $\mathcal{F}$ as the union $\mathcal{F}_1 \cup \mathcal{F}_2 \cup ... \cup \mathcal{F}_h$, where the sub-forest $\mathcal{F}_{h'}$ contains all trees of hight h'. By the same argument as in lemma 6.2.1 there exists (provided D and R are sufficiently large compared to $h(\mathcal{F})$) a restriction ρ (i.e. a partial bijection) such that for each $h' \in \{1, 2, .., h\}$ all trees in $\mathcal{F}^{\rho}_{h'}$ satisfies (1) and (2) in lemma 6.2.1 with h replaced by h'.

Consider the trees in $\mathcal{F}^{\rho}$ of some hight $h' < h$. Clearly there is a flux of trees from each $\mathcal{F}_{h''}$ with $h' < h'' \leq h$.

Definition 6.3.1 (Strong normal form) A forest $\mathcal{F} := \{T_1, T_2, ..., T_u\}$ is on *strong normal form* if for each $h' \leq h$:

(1) For each $d_1 < d_2 < ... < d_{h'-1}$ and for each permutation $\pi : \{1, 2, ..., h\} \rightarrow \{1, 2, .., h\}$ the number (modulo q) of trees of the form

$$(d_1, r_{\pi(1)})(d_2, r_{\pi(2)})...(d_{h-1}, r_{\pi(h-1)})(r_{\pi(h)})$$

only depends on residue classes modulo $q^{h-h'}$ of the elements $r_1 < r_2 < ... < r_h$.

(2) For each $d_1 < d_2 < ... < d_{h'-1}$, for each $d_{h'} \in D \setminus \{d_1, d_2, .., d_{h'-1}\}$ and for each permutation $\pi : \{1, 2, ..., h'-1\} \rightarrow \{1, 2, .., h'-1\}$ the number (modulo q) of trees of the form

$$(d_1, r_{\pi(1)})(d_2, r_{\pi(2)})...(d_{h'-1}, r_{\pi(h'-1)})(d_{h'})$$

only depends on residue classes modulo $q^{h-h'}$ of the elements $r_1 < r_2 < ... < r_{h-1}$. ♣

Lemma 6.3.2 (Stability) *If $\mathcal{F}$ is on strong normal form, then $\mathcal{F}^{\rho}$ is on strong normal from (provided that $h(\mathcal{F})+ \mid \rho \mid < \mid D \mid \leq \mid R \mid$).*

Proof: There are 3 ways the representation of a **PU**-tree T

$$(d_1, r_1)(d_2, r_2)....(d_{l-1}, r_{l-1})(d_l, h_l)...(d_{h-1}, r_{h-1})(r_h)$$

might change. In all cases suppose that the lowest place where it get 'hit' is on level l.

(1) $< d_l, r_l > \in \rho$. The tree T^ρ is of the form $(d_1, r_1)...(d_{l-1}, r_{l-1})(d_{l+1}, r_{l+1})...(r_h)$. If this tree get 'hit' by ρ again there are 3 ways this can happen...(repeat the reduction).

(2) $< d_l, r' > \in \rho$ where $r' \neq r_l$. The tree T^ρ has the representation $(d_1, r_1)(d_2, r_2)...(d_{l-1})$.

(3) $< d', r_l > \in \rho$ where $d' \neq d_l$. The tree T^ρ has the representation $(d_1, r_1)(d_2, r_2)....(d_{l-1}, r_{l-1})(d_l)$.

From this observation it is not hard to see that the regularity among the trees of hight l, are inherited (after the restriction) by trees of smaller hight. □

If we combine this lemma with lemma 6.1.4, 6.2.3 and lemma 6.2.1 we get:

Lemma 6.3.3 (Strong normal form) *Fix $q \in \mathbb{N} \setminus \{1\}$. For all $h, l \in \mathbb{N}$ there exists $d \in \mathbb{N}$ such that the following hold:*

Suppose that $\mathcal{F} := \{T_1, T_2, ..., T_u\}$ is a (D, R)-labelled forest where all trees have hight $\leq h$ and where $d \leq \mid D \mid$ and where $\mid R \mid = \mid D \mid + q^l$. Then there exists a partial bijection $\rho : D \to R$ such that $\mathcal{F}^\rho := \{T_1^\rho, T_2^\rho, ..., T_u^\rho\}$ is a forest of (D', R')-labelled trees (where $D' := D \setminus \mathrm{dom}(\rho)$ and $R' := R \setminus \mathrm{ran}(\rho)$) on the strong normal form.

