SOME PROBLEMS IN LOGIC AND NUMBER THEORY, AND THEIR CONNECTIONS

Alan Robert Woods

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Faculty of Science, Department of Mathematics, 1981.

ACKNOWLEDGEMENTS

I wish to express my gratitude to Jeff Paris who supervised the work and was a source of constant encouragement. Thanks are also due to all the members of the Manchester logic community who provided a stimulating environment with many interesting discussions on a wide range of topics, and especially to my fellow students, among them Jonathan Cox, Costas Dimitracopoulos, Adam Huby, Collin Seadon, Martin Merry, and José Ruiz. In particular several useful conversations related to the subject matter of part of this thesis were had with Costas Dimitracopoulos. I would also like to thank Jeff Paris and Alex Wilkie for sharing with me the knowledge gained from their recent work. In particular this contributed greatly to making my introduction to complexity theory relatively painless.

My sincere thanks go to Beryl Sweeney who typed the majority of the thesis and obviously put a great deal of effort and energy into it. Thanks too to Briget Iveson who typed the remainder.

Acknowledgement must also be made of the financial support received from the University of Manchester in the form of a research studentship in science and a Bishop Harvey Goodwin mathematical scholarship.

Finally I would like to thank my wife Ellisha for her understanding and help.

Alan Woods,
December, 1981.

ABSTRACT

This thesis is primarily concerned with some problems related to both logic and number theory.

In chapter 1, Wilkie's problem,which asks whether the existence of arbitrarily large prime numbers can be proved using only induction on bounded quantifier (Δ_0) formulas together with the usual "algebraic" axioms (in the first order language with $\leqslant,+,.$), is linked with a question of Macintyre about the provability of the pigeon hole principle in the same axiom system. It is shown that if an axiom schema asserting a version of the pigeon hole principle for Δ_0 formulas is added, then it is possible to prove Sylvester's theorem that for $y \geqslant x \geqslant 1$, some number among $y+1, y+2, \ldots, y+x$ has a prime divisor $p > x$. Alternatively, if function symbols corresponding to Grzegorczyk's $\mathcal{E}^2$ functions are added (with their definitions) and induction allowed on bounded formulas involving them, then the pigeon hole principle can be proved for such formulas, and the existence of infinitely many primes follows.

The problem of defining addition and multiplication on the natural numbers N by first order formulas involving the predicate $x \perp y$ ("x and y are coprime") is considered in chapter 2. The predicates $z = x+y$, $z = x.y$ are defined by bounded quantifier formulas involving only $\leqslant$ and $\perp$. As an application, a proof is given that the class of all rudimentary (i.e., Δ_0 definable) sets of positive numbers is contained in the class consisting of the spectra $S_\phi = \{|M| : M \text{ finite}, M \models \phi\}$ of those sentences ϕ having only graphs as models, with equality between these classes if and only if S_ϕ is rudimentary for <u>every</u> first order sentence ϕ. An analogous result holds with partial orderings in place of graphs. (It is noted in Appendix II that if every S_ϕ is rudimentary, then $NP \neq$ co-NP.)

Julia Robinson's question whether $+,\cdot$ are definable by formulas involving only $\perp$ and successor, is shown to be equivalent to an open problem in number theory, namely the conjecture that there is some k such that every $n \in N$ is determined uniquely by the sequence of sets of distinct primes $S_0,S_1,\dots,S_k$, where $S_i = \{p: p|n+i\}$. An unconditional proof is given that the theory of N in this language is undecidable.

In chapter 3 it is deduced from the linear case of Schinzel's hypothesis H that $z = x.y$ is definable in the language with $=,+$ and the predicate "x is a prime", but that the defining formula cannot be existential, since the conjecture provides an algorithm for deciding all existential sentences in this language.

CONTENTS

INTRODUCTION

The guiding philosophy underlying this thesis is the conviction that significant connections between logic and number theory do exist, and that the only way to find them is by actively searching for them.

There are three chapters. Each is essentially independent of the others (except that some definitions and simple results from the early part of chapter 2 are assumed in chapter 3). However several themes run through the work as a whole. One is the important role that the prime divisors of consecutive integers play. Another (particularly in chapters 1 and 2) is the appearance of problems in computational complexity theory intimately interwoven with the at first sight seemingly unrelated questions under consideration.

A third theme - the goal of much of the work on definability and (un)decidability - is the desirability of achieving a fuller understanding of *logical redundancy* in the combined additive and multiplicative structure of the natural numbers, its finite initial segments, and nonstandard models of related axiom systems. Here by "logical redundancy" is meant the ability to recover the whole structure from seemingly (and in certain senses genuinely) weak parts of it, through the use of first order definitions. Certainly thinking about the structure of nonstandard models played a considerable role in the development of the proofs of the definability and decidability theorems in chapters 2 and 3.

