Electronic Colloquium on Computational Complexity
ECCC Technical Reports Series 1994 Report Nr: TR94–014

The Independence of the modulo p Counting Principles

Miklós Ajtai[†]

Received *December 16, 1994*

Abstract. The modulo p counting principle is a first-order axiom schema saying that it is possible to count modulo p the number of elements of the first-order definable subsets of the universe (and of the finite Cartesian products of the universe with itself) in a consistent way. It trivially holds on every finite structure. An equivalent form of the the mod p counting principle is the following: there are no two first-order definable equivalence relations Φ and Ψ on a (first-order definable) subset X of the universe A (or of A^i for some $i = 1, 2, \ldots$) with the following properties: (a) each class of Φ contains exactly p elements, and (b) each class of Ψ with one exception contains exactly p elements, the exceptional class contains 1 element. In this paper we show that the mod p counting principles, for various prime numbers p, are independent in a strong sense.

Keywords: propositional proof-systems, forcing, constant-depth circuits, representations of the symmetric group

[†] IBM Almaden Research Center

Online access for ECCC:
FTP: ftp.eccc.uni-trier.de:/pub/eccc/
WWW: http://www.eccc.uni-trier.de/eccc/
Mail to: ftpmail@ftp.eccc.uni-trier.de, subject "MAIL ME CLEAR", body "pub/eccc/ftpmail.txt"

Introduction The first nonpolynomial lower bound for constant depth Frege proofs were given for the proofs of the Pigeonhole Principle ([Ajt1]). Later this result was improved both in terms of the length of the proof and the constructivity of the methods involved. (See [BIKPPW], [BPU], [KPW], [PBI]). In the search for further tautologies whose proofs are even more difficult than that of the PHP the next step was the so called Parity Principle which is the mod 2 counting principle as defined in the abstract. It was proved in [Ajt2] that even if we are allowed to use PHP_n as an axiom schema (in the sense described above) the Parity Principle has no constant depth polynomial size Frege proof. (The Parity Principle implies PHP_n, in this sense, if it is stated in the form that there is no one-to-one map of $\{1,\dots,n\}$ onto $\{1,\dots,n-1\}$. If we raplace "onto" by "into" then it is not known whether the implication holds.) The Parity-Pigeonhole theorem was further improved in the same sense as the original Pigeonhole result, by Beame and Pitassi, (see [BP]). The proof of the mentioned theorem about the Parity Principle can be modified so that it gives that generally for any fixed p, $CP_{p,n}$ does not imply PHP_n. In this paper we show that for various primes p the statements $CP_{p,n}$ are independent of each other in the sense that if p,q are different primes (of constant size) then $CP_{p,n}$ has no polynomial size constant depth Frege proof even if we use $CP_{q,n}$ as an axiom schema. (The proof yields, without any extra effort, that even if we use $CP_{q_1,n}$,...,$CP_{q_k,n}$ together as an axiom schema, where $q_1,\dots,q_k$ are primes distinct from p and of constant sizes, then $CP_{p,n}$ still has no constant depth polynomial size Frege proof.)

The proof uses a somwehat extended version of the combinatorial, and model theoretical part of the original Pigeonhole and Parity-Pigeonhole results. The essential new part of the proof is the application of a theorem about the solution of a symmetric system over a finite field. (We prove this theorem in [Ajt4].) The theorem essentially says that if there is a linear system (mod p) whose variables are indexed by sequences of length k from a set A of size n (k is fixed and n is sufficiently large), and the system is invariant under permutations of A, then a solution (if exists) can be given in a way which is essentially independent of n. We will use this theorem in the following way: we will have a symmetric system for each n and we have to decide, for which n does it have a soltuion. This presentation of a solution will guarantee that the the fact whether our specific system has a solution or not depends only on the behavior of n

modulo p^c where c is a constant. The classical solutions (like Cramer's rule) would have led to the hopelessly difficult problem of determining the value of a complicated $n \times n$ determinant mod p. Later we give an exact formulation of a corollary of this theroem (Theorem **4**) that we will actually use in our proof.

Finally we note that Soren Riis gave independently a different proof of the main theorem.

1. Suppose that $\mathcal{L}$ is a first-order language with equality and a finite number of relation and function symbols. Assume also that $\leq$ is a binary relation symbol of $\mathcal{L}$ and T is a first-order theory in $\mathcal{L}$. We will say that T is weakly finite if for any natural number n, T has a model whose universe is finite and contains more than n elements and in any model of T, the relation $\leq$ is a total order of the universe.

We will say that a language $\mathcal{L}'$ is an extension of $\mathcal{L}$ if it contains all of the relation and function symbols of $\mathcal{L}$ (and possibly others too). Assume that τ is an interprertation of $\mathcal{L}$ and τ' is an interpretation of $\mathcal{L}'$. We say that τ' is an extension of τ, if τ' as a function is an extension of τ, that is, the universe of the two interpretations are the same and the relation and function symbols of $\mathcal{L}$ has the same interpretation according to τ and τ'.

Definitions. 1. Assume that A is the universe of an interpretation σ of a firstorder language, $\tilde{\mathcal{L}}$ and i is a positive integer, $X \subseteq A^i$. We say that X is first-order definable if there is a finite sequence $a_1, \ldots, a_j$ from the elements of A and a first-order formula ϕ of $\mathcal{L}$ with $j + i$ free variables so that for each $b_1, \ldots, b_i \in A$ we have $\langle b_1, \ldots, b_i \rangle \in A$ iff $\models_\sigma \phi(b_1, \ldots, b_i, a_1, \ldots, a_j)$.

2. Assume that X is an arbitrary set and E is an equivalence relation on X and p is a prime number. We say that E is a p partition of X if each class of E contains exactly p elements. E will be called a p-exceptional partition of X if each class of E with one exception contains exactly p elements and the exceptional class contains 1 element.

3. Assume that that X is a subset of A or of A^i for some $i = 1, 2...$, where A is the universe of an interpretation of a first-order language. We will say that the p equipartition principle holds for X if there are no first-order definable equivalence

relations E_1, E_2 on X so that E_1 is a p-partition of X and E_2 is a p-exceptional partition of X.

4. We say that the modulo p counting principle holds in an interpretation of a first-order language if for any positive i and any first-order definable subset X of A^i, where A is the universe of the interpretation, the p-equipartition principle holds for X.

Clearly for any fixed language, the the modulo p counting principle can be stated as a first-order axiom-schema S, that is the principle holds on a structure iff each axiom of S holds on it. For an alternative definition of the modulo p counting principle and further motivation see [Ajt4]. (As we have indicated in the abstract, if there is a first-order definable ordering of the universe A, then the modulo p counting principle is equivalent to the existence of a function on the first-order definable subsets of A^i, $i = 1, 2, ...$ with values in F_p so that the function satisfies the usual properties of the cardinality function, that is, it is additive on disjoint sets, multiplicative on direct products, invariant under first-order definable one-to-one maps, and takes the value 1 on singletons. Such a function can be called a modulo p cardinality function.)

Suppose now that p and q are two distinct prime numbers. The following theorem says that the "modulo p counting principle" and the "modulo q counting principle" are independent in a strong sense.

Theorem 1 . *Suppose that T is a weakly finite theory of the language $\mathcal{L}$. Then there exists an extension $\mathcal{L}'$ of $\mathcal{L}$ so that the following theory T' in $\mathcal{L}'$ is consistent:*

T + "the axiom schema of the modulo p counting principle" + $\neg$ "the axiom schema of the modulo q counting principle".

$\mathcal{L}'$ will be the language that we get from the language $\mathcal{L}$ by adding a new binary relation symbol Ψ and the ternary relations $+$ and $\times$ which define the arithmetic operations in Peano Arithmetic. Let M be an arbitrary nonstandard model of Penao Arithmetic and let n be a nonstandard element of M so that there is an interpretation τ of $\mathcal{L}$ on the universe $\{0, 1, ..., n\}$ so that $\leq$ has the same interpretation according to τ as in Peano Arithmetic. (Since T is weakly finite there is such an interpretation of T for an infinite set of standard positive integers n. Moreover all of these interpretations

are elements of N. Therefore there is a nonstandard n as well with this property. E.g. if m is an arbitrary nonstandard element of N then the greatest $n \leq m$ with the given property will be necessarily nonstandard.) We assume for the moment that $n \equiv 0 \pmod{q}$. Later we will explain why this assumption can be dropped. We will show that there is an equivalence relation $\bar{\Psi}$ on the set $\{0, 1, ..., n\}$ (this is an infinite set) so that if τ' is the extension of τ defined by $\tau'(\Psi) = \bar{\Psi}$, then we have

(**2**) $\tau' \models T +$ "Ψ is an equivalence relation on the universe and each class of Ψ contains exactly q elements" + "the modulo p counting principle".

That is, we want to show that the mod q-equipartiton principle does not hold on the set $\{0, 1, ..., n\}$, but the mod p-equipartiton principle holds not only on $M_{n+1} = \{0, 1, ..., n\}$ but on every first-order definable subsets of M_{n+1} and also, for each standard i, on every first-order definable subsets of the Cartesian product $M_{n+1}^i = M_{n+1} \times \ldots \times M_{n+1}$ (i copies).

Since M is a model of Peano Arithmetic and the number of elements of the set $\{0, ..., n-1\}$ is divisible by q (in M) it is easy to give a first-order definable equivalence relation on $\{0, 1, ..., n-1\}$ so that the number of elements in each class is divisible by q. E.g we may divide this set into consecutive intervals of length q. This shows that the universe $\{0, 1, ..., n\}$ has a first-order definable q exceptional partition already in τ. Therefore the property of Ψ given in (2) implies that $\tau' \models \neg$"the modulo q counting principle" and so out theorem is a consequence of (2).

Sketch of the proof.

We will construct $\bar{\Psi}$ in the following way. For each $\epsilon = 1/k$ where k is a standard positive integer we pick a q-partition Ψ_ϵ of a subset of $\{1, \ldots, n\}$ with $n - n^\epsilon$ elements, so that $\Psi_\epsilon \in M$. Moreover we pick the sequence in a way that if $\epsilon < \epsilon'$ then Ψ_ϵ is an extension of $\Psi_{\epsilon'}$. The common extension of all of the partitions Ψ_ϵ will be Ψ. We show that the sequence can be picked in a way so that

(**P**) for any first-order formula ϕ the truth value of $\phi(a)$, $a \in \{1, \ldots, n\}$ can be decided in the knowledge of a suitbale Ψ_ϵ which does not depend on a, and in the knowledge of those classes of $\bar{\Psi}$ which contain a set $U_a \subseteq \{1, \ldots, n\}$ of constant size, where U_a does not depend on $\bar{\Psi}$.

(The method of constructing such a model of $M[\bar{\Psi}]$ is a variant of Cohen's method of forcing as described in [Ajt2].) The essential property of the extended model $M[\bar{\Psi}]$

is that first-order formulae over it can be decided using only partial information about $\bar{\Psi}$.

We want ot show that $M[\bar{\Psi}] \models$"mod p counting principle". For the sake of simplicity we assume that $p \neq 2$, the $p = 2$ case is only slightly more complicated. Suppose that the mod p counting principle does not hold. We show that, then there exists a first-order definable weight function χ on the edges of a complete directed graph G with n^c nodes, taking values in the finite field with p elements so that,

(1) for each fixed $x_0 \in G$ we have
$|\{y \in G | \chi(x_0, y) \neq 0\}| \leq p$, and
$|\{y \in G | \chi(y, x_0) \neq 0\}| \leq p$.

(2) $\chi(x, y) = -\chi(y, x)$ for all $x, y \in G$

(3) for all $x_0 \in G$ we have $\sum_{y \in G} \chi(x_0, y) = 0$. ((1) implies that the sum is defined.)

(4) there is an $y_1 \in G$ so that $\sum_{x \in G} \chi(x, y_1) = 1$ and
for all $y_0 \neq y_1$, $y_0 \in G$ we have $\sum_{x \in G} \chi(x, y_0) = 0$.

Such a function could not exist in a model of Peano Arithmetic since (2),(3) and the second part of (4) would imply that $\sum_{x \in G} \chi(x, y_1) = 0$ in contradiction to the first part of (4). Although $M[\bar{\Psi}]$ is not a model of Peano Arithmetic so this argument is not valid there, using property (P) we show that there is a structure in M, which is similar enough to $\langle G, \chi \rangle$ so that we still get a contradiction. This structure will provide a solution for a symmetric system E_n of linear equations. (χ itself was a solution of a symmetric system too, if we consider the premutations only of $G - \{y_1\}$).