Furthermore if $\mathcal{F}$ is (q, l)-exceptional, then $\mathcal{F}^\rho$ is (q, l)-exceptional.

6.4 Proof of the main lemma

Now we are ready to show the main lemma (lemma 6.1.5) in this section.

Proof: According to lemma 6.3.3 we lose no generality by assuming that $\mathcal{F}$ is on strong normal form. For each $d \in D$ and $r \in R$ we can consider the branch $\{< d, r >\}$ of length 1. It appears 0 modulo q^k times so we have the identity:

(1) $(d, *) \rightsquigarrow -(d, r) \rightsquigarrow +(r) = 0$ modulo q^k.

Here $(d, *) \rightsquigarrow$ of course denotes the number of **PU**-trees of the form

$$(d, r_1)(u_1^2, u_2^2)....(u_1^{l-1}, u_2^{l-1})(u^l)$$

Similarly $(d, r) \rightsquigarrow$ denotes the number of **PU**-trees of the form

$$(d, r)(u_1^2, u_2^2)....(u_1^{l-1}, u_2^{l-1})(u^l),$$

and (r) denote the number of appearances of the tree (r). The trees in (d) does not enter the equation because we assume $\mid R \mid = 0$ modulo q^k.

The number u of trees in the forest $\mathcal{F}$ is given by:

(2) $u = \Sigma_{d\in D}\ (d,*) \rightsquigarrow\ +\Sigma_{d\in D}\ (d)\ + \Sigma_{r\in R}\ (r)$

Now for each $d \in D$ consider $d_2, d_3, .., d_{l-1}, d_l$. Consider the set of trees of the form $(d,r_1)(d_2,r_2)...(d_{l-1},r_{l-1})(r_l)$ or $(d,r_1)(d_2,r_2)...(d_{l-1},r_{l-1})(d_l)$ where $r_1, r_2, .., r_l$ belongs to a certain type (expressed by the relative size of $r_1, .., r_l$, but also taking their residue classes modulo q^k into account). The number of such trees is 0 modulo q^{l-k} provided $\mid R \mid = \mid D \mid + q^l$. But then:

(3) $(d,*) \rightsquigarrow = 0$ modulo q^k for each $d \in D$.

But according to (1)

(4) $(d,r) \rightsquigarrow = (r)$ modulo q^k for all $d \in D$ and $r \in R$.

According to (2) and the assumption that $u \neq 0$ modulo q^k, for each $d \in D$

(5) $\Sigma_{d'\in D}\ (d')\ +\ \Sigma_{r\in R}(d,r) \rightsquigarrow \neq 0$ modulo q^k.

But in general

(6) $\Sigma_{r\in R}\ (d,r) \rightsquigarrow = (d,*) \rightsquigarrow$

so by combining (3) and (5)

(7) $\Sigma_{d\in D}\ (d) \neq 0$ modulo q^k.

But by the normal form theorem we can assume that D is divided into disjoint classes $D_1, D_2, ..., D_r$ which each have 0 modulo q^{l-h} elements. And thus if $l-h \geq k$ the sum $\Sigma_{d\in D}\ (d) = \Sigma_{j=1}^{r}\Sigma_{d\in D_j}\ (d)$. By the normal form theorem (d) the number of trees in (d) is constant on each $D_j,\ j = 1,2,..,r$ so the must equal 0 modulo q^{l-h}. This a is contradiction if $l-h \geq k$. □

6.5 Brief discussion of the general problem

The method in the last subsection only give an asymptotic classification of exceptional forest. This is good enough for a complete classification of the Count(q) versus Count(p) problem in the case of polynomial growth rate.

The fact that forests on the strong normal form remains on this form when 'hit' by a (randomly chosen?) restriction is very important. And it is very promising for the full classification (when n is large). The critical question is whether we can create order fast enough? Is it possible to create sufficiently much regularity before we have used the elements in $I = \{1,2,...,n\}$. This seems to be a race between different forces!

In the first version of this paper I tried to bring a hypothetically given exceptional forests $\mathcal{F}$ on a strong normal form. This was done by selecting a suitable collection G of group actions on $\mathcal{F}$. For each $g \in G$ I defined a forest $\mathcal{F}^g$ containing the same number of trees as $\mathcal{F}$. Now by a suitable choice of G (so $u\cdot \mid G \mid \neq 0$ modulo q) the forest $\cup_{g\in G}\ F^g$ remains exceptional. By a proper choice of G I was able to show

that the resulting forest gets efficiently closer to the strong normal form. This idea does not *a priory* require any strong assumptions on the hight of the forest. However the argument depend on the validity of a certain modular identity. At present this validity is open. Its validity could be important for the full classification in the general case of sub-exponential growth rate.