The question of to what extent these "redundancy" properties are shared by other finite structures (including "machines"), and the related problem of "measuring" the relative strengths of different finite structure theories, are only touched on here, but in the author's view, they are potentially very important and worthy of intensive study in their own right.

Finally, some connections between logic and number theory have been found, and these will be described in the pages that follow, however more and deeper links no doubt still remain hidden. The author hopes that the present work will challenge its readers to actively seek them.

LOGIC NOTATION

$\wedge$(*and*), $\vee$(*or*), $\neg$(*not*), $\rightarrow$(*implies*) $<=>$(*if and only if*), $\forall$(*for every*), $\exists$(*there exists*).

$\Sigma \vdash \phi$ (ϕ *is syntatically provable from the axioms* Σ.),

$M \models \phi(x_1,\ldots,x_n)[a_1,\ldots,a_m]$ or just $M \models \phi(a_1,\ldots,a_n)$($\phi(x_1,\ldots,x_n)$ *is true in the model* M *at* $x_1 = a_1,\ldots,x_n = a_n$. *Here* $a_1,\ldots,a_n \in M$.).

Note that there are some minor differences in notation between chapters. (For example the Δ_0 formulas of chapter 1 are called bounded $L_{\leq,+,\cdot}$ formulas in chapter 2.)

Other notation:

$[x]$ denotes the largest integer $n \leq x$.

CHAPTER 1 Bounded induction, the pigeon hole principle, and the existence of arbitrarily large prime numbers.

One of the benefits the study of logic can bring to other areas of mathematics is an understanding of which basic axioms are required to derive any given theorem. Although there has been considerable success in some areas (for example in general topology, infinite algebra and set theory with the axiom of choice, and more recently in finite combinatorics with variants of Ramsey's theorem - see Paris and Harrington [1977]), very little of any significance is known about the axiomatic requirements of number theory, where by "number theory" we mean that body of theorems about the natural numbers which is actually studied by number theorists. Even the case of the very basic theorem that there exist arbitrarily large prime numbers, is not properly understood. Wilkie [1977] highlighted the problem by asking whether this can be proved using an axiom system similar to Peano Arithmetic but having the restriction that the induction axiom is available only where the induction hypothesis is a bounded formula. (A formula $\phi(\vec{x})$ in the first order language with primitives $=, \leqslant, +, ., 0, 1$ is bounded if every quantifier in $\phi(\vec{x})$ occurs either in the form $\forall u \leqslant v$ or in the form $\exists u \leqslant v$, where these are abbreviations for $\forall u(u \leqslant v \to \ldots)$ and $\exists u(u \leqslant v \wedge \ldots)$.) The class of all such formulas will be denoted by Δ_0, and for definiteness we will consider the axiom system $I\Delta_0$ consisting of:

$$0 \leqslant 0 \wedge \neg 1 \leqslant 0$$

$$\forall x(x+0 = x \wedge x.0 = 0 \wedge x.1 = x)$$

$$\forall x \forall y(x+1 = y+1 \to x = y)$$

$$\forall x \forall y(x \leqslant y+1 \Longleftrightarrow x \leqslant y \vee x = y+1)$$

$$\forall x \forall y(x+(y+1) = (x+y)+1)$$

$$\forall x \forall y(x.(y+1) = (x.y)+x)$$

$$\forall \vec{x}(\theta(\vec{x},0) \wedge \forall y(\theta(\vec{x},y) \to \theta(\vec{x},y+1)) \to \forall y\theta(\vec{x},y))$$

for all Δ_0 *formulas* $\theta(\vec{x},y)$, together with the usual axioms for predicate calculus with identity.

Wilkie motivated his question by stating that

"An affirmative answer here (which unfortunately seems unlikely) would, we believe, give an essentially new proof of the infinity of primes. This is because all current proofs, as far as we know, introduce functions of exponential growth and it is known that induction on bounded quantifier formulae, while natural and quite strong in many respects, cannot define functions of greater than polynomial growth."

The present author agrees that a proof using only $I\Delta_0$ would probably necessarily involve some new, and possibly powerful, idea (although he certainly does not think that a proof of a negative answer would be unfortunate). However as we shall see, the present lack of a proof has more to do with our inability to prove a certain combinatorial principle in $I\Delta_0$ than with the lack of functions of exponential growth.