We use the following observation in the proof: if $q | n + 1$ then the construction of $\bar{\Psi}$ can be carried out in M, therefore the pair $\langle G, \chi \rangle$ is in M, what, as we have seen, is impossible. Therefore we know that if $q | n + 1$ then E_n has no solution. On the other hand using our theorem about the symmetric linear equation we show that the fact whether our system has a solution or not depends only on the residue class of n mod p^c, if n is sufficiently large, where c is a constant (standard) integer. Clearly each residue class mod p^c contains an integer n so that $q | n + 1$ therefore the system E_n has no solution for any sufficiently large integer n.

We will construct $\bar{\Psi}$ in the following way. For each $\epsilon = 1/k$ where k is a standard positive integer we pick a q-partition Ψ_ϵ of a subset of $\{1, \ldots, n\}$ with $n - n^\epsilon$ elements,

so that $\Psi_\epsilon \in M$. Moreover we pick the sequence in a way that if $\epsilon < \epsilon'$ then Ψ_ϵ is an extension of $\Psi_{\epsilon'}$. The common extension of all of the partitions Ψ_ϵ will be Ψ. We show that the sequence can be picked in a way so that

(**P**) for any first-order formula ϕ the truth value of $\phi(a)$, $a \in \{1, \ldots, n\}$ can be decided in the knowledge of a suitbale Ψ_ϵ which does not depend on a, and in the knowledge of those classes of $\bar{\Psi}$ which contain a set $U_a \subseteq \{1, \ldots, n\}$ of constant size, where U_a does not depend on $\bar{\Psi}$.

(The method of constructing such a model of $M[\bar{\Psi}]$ is a variant of Cohen's method of forcing as described in [Ajt2].) The essential property of the extended model $M[\bar{\Psi}]$ is that first-order formulae over it can be decided using only partial information about $\bar{\Psi}$.

We want ot show that $M[\bar{\Psi}] \models$"mod p counting principle". For the sake of simplicity we assume that $p \neq 2$, the $p = 2$ case is only slightly more complicated. Suppose that the mod p counting principle does not hold. We show that, then there exists a first-order definable weight function χ on the edges of a complete directed graph G with n^c nodes, taking values in the finite field with p elements so that,

(1) for each fixed $x_0 \in G$ we have
$|\{y \in G | \chi(x_0, y) \neq 0\}| \leq p$, and
$|\{y \in G | \chi(y, x_0) \neq 0\}| \leq p$.

(2) $\chi(x, y) = -\chi(y, x)$ for all $x, y \in G$

(3) for all $x_0 \in G$ we have $\sum_{y \in G} \chi(x_0, y) = 0$. ((1) implies that the sum is defined.)

(4) there is an $y_1 \in G$ so that $\sum_{x \in G} \chi(x, y_1) = 1$ and
for all $y_0 \neq y_1$, $y_0 \in G$ we have $\sum_{x \in G} \chi(x, y_0) = 0$.

Such a function could not exist in a model of Peano Arithmetic since (2),(3) and the second part of (4) would imply that $\sum_{x \in G} \chi(x, y_1) = 0$ in contradiction to the first part of (4). Although $M[\bar{\Psi}]$ is not a model of Peano Arithmetic so this argument is not valid there, using property (P) we show that there is a structure in M, which is similar enough to $\langle G, \chi \rangle$ so that we still get a contradiction. This structure will provide a solution for a symmetric system E_n of linear equations. (χ itself was a solution of a symmetric system too, if we consider the premutations only of $G - \{y_1\}$).

We use the following observation in the proof: if $q|n+1$ then the construction of $\bar{\Psi}$ can be carried out in M, therefore the pair $\langle G, \chi \rangle$ is in M, what, as we have seen, is impossible. Therefore we know that if $q|n+1$ then E_n has no solution. On the other hand using our theorem about the symmetric linear equation we show that the fact whether our system has a solution or not depends only on the residue class of n mod p^c, if n is sufficiently large, where c is a constant (standard) integer. Clearly each residue class mod p^c contains an integer n so that $q|n+1$ therefore the system E_n has no solution for any sufficiently large integer n.

Now we return to the actual proof. In the proof of (2) we will use the following notions which were introduced in [Ajt4] in a slightly more general form. Examples motivating the following definitions were also provided there.

We will be interested in linear equations over F_p, where F_p is the field with p elements. The variables in the equation will be associated with sequences of length k consisting of subsets of H and $H \times H$, where H is a fixed finite set. Actually we will have several group of variables each group will be associated with an element of a set Γ. More precisely we have the following definiton:

Defintion. Suppose that H is a finite set, $|H| = n$ and t, r are a positive integers. Let $H^{(t,r)}$ be the set of all sequences $D_1, ..., D_{t'}$, $1 \leq t' \leq t$ with the following properties:

(a) for each $i = 1, ..., t'$ either $D_i \subseteq H$ or $D_i \subseteq H \times H$,

(b) for each $i = 1, ..., t'$, $|D_i| \leq r$.

Let $\Delta = H^{(t,r)}$. With each element of δ of Δ we associate a variable x_δ. We will be concerned with systems of linear equations where each single equation is of the form $\sum_\delta a_\delta x_\delta \in \Delta = b$ where $a_\delta \in F_p$, $b \in F_p$ for all $\delta \in \Delta$. If we have several equations then we will use an other subscript $\kappa \in K$ that is our system of equations will be:

(**3**) $\sum_{\delta \in \Delta} a_{\delta,\kappa} x_\delta = b_\kappa$

where $a_{\delta,\kappa} \in F_p$, $b_\delta \in F_p$ for all $\kappa \in K, \delta \in \Delta$. We will be interested in solutions with values in F_p.

2. Assume now that π is a permutation of H. The map π can be extended in a natural way onto the subsets of H or $H \times H$ by $D\pi = \{x\pi |\ x \in D\}$ if $D \subseteq H$ and

$D\pi = \{\langle x\pi, y\pi\rangle |\ \langle x, y\rangle \in D\}$ if $D \subseteq D \times D$. We may further extend π onto Δ by $\langle D_1, ..., D_{t'}\rangle\pi = \langle D_1\pi, ..., D_{t'}\rangle\pi$. We will say that the linear system (3) is symmetric if for each $\kappa \in K$ the equation $\sum_{\delta\in\Delta} a_{\delta\pi,\kappa} x_{\delta\pi} = b_\kappa$ is also an equation of the system. The symmetric hull of a system will be the smallest symmetric system containing it.

3. Let H' be a subset of H and $\Delta' = H'^{(t,r)}$. Suppose that a' is a F_p-valued function defined on Δ' and $H_0 \subseteq H'$, $|H' - H_0| > rt$ and g is a F_p valued function defined on t. We will say that the system (3) is based on the quadruplet a', H_0, H', g if the following holds:

(a) $b_\kappa = g(\kappa)$ for all $\kappa \in K$

(b) for any permutation π of H which fixes each element of H_0, and for all $\delta \in \Delta$, $\kappa \in K$ if $\delta\pi \in \Delta'$ then $a_{\delta,\kappa} = a'_{\delta\pi,\kappa}$.

4. We will say that a symmetric system E is induced by the quadruplet a', H_0, H', g (over H) if E is the symmetric hull of a system based on this quadruplet.

The following theorem is an immediate consequence of Corollary **D6** of [Ajt3].

Theorem. 4 . *For all positive integers t, l, r and prime p there is a positive integer c so that if $H_0 \subseteq H'$, $|H' - H_0| > rt$, Γ, K are finite sets, $\Gamma \leq l$, $|H'| \leq l$, $|K| \leq l$ and a is a F_p valued function on $\Delta' = H'^{(t,r)}$ and g is a F_p valued function on K then the following holds:*

There is a subset Q of the residue classes modulo p^c, so that for any sufficiently large n, if H is a set containing n elements, $H' \subseteq H$, and the symmetric system E is induced by the quadruplet a', H_0, H', g over H, then E has a solution iff $n \equiv s \pmod{p^c}$ for some $s \in Q$.

Definitions. 1. A partition of a set S is a set of pairwise disjoint subsets of S, whose union is S.

2. We will say that the partiton P of the set S is a q-partition if each class of P is a set with exactly q elements.

3. A q-partition of a set $S' \subseteq S$ will be called a partial q-partiton of S.

4. The partial q-partitons P, Q of the set S will be called compatible if every class of P is either a class of Q too, or disjoint from every class of Q.

5. Suppose that P is a partial q-partition of the set S and $V \subseteq S$. We say that V covers P iff each class of P contains at least one element from V. We say that V is inside P iff V is a subset of $\bigcup P$. V supports P iff it is both inside P and covers P.

Suppose $\bar{\Psi}$ is a q-partition of $M_n = \{0, 1, ..., n\}$. (Of course $\bar{\Psi}$ is not an element of M.) Let τ' be the extension of τ with $\tau'(\Psi) = \bar{\Psi}$. Obviously $\models_{\tau'} T \wedge \neg$ "the modulo q counting prinicple for Ψ". For the proof of Theorem 1, it is sufficient to prove that for a suitable choice of $\bar{\Psi}$, we have $\models_{\tau'}$ "the axiom-schema of the modulo p counting principle".

We will construct $\bar{\Psi}$ with a variant of Cohen's method of forcing. The construction is described in section 2. It will be trivial that $\bar{\Psi}$ is a partition of M_{n+1} and each class of $\bar{\Psi}$ has exactly q element. The more difficult part of the proof is to show that the "modulo p counting prinicple" holds for τ', that is, in the structure $M_n[\bar{\Psi}]$. In section 2 we will show that if the "mod p counting principle" is not true in $M_n[\bar{\Psi}]$ then already in M there must be certain objects which have properties similar to that of $\bar{\Psi}$. In this section we only formulate these properties (see conditions (5) below) and show that they cannot hold in any model of Peano Arithmetic. The missing part of the proof will be given in section 2. We will prove there that if the "modulo p counting principle" does not hold in $M[\bar{\Psi}]$, and M was an elementary extension of the standard model of Peano Arithmetic then the following statement holds in the standard model of Peano Arithmetic: (for the sake of simplicity we formulate the statement only for the $p \neq 2$ case. The $p = 2$ case is only slightly more complicated. We will show at the end of this section how to handle that.)

(**5**) there exist a positive integer k so that for all m_0 there exists a positive integer $m > m_0$, a finite set H with m elements and two functions f, g with the following properties:

(**6**) f is a function of three variables and $f(U, V, Q)$ is defined iff $U \subseteq H$, $V \subseteq H$, $|U| = |V| = k$, and Q is a partial q-partition supported by U.

(**7**) each value of f is an element of F_p

(**8**) if the partial q partitions P and Q are compatible and both $f(U, V, P)$, and $f(V, U, Q)$ are defined, then $f(U, V, P) = -f(V, U, Q)$.

(**9**) the function g (of one variable) is defined on the set of all subsets of H with k elements.

(**10**) if $U \subseteq H$, $|U| = k$ and the partial q-partition P is supported by U then $\sum_{V \subseteq H, |V|=k} f(U,V,P) = g(U)$.

(**11**) $\sum_{U \subseteq H, |U|=k} g(U) = 1$.

Clearly the values of the functions f and g satisfiy certain linear equations. We want to write up this equations using the formalism of Theorem 4. Let $t = 3$ and $r = kq$. We will have three groups of equations. We get the groups from (8), (10) and (11).

Let K_1 be the set of all quadruplets $\langle U, V, P, Q\rangle$ so that U, V are subsets of H with k elements and P is a partial q partition supported by U and Q is a partial q partition supported by V, and P, Q are compatible. If $\kappa = \langle U, V, P, Q\rangle \in K_1$ then let $a_{\langle U,V,P\rangle,\kappa} = 1$, $a_{\langle V,U,Q\rangle,\kappa} = -1$. $a_{\delta,\kappa} = 0$ for all other values of $\delta \in \Delta$. Finally let $b_\kappa = 0$ for all $\kappa \in K_1$.

Let K_2 be the set of all pairs $\langle U, P\rangle$ so that U is a subset of H with k elements and P is a partial q partition supported by U. If $\kappa = \langle U, P\rangle \in K_2$ and V is any subset of H with k elements then let $a_{\langle U,V,P\rangle,\kappa} = 1$, $a_{\langle U\rangle,\kappa} = -1$ and $a_{\delta,\kappa} = 0$ for all other values of $\delta \in \Delta$. Finally let $b_\kappa = 0$ for all $\kappa \in K_2$.