7 Some applications

There are various alternative formulations of the classification. It is well known that complexity theory can be viewed as recursion theory done within a finite set of unspecified size. The levels in Arithmetical Hierarchy correspond to the levels in Polynomial Hierarchy [8]. It can be argued that low complexity reasoning is reasoning which can be formalised within (arbitrarily large) finite structures. Suppose that the universe is such an unspecified finite set. Although this is almost impossible to picture it is consistent. Such an "axiomatic finite" universe can be axiomatised in various ways. Its models (which are highly non-recursive) are of course not really finite.

As an example consider the following axiomatisation over second order logic. Suppose that we have the full Arithmetical comprehension axiom schema,

$$\forall z \; \exists \; \mathbf{X} \; \psi(x, z) \leftrightarrow x \in \mathbf{X}.$$

Here ψ is any first order formula. We allow ψ to contain set-variables. And assume that we have the usual induction axiom

$$0 \in \mathbf{X} \wedge \forall n \; (n \in \mathbf{X} \rightarrow n + 1 \in \mathbf{X}) \rightarrow \forall n \; n \in \mathbf{X}.$$

If the underlying universe was not assumed to be finite this would be the celebrated and powerful system **ACA** of analysis. If the underlying universe is axiomatic finite (e.g. satisfies the pigeon-hole principle) we denote the axiom system by $\mathbf{ACA}^{\mathrm{top}}$. For this system

Theorem 7.0.1 Count*(p) holds in all structures of* $\mathbf{ACA}^{\mathrm{top}}$ + Count*(q) exactly when all prime divisors in p appear in q.*

Proof: Combine the conservation results in [23],[24] with results for Bounded Arithmetic. By these results the system has the same deductive strength as Bounded Arithmetic axiomatised without functions symbols. By use of the usual coding methods the system is able to handle terms of polynomial growth rate. Thus the positive part of the classification can obtained. □

It is also possible to link the result to length of proofs in propositional logic. This type of link was first pointed out in [19].

Definition 7.0.2 A *Boolean formula* is a Boolean circuit where for each disjunction $\vee_j \pi_j$ and for each conjunction $\wedge_j \pi_j$ a particular bracketing is specified. The size and the depth of a Boolean formula is defined in the obvious way. In the calculation of the depth, disjunctions $\vee_j \pi_j$ and conjunctions $\wedge_j \pi_j$ are chosen maximally. ♣

Definition 7.0.3 A *general propositional proof system* $\mathbb{P}$ consists of:

(1) A finite number of substitution schemes.

A *substitution scheme* is a Boolean formula θ which only contains *special variables* (substitution variables). A *substitution instance* of θ is obtained by substituting the substitution variables $y_1, ..., y_k$ by Boolean formulas $\eta_1, ..., \eta_k$.

(2) A finite number of deduction rules.

A *deduction rule* $\frac{\theta_1, \theta_2, ..., \theta_k}{\theta}$ where $\theta_1, ..., \theta_k$ and θ are substitution schemes. A substitution instance is obtained by substituting the substitution variables $y_1, ..., y_k$ by Boolean formulas $\eta_1, ..., \eta_k$.

A $\mathbb{P}$-*proof* (in Hilbert style) of η is a sequence $\eta_1, ..., \eta_u = \eta$ of Boolean formulas, such that each η_j, $j = 1, 2, ..., u$ is either a substitution instance of a substitution scheme, or there are $i_1, ..., i_k < j$ such that $\frac{\eta_{i_1}, ..., \eta_{i_k}}{\eta_j}$ is a substitution instance of a deduction rule.

We only consider general propositional proof systems which are consistent and prove the usual tautologies.