The combinatorial principle referred to is a version of:

THE PIGEON HOLE PRINCIPLE: n *pigeons cannot fit, one per hole, into fewer than* n *holes.*

Specifically it will be shown that the existence of arbitrarily large primes can be proved if $I\Delta_0$ is augmented by the axiom schema:

$PHP(\Delta_0)$:

$$\forall x\, \forall y\, (\forall u_1 \leq y\ \forall u_2 \leq y\ \forall v \leq y\ (\theta(\vec{x},y,u_1,v) \wedge \theta(\vec{x},y,u_1,v) \rightarrow u_1 = u_2)$$
$$\wedge\ \forall u \leq y\ \exists v \leq y\ \theta(\vec{x},y,u,v) \rightarrow \forall v \leq y\ \exists u \leq y\ \theta(\vec{x},y,u,v)),$$

for each Δ_0 *formula* $\theta(\vec{x},y,u,v)$.

Admittedly each instance of this schema can be proved if we add to $I\Delta_0$ an additional axiom asserting

$$\forall x\ (2^x \textit{ exists}).$$

However, $I\Delta_0 + PHP(\Delta_0) \not\vdash \forall x(2^x$ *exists*), as can easily be seen by taking a nonstandard model $\langle M,\leq,+,\cdot\rangle$ of the theory of the natural numbers $\langle N,\leq,+,\cdot\rangle$ and considering the model $\langle A,\leq,+,\cdot\rangle \models I\Delta_0 + PHP(\Delta_0) + \neg\forall x\ (2^x$ *exists*) formed by choosing $a \in M\setminus N$ and letting $A = \{b \in M\colon \exists n \in N\ (b < a^n)\}$.

With the axiom system $I\Delta_0 + PHP(\Delta_0)$ it is in fact possible to prove:

SYLVESTER'S THEOREM: *If* $1 \leq x \leq y$ *then some number among* $y+1, y+2, \ldots, y+x$ *has a prime divisor* $p > x$.

The proof will be given in detail for the case $y \geq x^5$ since this case establishes the existence of arbitrarily large primes in $I\Delta_0 + PHP(\Delta_0)$ in a relatively simple way, while still requiring many of the essential ideas. §4 contains a sketch of how the method may be extended to the general case, essentially by recasting a variant of Sylvester's original argument as a proof in $I\Delta_0 + PHP(\Delta_0)$. Taking $y = x$ this case can be seen to include Chebychev's celebrated theorem:

BERTRAND'S POSTULATE: *For every* $x \geq 1$ *there is a prime* p *satisfying* $x < p \leq 2x$.

Presumably any of the standard "textbook" variants of the proof of Bertrand's Postulate can be similarly recast (with more or less difficulty).

As the reader may have surmised, it is not known whether $PHP(\Delta_0)$ can be proved from $I\Delta_0$. The question of whether the pigeon hole principle can be proved by bounded induction seems to have first been formulated by Angus Macintyre, who was led to it by the problem of proving that for each odd prime p there exist quadratic nonresidues modulo p, that is, residue classes $a(\text{mod } p)$ such that

$$\forall x(x^2 \not\equiv a\ (\text{mod } p)).$$

As this problem raises some interesting (and relevant) matters it is worth digressing for a moment to consider it.

As can easily be proved in $I\Delta_0$, for each prime p the $(p-1)/2$ numbers $1^2, 2^2, \ldots, ((p-1)/2)^2$ are incongruent (mod p) and include a representative of each quadratic residue a (mod p), that is, of each residue class $a \not\equiv 0 \pmod p$ for which

$$\exists x(x^2 \equiv a \pmod p).$$

Thus (using the notation $[1,n] = \{1,2,\ldots,n\}$) the function $f(x) = \min\{y: y \equiv x^2 \pmod p\}$ is a one to one map from $[1,(p-1)/2]$ onto those elements of $[1,p-1]$ which are quadratic residues. Obviously this is contrary to $PHP(\Delta_0)$ if there are no quadratic nonresidues. Also the stronger result that there exist $(p-1)/2$ quadratic nonresidues follows immediately using the following combinatorial principle, in the statement of which $f: A \leftrightarrow B$ means that f is a one to one map from the set A onto the set B.

BIJECTION EXTENSION PRINCIPLE: *If* $A,B \subseteq [1,n]$ *and* $f: A \leftrightarrow B$ *then there is an extension* f^* *of* f *such that* $f^*: [1,n] \leftrightarrow [1,n]$.

(More generally we could replace $[1,n]$ by a finite set C .)

Clearly for the function f considered above,

$$f^*: [(p-1)/2+1, p-1] \leftrightarrow \{a \in [1,p-1]: \forall x(x^2 \not\equiv a \pmod p)\}$$

(or rather the restriction of f^* to $[(p-1)/2+1, p-1]$ has this property) and hence there is a one to one map from $[1,(p-1)/2]$ onto the quadratic nonresidues.