Let K_3 be an arbitrary set with one element which is disjoint from $K_1 \cup K_2$. If $\kappa \in K_3$ then for all subset U of H with exactly k elements let $a_{\langle U\rangle,\kappa} = 1$ and for all other values of $\delta \in \Delta$, let $a_{\delta,\kappa} = 0$. Finally let $b_\kappa = 1$ for the single $\kappa \in K_3$.

If $K = K_1 \cup K_2 \cup K_3$ then E_H will be the following system of linear equations

(**12**) $\sum_{\delta \in \Delta} a_{\delta,\kappa} x_\delta = b_\kappa$, $\kappa \in K$

Clearly system E_H is symmetric. Properties (5.1),...,(11) imply that the following evaluation of the variables $x_\delta, \delta \in \Delta$ is a solution of E_H:

(a) if U, V are subsets of H with k elements and P is a partial q partition of H supported by U then $x_{\langle U,V,P\rangle} = f(U,V,P)$

(b) if U is a subset of H with k elements then $x_{\langle U\rangle} = g(U)$

(c) $x_\delta = 0$ for all other choices of $\delta \in \Delta$.

Now we give subsets $H_0 \subseteq H' \subseteq H$ so that $|H'|$ depends only on k (and not on $m = |H|$), a function a' defined on $\Delta' \times K'$, where $\Delta' = H'^{3,kq}$, $K' \subseteq K$, and a function g defined on K' so that the system E_H is induced by the quadruplet a', H_0, H', g over H.

Let H' be an arbitrary subset of H with $6kq$ elements and let $H_0 \subseteq H'$, $|H_0| = 2kq$.

Finally let $K' = K'_1 \cup K'_2 \cup K_3$, where

$\langle U, V, P, Q\rangle \in K_1$ will be an element of K'_1 too, iff $U, V \subseteq H_0$ and $P, Q \subseteq H_0 \times H_0$;

$\langle U, P\rangle \in K_2$ will be in K'_2 iff $U \subseteq H_0$, $P \subseteq H_0 \times H_0$.

We define g by $g(\kappa) = 0$ if $\kappa \in K'_1 \cup K'_2$ and $g(\kappa) = 1$ if $\kappa \in K_3$.

a' is the restriction of a onto $\Delta' \times K'$. It is easy to check that E_H is induced by the quadruplet a', H_0, H, g.

According to Theorem 4 if we consider now all of the possible sets $\bar{H} \supseteq H'$ the fact wether the system $E_{\bar{H}}$ has a solution or not depends only the value of $|\bar{H}|$ modulo p^c (at least if $|\bar{H}|$ is sufficiently large.) We will get a contradiction by showing that if $\bar{m} = |\bar{H}|$ is divisible by q then $E_{\bar{H}}$ has no solution. Indeed, in each residue class modulo p^c there are infinitely many integers which are divisible by q and so by Theorem 4, $E_{\bar{H}}$ has no solution if $|\bar{H}|$ is sufficiently large, in contradiction to (5).

Suppose now that $|H|$ is divisible by q. Then there is a q partition S of the whole set H. Suppose that E_H has a solution. Then there are functions f, g with (6),...,(11). We define binary function $\mathcal{F}$ on, $[H]^k$, the set of subsets of H with k elements. If $U, V \in [H]^k$ then let P be the partial q partition of H supported by U that we get from S by keeping only those classes of S which has a nonempty intersection with U. Let $\mathcal{F}(U, V) = f(U, V, P)$. According to (6),...,(11), the functions $\mathcal{F}$ and g have the following properties:

(a) $\mathcal{F}(U, V) = -\mathcal{F}(V, U)$ for all $U, V \in [H]^k$,

(b) $\sum_{V \in [H]^k} \mathcal{F}(U, V) = g(U)$ for all$U \in [H]^k$

(c) $\sum_{U \in [H]^k} g(U) = 1$.

This is however impossible since if we add the equations (b) for each $U \in [H]^k$ then we will get the following:

$\sum_{U,V \in [H]^k, U \neq V} (\mathcal{F}(U, V) + \mathcal{F}(V, U)) + \sum_{U \in [H]^k} \mathcal{F}(U, U) = \sum_{U \in [H]^k} g(U)$.

(a) implies that $\mathcal{F}(U,V)+\mathcal{F}(V,U)=0$ for any $U,v\in[H]^k$. Since $p\neq 2$ we we have $\mathcal{F}(U,U)=0$ too. Therefore $\sum_{U\in[H]^k} g(U)=0$ in contradiction to (c).

The $p=2$ case can be handled in the following way: we will show that in this case (5) remains true if we add the following property of the function f:

(**13**) $\quad f(U,U)=0$ for all $U\in[H^k]$.

The proof remains the same, only we have to include the corresponding equation in the system E_H.

Finally we explain why can we drop the $n\equiv 0 \pmod q$ condition.

Let $n+1\equiv a \pmod q$, $0<a\leq q$. Then there are positive integers n', r so that $n'\equiv 0 \pmod q$ and $n+1=a(n'+1)+rq$ where $r\leq q^2$. This means that the interval $[0,n]$ can be cut into $a+1$ disjoint subintervals so that the first a subintervals $I_1,...,I_a$ are of length $n'+1$ and the last subinterval I_{a+1} is of length rq. Applying the already proven part of the theorem with $n\to n'$ we get an equivalence relation $\bar{\Psi}$ on the set $\{0,...,n'\}$ so that each class of $\bar{\Psi}$ contains exactly q elements. Since $|I_1|\equiv 1 \pmod q$ there is an equivalence relation $\bar{\Theta}$ so that it is first-order definable in the structure $\{0,...,n,\leq,+,\times\}$, and it is q-exceptional on I_1.

Using $\bar{\Psi}$ and the arithmetic operations of Peano arithmetic restricted to the set $\{0,...,n\}$ we may easily define by a first-order formula an equvalence relation Θ_i on I_i so that each class of Θ_i contains exactly q elements for $i=1,...,a,a+1$. Therefore in the structure $\langle M_n,\leq,+,\times,\bar{\Psi}\rangle$ we may define by first-order formulas two equivalence relations $\tilde{\Psi}_0$, $\tilde{\Psi}_1$ so that

(a) $\quad \tilde{\Psi}_0$ is the common extension of the equivalence relations $\Theta_1,...,\Theta_{a+1}$ and therefore each class of $\tilde{\Psi}_0$ contains exactly q elements

(b) $\quad \tilde{\Psi}_1$ is the common extension of the equivalence relations $\bar{\Theta}$, $\Theta_2,...,\Theta_{a+1}$, and so $\tilde{\Psi}_1$ is q-exceptional on $\{0,...,n\}$.

Let $\mathcal{L}''$ be an extension of the language $\mathcal{L}$ with two binary relation symbols Ψ_0 and Ψ_1 and let τ'' be an extension of the interpretation τ of $\mathcal{L}$ defined by $\tau''(\Psi_0)=\tilde{\Psi}_0$ and $\tau''(\Psi_1)=\tilde{\Psi}_1$. We claim that

$\tau''\models T+$ "Ψ_0 is an equivalence relation on the universe and each class of Ψ_0 contains exactly q elements" + "Ψ_1 is a q-exceptional equivalence relation on the universe" + "the modulo p counting principle".

The first three statement trivally holds we have to show only that $\tau'' \models$ "the modulo p counting priniciple". The already proven part of the theorem implies that if τ_0 is an extension of τ to the language $\mathcal{L} \cup \{\Psi\}$ defined by $\tau_0(\Psi) = \bar{\Psi}$ then $\tau_0 \models$ "the modulo p counting principle". Since the first-order definable subsets of the structures defined by τ'' and τ are essentially the same, the definition of the modulo p counting principle implies that, $\tau \models$ "the modulo p counting principle". *Q.E.D.*(Theorem (1))

2. In this section we construct the equivalence relation $\bar{\Psi}$ and show that if the mod p counting principle does not hold for $\bar{\Psi}$ then (5) holds. This will complete the proof of theorem 1

Assume that $\bar{\Psi}$ is a q-partition of M_{n+1} and the mod p counting principle does not hold in τ'. Then there exist first-oder definable (in τ') equivalence relations E_1, E_2 on a first-order definable subset X of M_{n+1}^r (for some standard positive integer r) so that E_1 is a p-partition and E_2 is a p exceptional partition of X.

First we show that this implies that there is a standard positive integer i (depending only on p and r) and a function $\chi(x, y)$ of two variables defined on M_{n+1}^i with the following properties.

(**M1**) $\chi(x, y) \in F_p$ for all $x, y \in M_{n+1}^i$

(**M2**) χ is first-order definable in τ' in the sense that each set $\chi^{-1}(u) = \{\langle x, y\rangle | \chi(x, y) = u\}$, $u \in F_p$ is first-order definable.

(**M3**) for each fixed $x_0 \in M_{n+1}^i$ we have
$|\{y \in M_{n+1}^i | \chi(x_0, y) \neq 0\}| \leq p$, and
$|\{y \in M_{n+1}^i | \chi(y, x_0) \neq 0\}| \leq p$.

(**M4**) $\chi(x, y) = -\chi(y, x)$ for all $x, y \in M_{n+1}^i$

(**M5**) for all $x_0 \in M_{n+1}^i$ we have $\sum_{y \in M_{n+1}^i} \chi(x_0, y) = 0$. ((M3) implies that the sum is defined.)

(**M6**) there is an $y_1 \in M_{n+1}^i$ so that $\sum_{x \in M_{n+1}^i} \chi(x, y_1) = 1$ and
for all $y_0 \neq y_1$, $y_0 \in M_{n+1}^i$ we have $\sum_{x \in M_{n+1}^i} \chi(x, y_0) = 0$.

Proof. Let $i = rp + 2$ and let μ be a function on M_{n+1}^i with following properties

(a) for each $a \in M^i_{n+1}$, either $\mu(a) = 0$ or $\mu(a) = \langle B, \delta \rangle$, where $\delta \in \{0,1\}$ and B is a class of E_δ,

(b) for each class B of E_0 or E_1, there is exactly one element a of M^i_{n+1},

(c) the function μ is first-order definable in τ'.

The existence of such a function μ follows from the fact that each class of E_0 or E_1 contains at most p elements from M^r_{n+1} and therefore can be represented uniqualy by a strictly increasing sequence of length at most rp from the elements of M_{n+1}, therefore by an element of M^{rp}_{n+1} and a number giving the length of the sequence. Consequently an element of M^{pr+2}_{n+1} may code all of this information and the fact whether it is a class of E_0 or E_1. μ will be fistorder definble in the sense that the relation "$\mu(a)$ is an element of E_δ and $x \in \mu(a)$" is first-order definable for $a \in M^i_{n+1}$, $\delta \in \{0,1\}$ and $x \in M_{n+1}$.

Now we may define the function $\chi(x,y)$ for all $x, y \in M^i_{n+1}$ by the following rules:

(a) if $\mu(x) \neq 0$, $\mu(y) \neq 0$, $\mu(x) \in E_0$, $\mu(y) \in E_1$ then let $\chi(x,y)$ be the residue class of $-|\mu(x) \cap \mu(y)|$,

(b) if $\mu(x) \neq 0$, $\mu(y) \neq 0$, $\mu(x) \in E_1$, $\mu(y) \in E_0$ then let $\chi(x,y)$ be the residue class of $|\mu(x) \cap \mu(y)|$,

(c) in all of the cases not covered by (a) or (b) let $\chi(x,y) = 0$.

We show that properties (M1)-(M6) hold for χ. (M1)-(M4) are immediate consequences of the definition. For the proof of (M5) and (M6) it is sufficient to notice that $\sum_y \chi(x_0, y)$ is $-|\mu(x_0)|$, (mod p), provided that $\mu(x_0) \neq 0$. If $\mu(x) = 0$ then each term of the sum is 0. (In a similar way the sum $\sum_y \chi(y, x_0)$ is either $|\mu(x_0)|$ (mod p) or 0.) Each of the classes of E_0 has exactly p elements therefore (M5) holds. Each of the classes of E_1 but the exceptional one has p elements and the exceptional has 1 element. We get y_1 of (M6) from the exceptional class. *Q.E.D.*((M1)-(M6))

We will not construct only a single $\bar{\Psi}$ but a family of q partitions $\bar{\Psi}$ of M_{n+1}, whose set will be denoted by Γ. As we have shown, if the mod p counting principle does not hold in any of the interpretations $\tau'_{\bar{\Psi}}$, $\bar{\Psi} \in \Gamma$, then for each $\bar{\Psi} \in \Gamma$ there is a function $\chi = \chi_{\bar{\Psi}}$ with properties (M1),...,(M6). The domains of the functions $\chi_{\bar{\Psi}}$ in principle could be different for different elements $\bar{\Psi} \in \Gamma$ but our contstruction will

imply that the domain is the same set: M_{n+1}^i where i is a fixed standard integer. (As our proof shows if i is fixed then we can define $\chi_{\bar{\Psi}}$ from $\bar{\Psi}$ by a first-order formula.)