The *size* s of a propositional proof is $s := \Sigma_j s(\eta_j)$, and the *depth* d is $d := \max_j d(\eta_j)$. ♣

Definition 7.0.4 A *Frege proof system* (or a textbook proof system) is a general propositional proof system, where modus ponens $\frac{y_1, \neg y_1 \vee y_2}{y_2}$ is the only deduction rule. ♣

Definition 7.0.5 Let $\mathrm{Count}_n(p)$ denote the tautology:

$$(\vee_{i \leq n} \vee_{\{A: i \in A\}} \vee_{\{B: i \in B \wedge A \neq B\}} (p_A \wedge p_B)) \vee (\vee_{i \leq n} \wedge_{\{i \in A\}} \neg p_A)$$

where the sets A and B run through the p subsets of $\{1, 2, ..., n\}$. ♣

Theorem 7.0.6 *Fix $p \in \mathbb{N}$. Let $\mathcal{A}_p$ be the collection of all substitution schemes of the* Count*(q) principle for $q \in \mathbb{N}$ which contain all prime factors of p. Let $\mathbb{P}$ be any general proof system to which all the schemes in $\mathcal{A}_p$ are added. Then the tautologies* $\mathrm{Count}_n(p)$*, do not have bounded depth polynomial size $\mathbb{P}$-proofs.*

Proof: Suppose that for arbitrarily large $n \in \mathbb{N}$, there exists a $\mathbb{P}$-proof of depth $\leq d$ and size $\leq \exp(n^{\epsilon(n)})$. Let R be a suitable relation with domain $\mathbb{N}^r$, $r \in \omega$, which codes these proofs. Let $\mathbb{M}$ be a countable non-standard model of $\mathrm{Th}(\mathbb{N})$ over some countable language L which extends the language of Arithmetic and contains R. By overspill there exists a non-standard number $n \in \mathbb{M}$ which is not divisible by p, and there exists an $\mathbb{M}$-definable sequence $\theta_1, \theta_2, ..., \theta_u$ of formulas, which (within $\mathbb{M}$) is a general propositional $\mathbb{P}$-proof of $\mathrm{Count}_n(p)$. Furthermore, we can assume that the depth of the proof is $\leq d$, and that the size of the proof is $\leq \exp(n^t)$ for some $t < \frac{1}{\omega}$ (the map $\epsilon : \mathbb{N} \to \mathbb{Q}_+$ can without loss of generality be assumed to be L-definable, because otherwise L can be extended with a relation which defines ϵ).

Now choose a generic truth-table evaluation $\tilde{\rho}_G$. Such an evaluation exists according to lemma 2.3.2. Consider the sequence $\theta_1, ..., \theta_u$ (considered as circuits) and notice that $(\theta_u)^{\tilde{\rho}_G} = (\mathrm{Count}_n(p))^{\tilde{\rho}} = 0$. According to corollary 2.4.2 there exists $j_0 \leq u$ such that $(\theta_{j_0})^{\tilde{\rho}_G} = 0$ but $(\theta_j)^{\tilde{\rho}_G} = 1$ for all $j < j_0$. Now each substitution instance θ_j of a substitution scheme has $(\theta_j)^{\rho} = 1$ for each general truth-table evaluation ρ. If θ_j is obtained from a deduction rule then $(\theta_j)^{\rho} = 1$ provided that all the premises also have truth-value 1.

Finally I claim that all substitution instances of the $\mathrm{Count}(q)$ principle also get truth-value 1. Now if it got the truth value 0, then by the work in section 3 there would be a $\mathbb{M}$-definable generic system. By our refinement technique this would imply the existence of a specially labelled (I, p)-forest in which all branches appear 0 modulo q times. And the forest would contain a number of trees not divisible by q. According to the combinatorial results in section 6 this (first order) statement fails in the standard universe. We chose $\mathbb{M}$ to be a model of first order arithmetic, so this is a contradiction. □

Theorem 7.0.7 *Let $\mathbb{M}$ be a countable non-standard model of Th($\mathbb{N}$) over a countable first order language L (which contains the language of arithmetic). Suppose that $p \in \mathbb{N}$, $p \geq 2$ and $I := \{1, 2, ..., n\} \subseteq \mathbb{M}$ (for some $n \in \mathbb{M} \setminus \omega$ <u>not</u> divisible by p). Let*

$$\mathbb{M}_n^* := \{m \in \mathbb{M} : \ t(n) > m \text{ for some term } t \in L\}.$$

For any generic filter ρ_G the partition $\tilde{\rho}_G$ (see definition 2.2.5 page 12) partitions I into disjoint classes, each containing exactly p elements. If the terms $t \in L$ all have polynomial growth rate

(a) $(M_n^*, \rho_G) \models \neg\mathrm{Count}(p)$.

(b) (M_n^*, ρ_G) *satisfies induction for bounded $L(P)$-formulas.*

(c) $(M_n^*, \rho_G) \models \mathrm{Count}(q)$ *for all q which contains all prime factors in p.*

Proof: It suffices to show that the least number principle is valid for bounded $L(P)$-formulas with parameters in $\mathbb{M}_n^*$. Now each instance of the least number principle gets translated into a Boolean circuit (or Boolean formula if we specify the bracketing) of the form $\mathrm{LNP}_n(\pi_1, ..., \pi_u) := \pi_u \vee (\vee_{j \leq u}(\neg \pi_j \wedge (\wedge_{k<j} \pi_k)))$. Furthermore, according to earlier observation, each translated instance gets depth ≤ 4 and size $\leq \exp(n^t)$ for some $t < n^{\frac{1}{\omega}}$. According to the key lemma (lemma 2.4.1) for <u>any</u> generic filter ρ_G if $(\pi_u)^{\tilde{\rho}_G} = 1$, there exists $j_0 \leq u$ with $(\pi_{j_0})^{\rho} = 1$ and $(\pi_j)^{\rho} = 0$ for $j < j_0$. A simple argument shows that $\mathrm{LPN}_n(\pi_1, ..., \pi_u)^{\tilde{\rho}_G} = 1$. Using lemma 2.1.6 $(\mathbb{M}, \tilde{\rho}_G)$ satisfies induction for bounded $L(P)$-formulas with parameters in $\mathbb{M}_n^*$.

Again all $\mathrm{Count}(q)$ much be forced true when p contains a prime factor not in q. If not there would exists an $\mathbb{M}$-definable generic system. And thus by the refinement argument there would be a (I, p) forest with $\neq 0$ modulo q trees, in which each (type of) branch appears 0 modulo q times. This is a contradiction when p does not divide q. □

Theorem 7.0.8 *Suppose that all terms in L have polynomial growth rate, and contains at least one unspecified relation symbol. Then $I\Delta_0(L)+ \mathrm{Count}(q)$ prove $\mathrm{Count}(p)$ exactly when all prime factors in p divides q.*

Proof: $(\mathbb{M}_n^*, \tilde{\rho}_G) \models I\Delta_0(L) + \neg\ \mathrm{Count}(p) + \mathrm{Count}(q)$. □

8 Final remarks

The first version of this paper contained the complete reduction of the $\mathrm{Count}(q)$ versus $\mathrm{Count}(p)$ problem. This reduced the problem to a purely combinatorial problem. The revised version solves this problem explicitly (in the case of polynomial growth-rate). In addition the revised paper develops the underlying theory in more details.

To end, I am happy to learn that the topic is related to Hilberts Nullstellensatz [6], and by [4] also to representations of symmetrical groups. I hope these very interesting links will be further clarified and developed in the future.

References

[1] M.Ajtai; Definability in first order structures; Annals of Pure and Applied Logic 24, (1983) pp 1-48.

[2] M.Ajtai; On the complexity of the pigeonhole principle. 29^{th} Annual symp. on Found. Comp.Sci.(1988),pp 340-355.

[3] M.Ajtai; Parity and the pigeon-hole principle, in Feasible Mathematics Birkhauser, (1990), pp 1-24.

[4] M.Ajtai; The independence of the modulo p counting principles, preprint.

[5] Barrington, Beigel, Rudich; Representing Boolean functions as Polynomials Modulo composite numbers, 24th ACM STOC (1992) pp 455-461.

[6] P.Beame, R. Impagliazzo, J. Krajicek, T. Pitassi, P. Pudlak; Lower bounds on Hilberts Nullstellensatz and propositional proofs, preliminary version.

[7] A.Berrarducci, B.Intrigila; Combinatorial principles in elementary number theory, Annals of Pure and Applied Logic 55 (1991) pp 35-50.

[8] S.Buss; Bounded Arithmetic. Ph.D. dissertation, Princeton University, (1985). As book, Bibliopolis, Napoli (1986).

[9] S.Buss; The propositional pigeonhole principle has polynomial size Frege proofs, Journal of Symbolic logic 52 (1987)) pp 916-927

[10] P.Clote, J.Krajicek, ; Open problems in: Arithmetic, Proof theory and computorial complexity, Oxford university press (1993) pp 289-319.

[11] S.Cook, Recknow; The relative efficiency of Propositional proof systems, Journal of Symbolic logic 44 (1977) pp 36-50

[12] P.D'aquino; Local behavior of the Chebychev theorem in models of $I\Delta_0$; Journal of Symbolic Logic 57, (1992) pp 12-27.