We will show moreover that the family Γ can be given in a way that the following conditions are satisfied by a suitable $D \subseteq M_{n+1}$ and a standard positive integer k:

(**R1**) $D \in M$ and $M \models n - n^{1/k} \leq |D| \leq n - n^{1/(2k)}$.

(**R2**) each class of each $\bar{\Psi} \in \Gamma$ is either contained in D or disjoint from it. Moreover each $\bar{\Psi} \in \Gamma$ induces the same q-partition of D if we restrict it to D. (We will denote this q-partition of D by Ψ_D).

(**R3**) If P is a partial q-partition of M_{n+1} so that $|\bigcup P - D| \leq 4kq$ and P is an extension of Ψ_D then there is a $\bar{\Psi} \in \Gamma$ so that $\bar{\Psi}$ is compatible with P.

Remark. The next property (R4) roughly tells the following: the value of $\chi_{\bar{\Psi}}(x, y)$ can be decided without knowing the whole $\bar{\Psi}$, namely for each $x \in M_{n+1}^i$ there is a $U_x \subseteq M_{n+1} - D$, $|U_x| = k$ so that those classes of $\bar{\Psi}$ which intersect U_x, uniqualy determine $\chi(x, y)$ for any y. (In a similar way if we know $\bar{\Psi}$ on U_y it determines $\chi(x, y)$ for any x.) Moreover the function $x \to U_x$ is an element of M, and if η is the function which gives the value of $\chi_{\bar{\Psi}}(x, y)$ if x, y and the classes of $\bar{\Psi}$ intersecting U_x is given as input, then η is an element of M.

(**R4**) Let W be the set of all triplets $\langle x, P, \delta \rangle$, so that $x \in M_{n+1}^i$, P is a partial q-partition of M_{n+1} which is an extension of $\bar{\Psi}$, $|\bigcup P - D| \leq 4kq$ and $\delta = 0$ or $\delta = 1$. Then there is a function $x \to U_x$, in M, defined on M_{n+1}^i so that, $U_x \subseteq M_{n+1} - D$, $|U_x| = k$ for all $x \in M_{n+1}^i$, and there is a function η defined on $W \times M_{n+1}$ so that $\eta \in M$ and if $\langle x, P, \delta \rangle \in W$, $y \in M_{n+1}$, $\bar{\Psi} \in \Gamma$,$\bar{\Psi}$ is an extension of P and P covers U_x then $\chi_{\bar{\Psi}}(x, y) = \eta(\langle x, P, 0 \rangle, y)$ and $\chi_{\bar{\Psi}}(y, x) = \eta(\langle x, P, 1 \rangle, y)$.

We will prove in the next section that such a set Γ really exists. Using the assumption that a Γ exists with properties (R1)-(R4) now we are able to define the functions f and g of (5). Let $H = M_{n+1} - D$. Suppose that $U, V \subseteq H$, $|U| = |V| = k$ and Q is a partial p-partition supported by U. Let $\bar{\Psi}$ be an arbitrary element of Γ compatible with Q (such a $\bar{\Psi}$ exists according to (R3)). Let

$f(U, V, Q) = \sum\{\chi_{\bar{\Psi}(x,y)} | x \in M_{n+1}^i, U = U_x, y \in M_{n+1}^i, V = U_y\}$.

(R4) implies that this sum really can be defined in M and so the function f is an element of M, moreover the value $f(U, V, Q)$ does not depend on the choice of $\bar{\Psi}$.

Now we define the function $g(U)$. Suppose that $U \subseteq H$, $|U| = k$. If $U = U_{y_1}$, where y_1 is defined in (M6), then let $g(U) = 1$ otherwise let $g(U) = 0$.

We have to show that conditions (6)-(11) are satisfied. (6),(7) and (9) obviously hold since they were part of the definition.

Proof of (8). The definition of f implies that both $f(U,V,P) = \sum\{\chi_{\Psi_1}(x,y)|\langle x,y\rangle \in A\}$ and $f(V,U,Q) = \sum\{\chi_{\Psi_2}(x,y)|\langle x,y\rangle \in B\}$ for suitable $\Psi_1, \Psi_2 \in \Gamma$ and $A, B \subseteq M^i_{n+1} \times M^i_{n+1}$. Since P and Q are compatible we may assume according to (R4) that $\Psi_1 = \Psi_2 = \bar{\Psi}$ are identical. According to (M3), that is, the antisymmetricity of χ, it is sufficient to show that $\langle x,y\rangle \in A$ if $\langle y,x\rangle \in B$. This is however an immediate consequence of the definitions that is of $A = \{\langle x,y\rangle|U = U_x, V = U_y\}$ and $B = \{\langle x,y\rangle|U = U_y, V = U_x\}$

Proof of (10). Let $\bar{\Psi} \in \Gamma$ so that $\bar{\Psi}$ is compatible to P. (According to (R3) there is such a $\bar{\Psi}$). The definition of f implies that $\sum_{V\subseteq H,|V|=k} f(U,V,P) = \sum\{\chi_{\bar{\Psi}}(x,y)|U = U_x, x \in M^i_{n+1}, y \in M^i_{n+1}\}$. According to (M5) and (M6) this sum is 0 if $U_{y_1} \neq U$ and the value of the sum is 1 if $U_{y_1} = U$, that is according to the definition of g, $\sum_{V\subseteq H,|V|=k} f(U,V,P) = g(U)$.

Finally (11) trivially holds because g takes the value 1 exactly once and all of its other values are 0's.

Here we gave a pair of functions f, g with the required properties only on a fixed set H of size m where m was a nonstandard element of M. Assume now that M is an elementary extension of the standard model of Peano arithmetic. Since for all standard m_0 $M \models$ "there exists an $m > m_0$ and H, f, g with the required properties" this is true (for any fixed standard m_0) in the standard model of Peano Arithmetic as well, therefore (5) holds in the standard model. *Q.E.D.*(5)

3. If M is a model of Peano Arithmetic and $n \in M$ then M_n will denote the set $\{x \in M|M \models x < n\}$. Suppose that X is a k-ary relation defined on M where k is a natural number. We say that X is definable in M if there is a first-order formula $\phi(x_1,\ldots,x_k,y)$ of Peano Arithmetic with the free variables $x_1,\ldots,x_k,y$ and there is a $c \in M$ so that for all $x_1,\ldots,x_k \in M$ we have $X(x_1,\ldots,x_k)$ iff $M \models \phi(x_1,\ldots,x_k,c)$. If X is a k-ary relation on M_n then there exists a single

first-order formula $\phi(x_1, \ldots x_k, y)$ (which does not depend on X) so that if X is defined on M_n only then there exists a $c_X \in M$ so that for all $x_1, \ldots, x_k \in M_n$ we have $X(x_1, \ldots, x_k, c_X)$ iff $M \models \phi(x_1, \ldots, x_2, c_g)$. We will suppose that for each g a c_g is fixed (e.g. the smallest one with the required properties). This makes it possible to treat the relations on M_n as elements of M.

Assume that $\mathcal{M} = \langle M_n, +, \times, \leq, A_1, ..., A_j \rangle$ where M is a countable nonstandard model of Peano Arithmetic and n is a nonstandard integer in M, $+, \times, \leq$ are the arithmetic operations and the usual ordering of M, and $A_1, ..., A_j$ are relations with arities $r_1, ..., r_j$ where $j, r_1, ..., r_j$ are standard. (We may think that $\mathcal{M}$ is the structure defined by the interpretation τ of section 1.) We want to add a new binary relation ρ to this structure. (This will be $\bar{\Psi}$ in the proof of theorem 1.) We will denote by $\mathcal{M}[\rho]$ the structure that we get from $\mathcal{M}$ by adding the relation ρ to it, that is $\mathcal{M}[\rho] = \langle M_n, +, \times, \leq, A_1, ..., A_j, rho \rangle$

If H is a set and P is a partial q partition of H then $\bigcup P$ will denote the union of the set of all classes of H. Assume that P, Q are partial q partitions, $P \subseteq Q$ iff every class of P is also a class of Q. In this case we will say that Q is an extension of P. Since there is a ono-to-one correspondence between partitions and equivalence relations we will use this terminology for equivalence relations too.

We will construct ρ in the following way. We take a partially ordered set, whose elements are definable in M. In the case when we want ρ to be a q partition of M_n the elements of this partially ordered set will be partial q partitions P of M_n, which are defined in M and have the property $M \models |\bigcup P| \leq n - n^{\epsilon}$, for some standard ϵ. We will pick a sequence (outside M) from these partial q partitions, so that ϵ tends to 0 and the latter elements of the sequence are extensions of the earlier ones. Since M_n is countable we will be able to pick this sequence so that the common extension of the partial q partitions in it will be a q partition of the whole M_n.

First we define a partially ordered set. We may think of the elements of this set as approximations of the relation ρ. As in the previous example we will pick a sequence from this partially ordered set and the union of the relations in the sequence will be ρ.

We will frequently deal with sets which are not definable in M but still they are the union of a uniform sequence of definable sets. E.g. the set of all partial q

partitions $P \in M$ of M_n so that $|\bigcup P| \leq n - n^\epsilon$ for some standard ϵ. This situation motivates the following definition.

Definition. If k is a natural number and X is a k-ary relation on M we say that X is ω-definable in M iff there exists a first-order formula $\psi(x_1, \ldots, x_k, y, z)$ of Peano Arithmetic, and a $b \in M$ so that for all $a_1, \ldots, a_k \in M$ we have: $X(a_1, \ldots, a_k)$ iff " there exists a standard natural number y, so that $M \models \psi(a_1, \ldots, a_k, y, b)$". (The standardness of y is the essence of this definition.)

Definiton. Suppose that $\langle \wp, \leq \rangle$ is a partially ordered set, whose elements are binary relations on M_n and the ordering is: $p \leq q$ iff $q \subseteq p$. Assume further that

(a) each element $p \in \wp$ is definable in M,

(b) the set $\wp$ is ω-definable in M,

(c) $\wp$ has a greatest element $1_\wp$, and

(d) $\wp$ has no minimal elements.

We will call such a partially ordered set a notion of forcing.

Remarks. 1. Although the elements of $\wp$ are relations on M_n, we may treat them as elements of M, so requirement (b) is meaningful. (See remark in the definition in the definiton of definability).

2. Since $\wp$ has no minimal elements but it is covered by a set which is finite in M, it cannot be definable in M.

Example. Let $\wp_\epsilon = \{P | P \in M,\ P$ is a partial q partition of $M_n,\ M \models$"$|\bigcup P| \leq [n - n^\epsilon]$". $\wp^\circ = \bigcup_{1/k} \{\wp_{1/k} | k$ is a standard natural number $\}$. $\wp^\circ$ is a notion of forcing.

Definition. Assume that T is a subset of $\wp$, where $\wp$ is an arbitrary notion of forcing. We say that T is dense iff for all $g \in \wp$ there is a $h \in T$ with $h \leq g$. (We will be mainly interested in those dense subsets which are ω-definable in M.)

Example. In $\wp^\circ$ the following sets are ω-definable dense sets. (The sets are not definable in M since $\wp^\circ$ itself is not definable in M)

1. For each fixed standard rational $\delta > 0$, $T_\delta = \wp^\circ - \wp_\delta$.

2. For each fixed $x \in M_n$, $T_x = \{p \in \wp^\circ \mid x \in \bigcup p\}$.

Definition. Let G be a subset of $\wp$, where $\wp$ is an arbitrary notion of forcing. We say that G is $\wp$ generic over M iff the following three conditions are satisfied:

(a) $g \in G$, $h \in \wp$, $g \leq h$ implies $h \in G$,

(b) for all $g, g' \in G$ there is a $h \in G$ with $h \leq g$ and $h \leq g'$,

(c) if T is a dense subset of $\wp$, which is ω-definable in M, then $G \cap T$ is nonempty.

Since M is countable it is possible to pick (outside M) a decreasing sequence $p_1, p_2, \ldots$ form the elements of $\wp$ so that the sequence contains at least one element from every dense subset of $\wp$ which is ω-definable in M. The filter generated by this sequence is a generic subset of $\wp$ over M. In the example with the partial q partitions $f \leq g$ iff f is an extension of g. $p_1, p_2, \ldots$ is the sequence mentioned whose common extension is the required function. This will be also the common extension of all of the functions in the filter G generated by $p_1, p_2, \ldots$.

Example. Assume that G is $\wp^\circ$ generic over M and let $\rho = \bigcup_{p \in G} p$. Clearly ρ is a partial q partition of M_n. We claim that it is actually a q partition of M_n that is $\bigcup \rho = M_n$. Indeed as we have remarked earlier for each fixed $x \in M_n$, $T_x = \{p \in \wp^\circ \mid x \in \bigcup p\}$ is dense and therefore according to the definition of generic sets contains an element from G. Thus we have $x \in \bigcup \rho$.

Definitions. 1. Suppose that $\phi(y_0, \ldots, y_i)$ is a first-order formula of L', $a_0, \ldots, a_i \in M_n$, $g \in \wp$. We say that $g \Vdash \phi(a_0, \ldots, a_i)$ ("g forces $\phi(a_0, \ldots, a_i)$) iff for any generic subset G of $\wp$ with $g \in G$ we have that $\rho = \bigcup G$ implies $\mathcal{M}[\rho] \models \phi(a_0, \ldots, a_i)$.

We will be interested in the properties of those relations on M_n which can be defined by a first-order formula from ρ and the relations given in $\mathcal{M}$.

3. Suppose that i is a natural number and X is a relation on M_n^i. We say that X is in $\mathcal{M}[\rho]$ (or definable in $\mathcal{M}[\rho]$), if there exists a natural number j and a first-order formula $\phi(x_0, \ldots, x_{i-1}, y_0, \ldots, y_{j-1})$ so that for some $b_0, \ldots, b_j \in M_n$ we have that for all $a_0, \ldots, a_{i-1} \in M_n$: $X(a_0, \ldots, a_{i-1})$ iff $\mathcal{M}[\rho] \models \phi(a_0, \ldots, a_{i-1}, b_0, \ldots, b_j)$.

The following lemma has three different assertions which are essentially are more formalized versions of the following ideas:

(a) if G is $\wp^\circ$ generic over M then for each first-order statement X about $\mathcal{M}[\rho]$ there is an element g of G which either forces X or its negation.

(b) the relation $p \Vdash X(a_0, ..., a_{i-1})$ (as a subset of $\{\langle p, a_0, ..., a_{i-1}\rangle | p \in \wp, a_j \in M_n)$ is ω-definable in M, and if we restrict p to $\wp_\epsilon$ then it is definable in M.

(c) asssume that we are interested in the turth values of the relations $X(a_0, ..., a_{i-1})$, $a_0, ..., a_{i-1} \in M_n$. In every generic set there is a p so that for each fixed $a_0, ..., a_{i-1}$ we have to extend p with only k elements, where k is a standard number depending only on i, so that p decides $X(a_0, ..., a_{i-1})$

Lemma 14. *Suppose that i is standard a natural number and X is a relation on M_n^i so that X is in $\mathcal{M}[\rho]$, where $\rho = \bigcup G$ and G is $\wp^\circ$ generic over M, then the following conditions hold:*

*(**15**) for all $a_0, \ldots, a_{i-1} \in M_n$ there is a $g \in G$ so that $g \Vdash X(a_0, \ldots, a_{i-1})$ or $g \Vdash \neg X(a_0, \ldots, a_{i-1})$.*

*(**16**) for each $q \in \wp^\circ$ there is a $q' \in \wp^\circ$, $q' \leq q$ so that the relation $p \Vdash X(a_0, \ldots, a_{i-1})$ resricted to the set $p \leq q'$, $p \in \wp^\circ$, $a_0, \ldots, a_{i-1} \in M_n$ is ω-definable in M, and for any standard rational $\epsilon > 0$ the relation $p \Vdash X(a_0, \ldots, a_{i-1})$ resricted to the set $p \leq q'$, $p \in \wp^\circ_\epsilon$, $a_0, \ldots, a_{i-1} \in M_n$ is definable in M*

*(**17**) for all $q \in \wp^\circ$ there exist a $q' \in \wp^\circ$, $q' \leq q$, a standard natural number k and a function U which is definable in M so that for all $a \in M_n^i$, $U(a)$ is a subset of M_n with k elements, and for all $p \in \wp^\circ$ if $p \leq q'$ and $U(a) \subseteq \bigcup p$, then either $p \Vdash X(a)$ or $p \Vdash \neg X(a)$.*

Remark. Since $\wp^\circ$ is not definable in M the relation $p \Vdash X(a_0, \ldots, a_{i-1})$ is not definable in M. However if we restrict p to a $\wp^\circ_\epsilon$ as described in (16) it will be definable. (17) shows that it is enough to consider this restricted relation since already in such a $\wp^\circ_\epsilon$ we will find an element p with $p \Vdash X(a)$ or $p \neg X(a)$.

(17) means that if a set of formulae is given with parameters in M_n then they can be almost decided simultaneously, that is there is a $p \in G$ so that for each fixed formula there is a set $(U(a))$ containing a standard number of elements so that if we give the classes of ρ there and the fact $p \subseteq \rho$, these together decide already whether the formula is true or false in $\mathcal{M}[\rho]$.

According to the following corollary if $f \in \mathcal{M}[\rho]$ is a first-order definable function from M_n^i into M_n^j where i, j are standard integers, then there exists $p \in G$ so that for each $a \in M_n^i$ there is a set U_a, containing a standard number of elements, so that if we give the classes of ρ there then this (tohether with $p \subseteq \rho$) already decides the value of the function at a.

Corollary 18 . *Suppose that i, j are standard natural numbers and X is a relation on M_n^{i+j} so that X is in $\mathcal{M}[\rho]$, where $\rho = \bigcup G$ and G is $\wp^\circ$ generic over M, and for each $x \in M_n^i$ there is exactly one $y \in M_n^j$ with $X(x,y)$. (That is "$y = f(x)$ iff $X(x,y)$" is a function). Then the following holds:*

(**19**) *for all $q \in G$ there exist a $q' \in G$, $q' \leq q$, a standard natural number k and a function U which is definable in M so that for all $a \in M_n^i$, $U(a)$ is a subset of M_n with k elements, and for all $p \in \wp^\circ$ if $p \leq q'$ and $U(a) \subseteq \bigcup p$, then there exists a $b \in M_n^j$ with $p \Vdash X(a,b)$.*

Now we are able to construct the family Γ with properties (R1)-(R4).

In Lemma 14 and Corollary 18 $\wp^\circ$ consists of partial q-partitions of the set M_n. In the following application of this lemma and its corollary M_{n+1} will take this role that is we apply the lemma with $n \to n+1$.

Assume that G is $\wp^\circ$ generic over M and $\rho = \bigcup G$ and $\mathcal{M}[\rho] \models \neg$"the modulo p counting principle". This means that there is a standard integer r, a set $X \subseteq M_{n+1}^r$ and equivalence relations E_0, E_1 on X, so that E_0 is a p partition of X and E_1 is a p-exceptional partition of X and X, E_0, E_1 are first-order definbale in the structure $\mathcal{M}[\rho]$ by first-order formulae ξ, ϕ_0, ϕ_1.

According to Lemma 14 there is a $p_1 \in G$ so that

(**20**) $p_1 \Vdash$"the equivalence relation defined by ϕ_0 is a p partition of the set defined by ξ and the equivalence relation defined by ϕ_1 is a p-exceptional partition of the same set".

We want to define Γ as $\Gamma = \{\bigcup G | G$ is $\wp^\circ$ generic over M, $p_1 \in G\}$. To be able to prove properties (R1)-(R4) we have to pick p_1 in a way that it forces certain first-order properties of $\rho = \cup G$. Lemma 14 and Corollary 18 will guarantee the existence of such a p_1. (20) already implies that, as we have promised, r and $i = pr + 2$ does not depend on the choice $\bar{\Psi} \in \Gamma$.

For any fixed $\rho = \bigcup G$, $p_1 \in G$ the function χ_ρ with properties (M1)-(M6) is first-order definable in $\mathcal{M}[\rho]$. According to (M3) for each fixed $x_0 \in M_{n+1}^i$ there are at most $2p$ elements $y \in M_{n+1}^i$ so that $\neg(\chi_\rho(x_0, y) = 0 \wedge \chi_\rho(y, x_0) = 0)$. We will call such a y an exceptional element with respect to x_0. Let $j = 4pi$. We define a function f on M_n^i with values in M_n^j. If $x_0 \in M_n^i$, $f(x_0)$ will be a sequence containing all

exceptional elements y and the corresponding values of $\chi(x_0, y)$ resp. $\chi(y, x_0)$. To make f first-order definable we may agree that the various elements y are arranged in lexicographic ordering. Clearly $f(x_0)$ determines all of the values $f(u, v)$, $x_0 \in \{u, v\}$.

Therefore according to Corollary 18 we may assume that p_1 was chosen with the following property: there is a standard positive integer k and a function U in M so that, $p_1 \in \wp^\circ_{1/k}$ for each $a \in M^i_{n+1}$ if $p_2 \leq p_1$ and $U(a) \subseteq \bigcup p_2$ then there is a $b \in M^j_{n+1}$ so that $p_2 \Vdash f(a) = b$. Acording to (16) we may also assume that the relation

(**21**) "$p_2 \Vdash f(a) = b$", $p_2 \in \wp^\circ_{1/(2k)}, a \in M^i_{n+1}, b \in M^j_{n+1}$ is in M.

The described properties of the function U imply that

(**22**) the classes of ρ intersecting $U(a)$ uniqualy determine all of the values $\chi_\rho(u, v)$ where at least one of the elements u, v is identical to a, provided that $\rho = \bigcup G$, $p_1 \in G$ and G is $\wp^\circ$-generic over M.

Let $\Gamma = \{\bigcup G | G$ is $\wp^\circ$ generic over M and $p_1 \in G\}$ and let $D = \bigcup p_1$. We show that Γ, D and k satisfy properties (R1)-(R4).

$p_1 \in M$ and $p_1 \in \wp^\circ_{1/k}$ implies (R1).

(R2) follows from the fact that each $\rho \in \Gamma$ is an extension of p_1.

Assume that P is a partial q-partition with the properties given in (R4). Since $\Psi_D = p_2$ we have that $P \in \wp^\circ$ and therefore for any $\wp^\circ$ generic G with $P \in G$ we have that $\bar{\Psi} = \bigcup G \in \Gamma$ and it is compatible with P.

Finally let $U_x = U(x)$. (21) and (22) implies (R4). *Q.E.D.*((R1)-(R4))

4. Proof of Lemma 14. As we have seen the relation $\rho = \bigcup G$ where G is $\wp^\circ$ generic over M is a q partition of M_n. In the following definitions $\tilde{f}$ will be an arbitrary q partition of M_n. but it will be of interest in the $\tilde{f} = \rho$ case. We will consider such a q partition $\tilde{f}$ as an evaluation of certain Boolean variables. This motivates the following definitions.

Definition. Suppose that D is a finite set. For each unordered pair (a, b), $a \in D$, $b \in D$ let $x_{a,b}$ be a Boolean variable. (That is the variable $x_{a,b}$ and $x_{b,a}$ are identical.)

We will use this definition in the case $D = M_n$.

If $\tilde{f}$ is a q partition of M_n then we may associate with it the following 0,1-evaluation e of the Boolean variables $x_{a,b}$, $a \in M_n$, $b \in M_n$: $e(x_{a,b}) = 1$ iff $\tilde{f}(a,b)$ that is if a and b are in the same class of $\tilde{f}$. (We will also denote this evaluation by val($\tilde{f}$).)