[13] Furst, Saxe, Simpser; Parity circuits and the polynomial time hirarchy, Proc 22nd IEEE FOCS (1981) pp 260-270.

[14] J.Haastad; Almost optimal lower bounds for small depth circuits, Proc 18th ACM STOC (1986) pp 6-20. J. Haastad; Computorial limitations for small depth circuits, Ph.D thesis M.I.T. (1986)

[15] J.Krajicek; On Frege and Extended Frege Proof Systems. To appear.

[16] J.Krajicek, P.Pudlak, and A.Wood, Exponential lower bound to the size of bounded depth Frege proofs of the pigeonhole principle, submitted (1991).

[17] A.Macintyre; The strength of weak systems, in: Schriftenreite der Wittgenstein Gessellschaft 13, Logic, Philosophy of Science and Epistemology, Wien (1987) pp43-50.

[18] J.Paris, A.Wilkie, A.Woods; Provability of the pigeonhole principle and the existence of infinitely many primes. Journal of Symbolic Logic 53, (1988), pp 1231-1244.

[19] J.Paris, A.Wilkie; Counting problems in Bounded Arithmetic, in: Methods in Mathematical Logic, LNM 1130, Springer (1985), pp 317-340.

[20] T.Pitassi, P.Beame, and R.Impagliazzo; Exponential lower bounds for the pigeonhole principle, preprint (1991).

[21] T.Pitassi, P.Beame; An Exponential separation between the Matching Principle and the pigeonhole principle. Proceedings 8^{th}-annual IEEE symposium on computer science (1993), pp 308-319

[22] A.Razborov; On the method of approximation, in: Proc. 21th ACM STOC (1989) pp 168-176.

[23] S.M.Riis; Making infinite structures finite in models of Second Order Bounded Arithmetic, in: Arithmetic, Proof theory and computorial complexity, Oxford university press (1993) pp 289-319.

[24] S.M.Riis; Independence in Bounded Arithmetic; DPhil dissertation, Oxford University (1993)

[25] S.M.Riis; Count(p) versus the pigeon-hole principle; To appear.

[26] S.M.Riis; Finitisation in Bounded Arithmetic; To appear.

[27] R.Smolensky; Algebraic methods in the theory of lower bounds for boolean circuit complexity; Proc. 19th ACM STOC (1987) pp 77-82.

Recent Publications in the BRICS Report Series

RS-94-21 **Søren Riis.** ***Count(q) does not imply Count(p).*** **July 1994, 55 pp.**

RS-94-20 **Peter D. Mosses and Martín Musicante.** ***An Action Semantics for ML Concurrency Primitives.*** **July 1994, 21 pp. To appear in Proc. FME '94 (Formal Methods Europe, Symposium on Industrial Benefit of Formal Methods), LNCS, 1994.**

RS-94-19 **Jens Chr. Godskesen, Kim G. Larsen, and Arne Skou.** ***Automatic Verification of Real–Timed Systems Using*** EPSILON**. June 1994, 8 pp. Appears in: Protocols, Specification, Testing and Verification PSTV '94.**

RS-94-18 **Sten Agerholm.** ***LCF Examples in HOL.*** **June 1994, 16 pp. To appear in:** ***Proceedings of the 7th International Workshop on Higher Order Logic Theorem Proving and its Applications,*** **LNCS, 1994.**

RS-94-17 **Allan Cheng.** ***Local Model Checking and Traces.*** **June 1994, 30 pp.**

RS-94-16 **Lars Arge.** ***External-Storage Data Structures for Plane-Sweep Algorithms.*** **June 1994, 37 pp.**

RS-94-15 **Mogens Nielsen and Glynn Winskel.** ***Petri Nets and Bisimulations.*** **May 1994, 36 pp.**

RS-94-14 **Nils Klarlund.** ***The Limit View of Infinite Computations.*** **May 1994, 16 pp. To appear in the LNCS proceedings of Concur '94, LNCS, 1994.**

RS-94-13 **Glynn Winskel.** ***Stable Bistructure Models of PCF.*** **May 1994, 26 pp.** ***Preliminary draft.*** **Invited lecture for MFCS '94. To appear in the proceedings of MFCS '94, LNCS, 1994.**

RS-94-12 **Glynn Winskel and Mogens Nielsen.** ***Models for Concurrency.*** **May 1994, 144 pp. To appear as a chapter in the** ***Handbook of Logic and the Foundations of Computer Science,*** **Oxford University Press.**