Lemma 14 is an assertion about a first-order formula $\phi(x)$ of the language L'. Suppose that x is fixed. The truth value of ϕ is a function of the partial q partition $\tilde{f}$. It is easy to see that there is a constant depth Boolean formula $\Gamma \in M$ on the variables $x_{a,b}$ whose value at the evaluation e is the same as the truth value of ϕ. (The evaluation e is not in M but since the Boolean formula is of constant depth an evaluation can be defined in the natural way outside M). We will try to replace Γ by a simpler Boolean formula Γ' so that $\Gamma(e) = \Gamma'(e)$ for all of the possible $\tilde{f}$. We will construct Γ' in M but since the evaluation e is not in M, Γ' cannot be any Boolean formula which is equivalent to Γ in M. Still there are possibilities to construct a good Γ'. For example we may apply one of the Boolean identities (commutativity, associativity, distributivity, etc.) to Γ. If Γ' is the new formula what we get this way, clearly $\Gamma(e) = \Gamma'(e)$. Even if we perfom such transformations on a set of disjoint subformulae of Γ still we get a good Γ', or we may perform a finite number of transformations one after the other of this type. (The number of transformations is counted in the world, not in M). To describe these things in a rigorous way first we define formally what is a constant depth, unlimited fan-in Boolean formula, then we define the mentioned operations on them.

Definition. Suppose that X is a set of Boolean variables. We define unlimited fan-in Boolean formulae in the following way. We define the formulae recursively according to their complexity. Let $F_0 = X \cup \{0,1\}$. Suppose that F_{k-1} is already defined. If H is a finite set of natural numbers and h is a function defined on H with values in F_{k-1} then let $\bigvee_{x \in H} h(x)$ and $\bigwedge_{x \in H} h(x)$ be elements of F_k. Moreover if g is an element of F_{k-1} then let both g and $\neg g$ be an element of F_k. We define F_k as the set of all elements that we can get through one of the described ways. $F = \bigcup_{k=0,1,\ldots} F_k$ is the set of unlimited fan-in Boolean formulae with variables in X. In the following Boolean formula will mean always an unlimited fan-in Boolean formula. The depth of a Boolean formula g will be the smallest integer k with $g \in F_k$. We may define the size of the formula by induction on its depth k. For $k = 0$ the

size is 1 and size($\bigwedge_{x\in H} h(x)$) = $\sum_{x\in H}$size($h(x)$) (and similarily for $\bigvee$), moreover size($\neg s$) =size(s) + 1. This definition of the size is not the same as the corresponding notion for Boolean circuits. However if we want only to define constant depth polynomial size circuits/formulae the two notion is the same.

We will call two Boolean formulae equivalent if their value is the same under any 0,1-evaluation of the variables. We will consider Boolean formulae in a nonstandard model M of Peano Arithmetic, whose depth is a standard natural number. For such formulae it is possible to define the value of the formula even for an evaluation which is not in M. It is possible that two such formulae are equivalent in M still there is an evaluation (not in M) so that the corresponding values of the formulae are different. In the following we will define relations in M which will be finer than the equivalence of the formulae, and will have the property that if two formulae are in relation with eachother than their values are the same for any evaluations (not necessarily in M).

Defintion. In the following we give some of the usual Boolean identities for unlimited fan-in formulae. (For our purposes it is important that they are given in the unlimited fan-in form.) Each identity has a dual form that we get by changing the role of the operations $\bigvee$ and $\bigwedge$. Although we will give here only one of the two forms later referring to these identities we will mean both of them.

(B1) If the ranges of the functions h and g coincide then $\bigwedge_{x\in H} h(x) \equiv \bigwedge_{x\in G} g(x)$.

(B2) If $H = \bigcup_{i\in I} H_i$ where $\{H_i\}$ is a family of pairwise disjoint sets, h_i is a function defined on H_i for all $i \in I$, h is the common extension of all h_i to H and $\bigwedge_{x\in H} h(x) \in F$ then

$\bigwedge_{x\in H} h(x) \equiv \bigwedge_{i\in I}(\bigwedge_{x\in H_i} h_i(x))$.

(B4) if $s \in F$ and $\bigwedge_{x\in H} h(x) \in F$ then

$s \vee \bigwedge_{x\in H} h(x) \equiv \bigwedge_{x\in H}(s \vee h(x))$.

(B5) if $\bigwedge_{x\in H} h(x) \in F$ then $\neg \bigwedge_{x\in H} h(x) \equiv \bigvee_{x\in H} \neg h(x)$.

Apart from these identities for unlimited fan-in formulae we will need the usual Boolean identities fixing the role of $0, 1$ and the operation $\neg$.

(B6) if $s \in F$ then $0 \vee s \equiv s$, $0 \wedge s \equiv 0$, $1 \vee s \equiv 1$, $1 \wedge s \equiv s$, $s \vee \neg s \equiv 1$, $s \wedge \neg s \equiv 0$, $\neg\neg s \equiv s$.

As we mentioned before we want to define a relation L between Boolean formulae so that $\Gamma L \Gamma'$ implies $\Gamma(e) = \Gamma'(e)$ for any $\tilde{f}$ where, e is the evaluation corresponding to e. Since $\tilde{f}$ is a partial q partition, there will be Boolean equations between the variables $x_{a,b}$ which do not follow from the general Boolean identities given in (B1)-(B7) still they hold for all of the evaluations of type e.

Definitions. 1. Let $B = B(D)$ denote the set of unlimited fan-in Boolean formulae with the variables $\{x_{u,v}\}$, $u, v \in D$. A $\kappa \in B$ is called a k,q-partition if there is a q partition P of a set $D(\kappa) \subseteq D$ so that $\kappa = \bigwedge_{\langle u,v\rangle \in P} x_{u,v}$ and $|D(\kappa)|/q = k$. We will use the notation $D(\kappa) = \bigcup P$, $P = P_\kappa$, $k = |\kappa|$.

We say that a set $V \subset D$ covers the k,q partition map $\kappa \in B$ if each class of P_κ has an element in V.

Assume that κ, κ' are k,q, resp. k',q partitions. We say that κ and κ' are contradictory if the a partitions P_κ, $P_{\kappa'}$ are not compatible. (That is there is no partial q partition of D which is an extension of both.)

2. We call a formula $h \in B$ a k-disjunction if $h = \bigvee_{\kappa \in K} \kappa$, where each $\kappa \in K$ is a k',q-partition for some $k' \leq k$.

The set V covers the k-disjunction $h = \bigvee_{\kappa \in K} \kappa$, if it covers all $\kappa \in K$.

3. Let Eq be a Boolean formula so that if $\mathrm{Eq}(\mathrm{val}(\tilde{f})) = 1$ then $\tilde{f}$ is an equivalence relation. e.g. Eq can be the conjunction of the following formulae:

(a) $\bigwedge_{a \in D} x_{a,a}$

(b) $\bigwedge_{a,b,c \in D}((x_{a,b} \wedge x_{b,c}) \rightarrow x_{a,c})$

(The formula $\bigwedge_{a,b \in D} x_{a,b} \leftrightarrow x_{b,a}$ is not needed since we have assumed that the variables $x_{a,b}$ and $x_{b,a}$ are identical.)

For each $u \in D$ let $D_u^{(q)}$ be the set of all subsets of D with exactly q elements. Let F_u be the Boolean formula $\bigvee\{x_{u,a_1} \wedge \ldots x_{u,a_q} |\ \{a_1, ..., a_q\} \in D_u^{(q)}\}$, saying that there are q elements that u is in relation all of them.

For each $a = \{a_1, ..., a_q\} \in [D]^k$ let G_a be the formula $(\bigwedge_{i=1}^{q} x_{a_1,a_i}) \rightarrow \bigwedge_{b \in D-a} \neg x_{b,a_1}$. (Meaning that an element can be in relation with at most q other element, including itself.)

Clearly the formula $O(D) \equiv \text{Eq} \wedge \bigwedge_{u \in D} F_u \bigwedge_{a \in [D]^q} G_a$ evaluated according to $\text{val}(\tilde{f})$ will be true iff $\tilde{f}$ is a q-partition of D. Therefore if $|D|$ is not divisible by q the Boolean equation $O(D) = 1$ has no solution.

4. Suppose that $h = \bigvee_{\kappa \in K} \kappa$ is a k-disjunction and V covers h, $|V| = l$. We define an l-disjunction $c(h, V)$. ($c(h, V)$ will act as a complement for h if we restrict our attention to evaluations of the variables which define partial q partition covered by V). Let $N = \{\mu |$ μ is a j, q-partition for some $j \leq l$; μ is covered by V and $\forall \kappa \in K$ μ is contradictory to $\kappa\}$ and

$c(h, V) = \bigvee_{\mu \in N} \mu$.

It is easy to check that if $\tilde{f}$ is a q partition then $\neg h$ and $c(h, V)$ has the same value under the evaluation $\text{val}(f)$. We say that the formulae $\neg h$ and $c(h, V)$ are k-equivalent.

5. If k is a natural number then we define a binary relation L_k between Boolean formulae. We say that $\Gamma L_k \Gamma'$ if there is a set S of pairwise disjoint subformulae of Γ so that if we replace each formula in S by another wich is equivalent to it according to (B1),...,(B6) or by a formula which is k'-equivalent to it for some $k' \leq k$.

If k, r are both natural numbers we define the relation $L_{k,r}$ by $aL_{k,r}b$ iff there exists a sequence $a_0 = a, a_1, \ldots, a_r = b$ so that for all $j = 0, \ldots, r-1$ we have $a_j L_k a_{j+1}$.

7. Suppose now that $|D| = n$, $\epsilon > 0$ and Q is a 0,1 assignment on a subset of X. We say that Q is an ϵ-partial assignment if there is a partial q partition P of D with $\bigcup P = [n - n^{\epsilon}]$ so that Q asssigns a value to a variable $x_{u,v}$ iff either $u \in \bigcup P$ or $v \in \bigcup P$, moreover for all pairs u, v where Q is evaluated we have $Q(x_{u,v}) = 1$ iff $P(u, v)$. We will use the notations $P = \text{part}(Q)$ $Q = \text{val}(P)$ and $\text{set}(Q) = \bigcup P$.

If λ is a Boolean formula then we will denote by λ^Q the Boolean formula that we get from λ if we perform the substitutions prescribed in Q.

Let P be a partial q-partition of D, and let $\epsilon > 0$.

We define a random variable $R = R_{\epsilon}^{(P)}$ which takes its values with uniform distribution on the set of all ϵ-partial assignments Q satisfying the condition that $\text{part}(Q)$ is extension of P.

Theorem 23 *$\forall q, s, d, u, \delta > 0 \exists \epsilon > 0, k, r$ so that for all sufficiently large n if $|D| = n$ and $\phi \in B(D)$ is a Boolean formula of size at most n^s and depth d, P is a partial q partition of D with $|\bigcup P| \leq n - n^\delta$ elements and $R = R_\epsilon^{(P)}$ is the random assignment defined earlier, then with a probability of at least $1 - n^{-u}$ the following holds. There exists a k-disjunction g and a set $V \subset D$ so that g is covered by V, $|V| = k$ and $\phi^R L_{k,r} g$.*

Using Theorem 23 we may complete the proof of Lemma 14. According to the original definiton of $\wp^\circ$ the elements of $\wp^\circ$ are partial q-partitions of M_n.

First we define two relations W_0 W_1. $W_0(p, a_0, \ldots, a_{i-1})$ will imply $p \Vdash X(a)$, $W_1(p, a_0, \ldots, a_{i-1})$ will imply $p \Vdash \neg X(a)$. For each fixed $a \in M_n^i$ let $\phi_a \in B(M_n)$ be the Boolean formula expressing the relation $X(a_0, \ldots, a_1)$. (Since X is definable in $\mathcal{M}[\rho]$ and ρ can be considered as an evaluation of the variables $x_{s,t}, s, t \in M_n$, there is such a formula ϕ_a.) We may assume that each ϕ_a is of depth at most d and size at most n^s, where the standard integers d, s depend only on the size of the first-order formula defining X but not on n or a). We apply Theorem 23 with $u = i + 1$ for each fixed ϕ_a, $a \in M_n^i$. Let $\epsilon > 0, k, r$ be the numbers whose existence is guaranteed by Theorem 23 and let P' be a value of $R_\epsilon^{(P)}$ satisfying the conclusion of Theorem 23 simultaneously for each fixed ϕ_a, $a \in M_n^i$. (Since $u > i$ there is such a q'.) We define a relation W_1 by $W_1(p, a_0, \ldots, a_i)$ iff "there exists a standard j so that $p \in \wp_{1/j}^\circ$ and $\phi^p L_{j,j} 1$". (We get the definition of W_0 if we substitute the last formula by $\phi^p L_{j,j} 0$"). Clealry W_0, W_1 are ω-definable. The conclusion of Theorem 23 implies that W_1 is equivalent to the relation $p \Vdash X(a_0, \ldots, a_1)$ if $p \leq q'$. This implies the first part of (16).

Let $\delta > 0$ be a standard rational. Then, according to Theorem 23 the relation W_1 with $p \leq q'$ restricted to $\wp_\delta^\circ$ is equivalent to "$p \in \wp_\epsilon^\circ$ and $\phi^p L_{k,r} 1$." where k and r may depend only on i and the size of the formula defining X but does not depend on the choice of $a_0, \ldots, a_{i-1}$. That is $p \Vdash X(a_0, \ldots, a_{i-1})$ is indeed definable in M, if $p \leq q'$, $p \in \wp_\epsilon^\circ$.

If we pick $U(a)$ as the set V belonging to ϕ_a then our previous argument shows that (17) holds.

(15) follows from (16). *Q.E.D.*(Lemma 14)

Proof of Corllary 18. Let $y = f(x)$ if $X(x,y)$. For each fixed x $f(x)$ can be considered as a $0,1$-sequence of length $j \log n$. For each fixed $r = 0,1,...,j \log n$ let $X(a,r)$ be the following relation on M_n^{i+1}: if $r \leq j \log n$ then the rth element of the sequence corresponding to $f(x)$ is a 1. Clearly we may assume that $X(a,r)$ is first-order definable. Therfore applying (17) of Lemma 14 with $X \to X(a,r)$, $i \to i+1$ we get a $q' \in G, q' \leq q$ and a function $U(a,r)$ so that if $p \leq q'$, $p \in \wp^\circ$ and $U(a,r) \subseteq \bigcup p$ then $p \Vdash X(a,r)$ or $p \Vdash \neg X(a,r)$. Since the sequence $\langle X(a,r) \rangle | r = 0,...,\log n \rangle$ determines $f(a)$ uniqualy, we have that for some $b \in M_n^j$, $p \Vdash b = f(a)$.*Q.E.D.*(Corollary 18)

5. Proof of Theorem 23 First we show that it is enough to prove the theorem for the special case when g is an s-disjunction. In this case (supposing that there are no identical terms in the disjunction) the size of g is not more than n^{s+q}, so we may drop the condition about the size of the formula. In the the formulation of the result for s disjunctions we may substitute the relation $L_{k,r}$ with a simpler one.

Definition. Suppose that ϕ, ψ are s disjunctions. We say that $\phi \mathcal{L} \psi$ if $\phi = \bigvee_{i \in I} d(i)$, $\psi = \bigvee_{i \in I'} d(i)$, $I' = \{i \in I | \forall j \in I \ d(i) \neq d(j)$ implies that part$(d(i))$ is not an extension of part$(d(j))\}$. (That is we get ψ from ϕ by deleting from ϕ those terms which are not "minimal").

Clearly there are absolute constants k, r so that for all ϕ, ψ $\phi \mathcal{L} \psi$ implies $\phi L_{k,r} \psi$. We will denote the (essentially) unique ψ with $\phi \mathcal{L} \psi$ by $\min(\phi)$. It is easy to see that if Q is an evaluation of the variables then $\min(\phi^Q) = \min((\min(\phi))^Q)$. (More precisely the two formulae are equivalent according to (B1)).

Lemma 24. *$\forall s, u \exists \epsilon > 0, k$ so that for all sufficiently large n if $|D| = n$, and $\phi \in B(D)$ is an s disjunction and $R = R_\epsilon$ is the random ϵ partial assignment, then with a probability of at least $1 - n^{-u}$ we have; there exists a set $V \subset D$ so that $min(\phi^R)$ is covered by V and $|V| \leq k$.*

Lemma 25. *Lemma 24 implies Theorem 23.*

Proof of Lemma 25. Let K_j be the set of formulae of size at most n^j from $B(D)$. For each positive integer let $U^j_{0,l} = U_{0,l}$ be the set of l disjunctions in K_j. Suppose now that $U_{d-1,l}$ is already defined then let $U_{d,l}$ be the set of all formulae from K_j which are either of the form $\bigvee_{x\in H} h(x)$ where $h(x) \in U_{d,l}$ for all $x \in H$ or of the form $\neg h$ where $h \in U_{d,l}$.

Claim 26. *If $g \in K_j$ and g is of depth at most d then there is a g' in $U_{2d,1}$ and there are positive integers k, r depending on only d so that $gL_{k,r}g'$.*

Proof. Using the identities in (B1),...,(B6) we may transform g into a formula which uses only $\bigvee$ and $\neg$ as logical connectives and still its depth is not greater than 2d. A single variable $x_{a,b}$ may be considered as a 1 disjunction.

Now we may prove Lemma 25 by induction on d. We give the proof for $d = 1$. Suppose that $g \in U(1,k)$ if g is of the form $\bigvee h(x)$ then using (B1) and (B2) we may transform g into a formula in $U(0,l)$ so Lemma 24 can be directly applied.

Assume now that h is of the form $\neg\phi$ where $\phi \in U(0,1)$. According to Lemma 24 with high probability we have $\phi^R \mathcal{L} g$ where g is a k disjunction covered by a set V, where $|V| = k$. g is k equivalent to $c(g,V)$ so we have $\phi^R L_{k,r+1} c(g,V)$. *Q.E.D.*(Lemma 25)

Before we start the proof of Lemma 24 we formulate two combinatorial Lemmas which will be repeatedly used throughout the proof. (The proofs of these lemmas is given in [Ajt1].) The first Lemma essentially states that if there is a function defined on a finite set H so that at each point x the value of the function is a small subset of H not containing x, then inside a small random subset H' the function will be almost trivial, that is H' will have only a constant number of points which are contained in a value of the function taken at a point in H'. The second Lemma is a generalization of the first, for functions with more than one variables.

Lemma C. *Suppose that* $0 < \epsilon < 1/2$, $0 < \delta < \epsilon/4$ *and* g *is a fuction defined on the finite set* H *with* n *elements such that* $g(x) \subseteq H$, $|g(x)| \leq |H|^{1-\epsilon}$ *and* $x \notin g(x)$ *for all* $x \in H$. *If* $j < |H|^{\delta}$ *and* H' *is a random subset of* $|H|$ *with* j *elements, then for all* $t > 0$ *we have* $P(|\{y|y \in H' \text{ and } y \in g(x) \text{ for some } x \in H'\}| \geq t) < n^{-c_1 t + c_2}$ *where* $c_1 > 0$ *and* c_1, c_2 *depend only on* ϵ.

Lemma C'. *Suppose that* $0 < \epsilon < 1/2$ *and* k *is a positive integer. Then there exists a* $\delta > 0$ *such that for any finite set* H *if* g *is a function defined on the Cartesian product* $\prod_k H$ *with* $g(x) \subseteq H$, $|g(x)| \leq |H|^{1-\epsilon}$, $g(\langle x_0, \ldots, x_{k-1}\rangle) \cap \{x_0, \ldots, x_{k-1}\} = \emptyset$ *for all* $x = \langle x_0, \ldots, x_{k-1}\rangle \in \prod_k H$, $j \leq [|H|^{\delta}]$ *and* H' *is a random subset of* H *with* j *elements, then for all* $t > 0$ *we have* $P(|\{y \in H' | y \in g(x) \text{ for some } x \in \prod_k H'\}| > t) < n^{-c_1 t + c_2}$ *where* $c_1 > 0$ *and* c_1, c_2 *depend only on* ϵ *and* k.

Now we continue the proof of Lemma 24.

Definition. Suppose R_ϵ is the random ϵ partial assignment and $D' = D - \text{set}(R_\epsilon)$. For each fixed value of D', D' let R'_δ be a δ partial assignment on the universe D'. Let $R_\epsilon \circ R'_\delta$ be the common extension of the partial q partitions R_ϵ, and R'_δ. Each value of $R_\epsilon \circ R'_\delta$ is a partial q partition on D, moreover the distribution of $R_\epsilon \circ R'_\delta$ is the same as the distribution of R_δ that is the random variables R_δ, and $R_\epsilon \circ R'_\delta$ are identical.

We will give the random ϵ partial assignment of Lemma 24 in the form $R_\epsilon = R_{\epsilon(1)} \circ \ldots \circ R_{\epsilon(j)}$ where $\epsilon(j) = \epsilon$ and j depends only on s, u. We will construct a sequence of k_i disjunctions for $i = 1, \ldots, j$: $\phi = \phi_1, \ldots, \phi_j = \min((\phi)^R)$ so that $\phi_t^{R_{\epsilon(t)}} \mathcal{L} \phi_{t+1}$ for $t = 1, \ldots, j-1$. $\phi = \phi_1$ is an arbitrary s disjunction. The next element of the sequence ϕ_2 will be an s disjunction with some additional property. As we will construct the elements of the sequence ϕ_t we will add more and more additional properties apart from being an s disjunction, and at the last step we get the s disjunction $\phi_j = \min((\phi)^R)$ with the property described in Lemma 24. So our proof will consist of several statement of the type:

if an s disjunction ψ has property P1 then with a probability of at least $1 - n^u$ we have $\psi^{R_{\epsilon'}} \mathcal{L} \psi'$ where $\epsilon' > 0$ depends only on s and u and ψ' is an s disjunction with property P2.

For the sake of notational simplicity we formulate these Lemmas on a universe D of size n but we will actually apply them for a universe of size $n^{\epsilon(t)}$.

Defintion. To make the definition of the notion of "property" simpler, let us assume that we will consider only sets D where the elements of D are natural numbers. A property P of s disjunctions will be a binary relation defined on all of the pairs ϕ, k where ϕ is an s disjunction for some D and k is a natural number (In this definition we assume that s and q are fixed).

Examples. 1. In the conclusion of Lemma 24 we defined a property of s disjunctions. Now we will denote by ψ the s disjunction corresponding to the s disjunction $\min((\phi)^R)$ in the lemma. We will denote this property by Π, that is $\Pi(\psi, k)$ iff "there exists a set $V \subseteq D$ so that ψ is covered by V and $|V| \leq k$". If this holds we will say that the weight of ψ is at most k.

2. We will denote by Ω the trivial property which holds for each s disjunction ϕ and natural number k. If we want to emphasize the dependence of Ω on s, q then we will write $\Omega_{s,q}$.

In the sequence ϕ_t described earlier, for each t we will have a property Q_t so that $Q_t(\phi_t, k)$ holds for all sufficiently large k, more precisely for all $k > k_0$ where k_0 depends only on s and u. E.g. the last Q_t will be the property that ϕ_t is of weight at most k. As we are going along the sequence each property will be the consequence of the previous one in a sense given in the following definition.

Definition. Suppose s, q are a positive integers and P, Q are properties of s disjunctions. We will say that property P can be reduced to property Q if the following holds:

$\forall k' \exists \epsilon > 0, k_0, h \in^\omega \omega$ with $\lim_{x \to \infty} h(x) = \infty$ so that $\forall k > k_0$ if n is sufficicently large, $|D| = n$ and ϕ is an s disjunction with $P(\phi, k')$ then with a probability of at least $1 - n^{-h(k)}$ there is an s disjunction g with $\phi^{R_\epsilon} \mathcal{L} g$ and $Q(g, k)$.

An immediate consequence of the definition is the following statement: if P, Q, T are properties of s disjunctions, P can be reduced to Q and Q can be reduced to T, then P can be reduced to T.

Lemma 27. *If $s = 1$ then for all positive integers q the property Ω can be reduced to the property "for each $a \in D$ there exists a $W_a \subseteq [D]^q$ and a $V(a) \subseteq D - \{A\}$ with the following properties:*

(a) $H \in W_a$ *implies* $a \in H$ *for all* $a \in D$

(b) $|V_(a)| \leq k$ *for all* $a \in D$

(c) $V(a) \cap H \neq \emptyset$ *for all* $a \in D$, $H \in W_a$

(d) ϕ *can be written in the form of* $\phi = \phi_1 \vee \phi_2$ *where* ϕ_1 *is a 1 disjunction of wight at most* k *and* $\phi_2 = \bigvee_{a \in D} \bigvee_{H \in W_a} \bigwedge_{u,v \in H} x_{u,v}$*"*.

Proof. If ϕ is a 1 disjunction it can be written in the form $\bigvee_{H \in W'} \bigwedge_{u,v \in H} x_{u,v}$ where $W' \subseteq [D]^q$. We may also write this in the form of $\bigvee_{a \in D} \bigvee_{H' \in W'_a} \bigwedge_{u,v \in H'} x_{u,v}$, where for each $a \in D$, W'_a is a set of subsets of D containing the element a and together with it exactly q elements. Let k be sufficiently large and let $G = \{a \in D| \ |W'_a| \geq n^{q-1-\frac{1}{k}}\}$, $(n = |D|)$.

Case I. $|G| > n^{\frac{2}{k}}$. We claim that with high probability after we perform the substitutions according to R_ϵ, where $\epsilon = \frac{1}{k}$, there will be a $H \in W'$ so that the value of $\bigwedge_{u,v \in H} x_{u,v}$ is 1. So $\min((\phi)^R)$ is covered by the empty set. Indeed assume that we randomize R_ϵ in the following way.

We will pick a sequence $a_1, ..., ai, ...$, $i \leq n^{\frac{2}{k}} + 1$ recursively and for each a_i we pick a class A_i of $R\epsilon$ containing A_i at random. Assume that the elements $a_1, ..., a_i$ and the classes $A_1, ..., A_{i-1}$ has been already chosen, where $i \leq n^{\frac{2}{k}}$. We will pick now the pair a_i, A_i. Let a_i be an element of $G - (A_1 \cup ... \cup A_{i-1})$. (Since $|G| \geq n^{\frac{3}{k}}$, $|A_j| = q < n^{\frac{1}{k}}$, $i - 1 \leq n^{\frac{2}{k}}$ such an a_i always exists.) We pick A_i with uniform distribution from the set of those elements of $G - (A_1 \cup ... \cup A_{i-1})$ which contain a_i. The definition of G implies that with a probability of at least $1 - n^{-\frac{1}{k}}$ we have $A_i \in W$. Therefore the probability of the event that there is at least one $i = 1, ..., [n^{2\epsilon}] + 1$ with $A_i \in W$ is at

least $1-(1-n^{1-\frac{1}{k}})n^{\frac{2}{k}} \geq 1-e^{-n^{\epsilon}} \geq 1-n^{-u}$ (for all fixed u and k if n is sufficiently large.) *Q.E.D.*(Case I.)

Case II. $|G| < n^{\frac{2}{k}}$. If $\epsilon > 0$ is sufficiently small with respect to k, then clearly $P((|D-\text{set}(R_\epsilon)) \cap G| > t) < |D|^{-c_1 t + c_2}$ where $c_1, c_2 > 0$ depend only on ϵ. Therefore if $\psi = \bigvee_{a \in G} \bigvee_{H \in W_a} \bigwedge_{u,v \in H}$ then ψ^{R_ϵ} is a 1 disjunction with weight of $1-n^{-ck}$. Let $\phi_1 = \psi^{(R_\epsilon)}$.

Let $\psi = \bigvee_{a \in D-G} \bigvee_{H \in W_a} \bigwedge_{u,v \in H}$. We want to show that $\phi_2 = \psi_2^{(R_\epsilon)}$ satisfies the condition of the Lemma. Assume that $a \notin G$. We claim that if $\epsilon > 0$ is sufficiently small then the probability p_t that $D-\text{set}(R_\epsilon)$ contains t pairwise disjoint set which are in W_a' is smaller than $1-|D|^{-\epsilon' t}$ where $\epsilon' > 0$ depends only on ϵ. Indeed, assume that $U_1, ...U_t \in W_a$ are pairwise disjoint sets. The probability of $U_1, ..., U_t \subseteq D - \text{set}(R_\epsilon)$ is at most $n^{(\epsilon-1)(q-1)t}$. On the other hand the number of all possible choices for $U_1, ..., U_t$ is at most $n^{(q-1-\frac{1}{k})t}$. (Here we used that $|W_a| < n^{q-1-\frac{1}{k}}$ for all $a \notin G$). Therefore we have that $p_t \leq n^{(\epsilon-1)(q-1)t} n^{(q-1-\frac{1}{k})t} = n^{(-\frac{1}{k}+\epsilon(q-1))t} \leq n^{-\epsilon' t}$. Therefore with a probability higher than $1-n^{-\epsilon' t}$ we have less than t disjoint sets from W_a' contained in $D - \text{set}(R_\epsilon)$. Therefore with a probability higher than $1-n^{\epsilon'' t}$ we may assume that this holds simutaneaously for all $a \notin G$ where $\epsilon'' > 0$ depends only on ϵ'. Suppose that such a value of R_ϵ is fixed. For each fixed $a \notin G$ let $U_1, ..., U_t$ be a maximal subset of W_a containing pairwise disjoint sets from $D - \text{set}(R_\epsilon)$ and let $V(a) = \bigcup_{i=1}^{t'} U(t)$. Clearly $|V(a)| \leq qt$ and $V(a)$ has a nonempty intersection with any element of W_a' contained in $D - \text{set}(R_\epsilon)$. Therfore in the new universe $D' = D - \text{set}(R_\epsilon)$ the formula $\phi^{(R_\epsilon)}$ can be written in the required form.

Lemma 28. *If $s = 1$ then for all positive integers q the the property*

"for each $a \in D$ there exists a $W_a \subseteq [D]^q$ and a $V(a) \subseteq D - \{A\}$ with the following properties:

(a) $H \in W_a$ *implies* $a \in H$ *for all* $a \in D$

(b) $|V(a)| \leq k$ *for all* $a \in D$

(c) $V(a) \cap H \neq \emptyset$ *for all* $a \in D$, $H \in W_a$

(d) ϕ *can be written in the form of* $\phi = \phi_1 \vee \phi_2$ *where* ϕ_1 *is a 1 disjunction of weight at most* k *and* $\phi_2 = \bigvee_{a \in D} \bigvee_{H \in W_a} \bigwedge_{u,v \in H} x_{u,v}$*". can be reduced to the property*

"ϕ is of weight at most k"

Proof. For each fixed $a \in D$ suppose that the set $V(a)$ covers the 1 disjunction $\bigvee_{H \in W_a} \bigwedge_{u,v \in H} x_{u,v}$.

Let us apply Lemma C with $H \to D$ and $f(x) = V(x)$. Let $H' = D-\text{set}(R_\epsilon)$. and $V = \{y \in H' | \exists x \in H'\ y \in f(x)\}$. Clearly V covers the 1 disjunction ϕ^{R_ϵ} and Lemma C implies that the requirement about the size of V is met with sufficiently large probability.

Lemma 24 clearly follows from the following assertion:

Lemma 29. *For all positive integers s, and q, $\Omega_{s,q}$ can be reduced to the property* Π*: "ϕ is of wheight at most k".*

Proof. The case $s = 1$ is an immediate consequence of Lemma 27 and Lemma 28 and the mentioned transitivity property of reducibility.

Now we start the proof of Lemma 29 without any restriction on s. We will proceed in the following way. If ϕ is an s disjunction then we will prove that there is an s disjunction ψ and an $s-1$ disjunction ψ' so that

(**30**) $\quad \phi^{R_\epsilon} \mathcal{L}(\psi \vee \psi')$

and ψ is of wheight at most k. Then applying the inductive (on s) assumption to ψ' we get the required result. Again we will proceed through several steps of type (30).

Definitions. 1. Suppose that ϕ is an s disjunction $\phi = \bigvee_{i \in I} d(i)$ where each $d(i)$ is an s', q partition for some $s' \leq s$ and $\min(\phi) = \bigvee_{i \in I'} d(i)$ for some $I' \subseteq I$. Let $(\phi)_s = \bigvee_{i \in I''} d(i)$ where $I'' = \{i \in I' | \text{part}(d(i))$ is an s, q partition$\}$. (In other words we get $(\psi)_s$ from ψ by keeping only those terms which describe a q partition with exactly s classes and which are not consequences of any terms of smaller size).

2. Assume that P is a property of s disjunctions. We define the property $(P)_s$ by $P_s(\phi, k)$ iff $P((\phi)_s, k)$.

Lemma 31. *For any positive integer s, Ω can be reduced to $(\Pi)_s$.*

We may get Lemma 29 easily from this lemma and the already proven case $s = 1$ by induction on s, using the transitive property of s reducibility.

Proof of Lemma 31. Suppose that $\phi = \bigvee_{i \in I} d(i)$ is an s disjunction. If $X \in [D]^q$ then we will denote by ϕ^X the s disjunction $\bigvee_{i \in I'} d(i)$ where $I' = \{i \in I |\ X$ is a class of part$(d(i))\}$.

We will prove the Lemma by using the following two assertions and the transitivity of reducibility.

(a) Ω can be reduced to Y, where Y is the following property: "for all $B \in [D]^k$ the weight of ϕ^B is at most k.

(b) Y can be reduced to $(\Pi)_s$.

Proof of (a). Suppose that $B \in [D]^k$ is fixed, $\phi^B = \bigvee_{i \in I'} d(i)$. For each fixed $i \in I'$, $d(i)$ is an s', q-partition for some $s' \leq s$. Let $d'(i)$ be the $s' - 1, q$ partition that we get from $d(i)$ by deleting the term $\bigwedge_{u,v \in B} x_{u,v}$. Let $\psi = \bigvee_{i \in I' d'(i)}$. ψ is an $s - 1$ disjunction, so by the inductive hypothesis with high probability there is an $s - 1$ disjunction g so that $\psi^{R_\epsilon} \mathcal{L} g$ and g is covered by a set of size $k - 1$, which implies our assertion.

Proof of (b). We apply Lemma C' with $H \to D$, $k \to q$. The function f is defined in the following way: if $a_1, ..., a_q \in D$ are q distinct elements, then according to Y there is a set V of size of at most k so that V covers $\phi^{\{a_1, ..., a_q\}}$. In this case let $f(\langle a_1, ..., a_q \rangle) = V$. For all other $\langle a_1, ..., a_q \rangle$ let $f(a_1, ..., a_q) = \emptyset$. Let $R = R_\epsilon$, $H' = D-$set(R). Lemma C' holds for H'. Let $X = \{y \in H' |\ y \in f(a_1, ..., a_q)$ for some $a_1, ..., a_q \in H'\}$. According to the Lemma $P(|X| > k) < n^{-c_1 k + c_2}$, where $c_1 > 0$ and c_1, c_2 depends only on ϵ. We claim that X covers $(\phi)_s^{R_\epsilon}$.

Let $\phi = \bigvee_{i \in I} d(i)$, and suppose that for a fixed $i \in I$, we have that B is a class of part$(d(i))$. It is sufficient to prove that if $|\bigcup$part$(d(i)) -$set$(R_\epsilon)| = sq$ then $B \cap X \neq \emptyset$. $s \geq 2$ implies that there is a $B' \in D$ with $B' \neq B$ so that B' is a class of part$(d(i))$. As we have seen X covers $\phi^{B'}$ which implies our statement.

REFERENCES

[Ajt1] M. Ajtai, Firstorder definability on finite structures, to appear in Annals of Pure and Applied Logic 1989.
[Ajt2] M. Ajtai, The complexity of the Pigeonhole Principle 29-th FOCS, 1988, 346-358. (Combinatorica, accepted for publication)
[Ajt3] Symmetric Systems of Linear Equations modulo p
[Ajt4] On the Existence of mod p Cardinality functions, in Feasible Mathematics II, accepted for publication.

[BIKPPW] P. Beame, R. Impagliazzo, J. Krajicek, T.Pitassi, P.Pudlak and A. Woods. Exponential lower bounds for the pigeonhole principle. 24th STOC, 1992, pp 200-220.

[BP] P. Beame, T. Pitassi, An exponential Separation between the matching Principle and the Pigeonhole Principle.

[BPU] Approximation and small depth Frege proofs. SIAM Journal of Computing, Dec. 1992.

[CR] S. Cook and R. Rechkhow, The relative efficiency of propositional proof systems, Journal of Symbolic Logic 44 (1977) (36-50).

[KPW] J.Krajicek, P.Pudlak, A. Woods. Exponential lower bounds to the size of bounded-depth Frege proofs of the pigeonhole principle. 1991.

[PBI] T. Pitassi, P.Beame, R. Impagliazzo. Exponential lower bounds for the pigeonhole principle.

[PW] J. Paris and A. Wilkie, Counting Problems in Bounded Arithmetic, in Methods in Mathematical Logic, Proc. Caracas 1983, Springer-Verlag Lecture Notes in Mathematics no. 1130. Ed: A. Dold and B. Eckman, Springer-Verlag, 1985, pp. 317-340.

[Wi] A. Wilkie, talk presented at the ASL Summer meeting in Manchester, England, 1984.

[Wo] A. Woods, Some problems in logic and number theory and their connections, Ph.D. dissertation, Department of Mathematics, Manchester University, 1981.