At first sight the bijection extension principle would appear to be quite powerful, since as far as the author is aware, it is not known whether there is a Δ_0 formula $\theta(x,y,p)$ which (on the natural numbers N) defines for each odd prime p a function $y = g_p(x)$ with

$$g_p: [1,(p-1)/2] \leftrightarrow \{a \in [1,p-1]: \forall x(x^2 \not\equiv a \pmod p)\}.$$

Certainly there is such a predicate $\theta(x,y,p)$ in the class $\mathcal{E}^2_*$

defined by Grzegorczyk [1953], but although every predicate on N which can be defined by a Δ_0 formula is in $\mathcal{E}^2_*$, it is not known whether these classes are equal, that is, whether $\Delta_0 = \mathcal{E}^2_*$. (Ritchie [1963] showed that for N, $\mathcal{E}^2_*$ is identical with the class of predicates whose characteristic functions can be computed by deterministic Turing machines working in linear space, and in fact there is a function $y = g_p(x)$ satisfying the above which can be computed <u>nondeterministically</u> in linear space and polynomial time (where these are measured, as usual, with respect to the number of digits in x and p), but even in this case it is not known whether all such functions have Δ_0 definable graphs.)

$\mathcal{E}^2_*$ is defined to be the class of all predicates having $\mathcal{E}^2$ characteristic functions, where a function is in $\mathcal{E}^2$ if it can be <u>defined</u> by the satisfaction of a finite sequence of the following schemes:

(1) $y = 0$, $z = 1$, $z = x + y$, $z = x.y$ are in $\mathcal{E}^2$.

(2) Substitution: *If* $z = f(x_1,\dots,x_k,\dots,x_m)$ *and* $z = g(y_1,\dots,y_n)$ *are in* $\mathcal{E}^2$ *then* $z = f(x_1,\dots,x_{k-1}, g(y_1,\dots,y_m), x_{k+1},\dots,x_m)$ *is in* $\mathcal{E}^2$.

(3) Limited recursion: If $z = g(\vec{y})$, $z = h(x,\vec{y},w)$ *and* $z = b(x,\vec{y})$ *are in* $\mathcal{E}^2$ *then so also is the function* $z = f(x,\vec{y})$ *satisfying*:

(i) $f(0,\vec{y}) = g(\vec{y})$

(ii) $f(x+1,\vec{y}) = h(x,\vec{y},f(x,y))$

(iii) $f(x,\vec{y}) \leq b(x,\vec{y})$.

(Note that it is easy to prove by induction on the complexity of definition, that every $\mathcal{E}^2$ function has at most polynomial growth.)

Now suppose we associate a function symbol f (say) with each $\mathcal{E}^2$ definition of a function, and turn each such definition into an axiom DEF(f). If we add these axioms to $I\Delta_0$ and allow induction on bounded formulas $\theta(\vec{x},y)$ involving the new function symbols as well as $\leq,+,.$ we obtain a new axiom system which will be denoted by $I\mathcal{E}^2_*$. Similarly we can formulate a schema $\mathrm{PHP}(\mathcal{E}^2_*)$ analogous to $\mathrm{PHP}(\Delta_0)$ by allowing $\theta(\vec{x},y,u,v)$ to have these new function symbols, and ask the obvious question whether $I\mathcal{E}^2_* \vdash \mathrm{PHP}(\mathcal{E}^2_*)$. The answer is affirmative and of course it follows that $I\mathcal{E}^2_* \vdash$ *there exist arbitrarily large prime numbers*. Thus <u>if</u> one could show that every $\mathcal{E}^2$ definition (defining a function $y = g(\vec{x})$, say) corresponds to a Δ_0 formula $\phi_g(\vec{x},y)$ in such a way that replacing every function symbol in the axioms DEF(f) using the ϕ_g's yields sentences provable in $I\Delta_0$, then it would follow that $I\Delta_0 \vdash \mathrm{PHP}(\Delta_0)$ and hence $I\Delta_0 \vdash$ *there exist arbitrarily large primes*.

Although it seems unlikely that this approach can succeed, it does suggest analysing $I\mathcal{E}^2_*$ proofs of the existence of arbitrarily large primes to see exactly which $\mathcal{E}^2$ functions are involved.

A standard example of an $\mathcal{E}^2$ function not known to be definable by a bounded formula is $\pi(x)$, the number of primes not exceeding x. (A proof that $y = \pi(x)$ cannot be defined by a bounded formula would help explain why no computationally efficient formulas for $\pi(x)$ or its "inverse" $y = p_n$, the nth prime number, have been found.)

This leads the author to conjecture that there should be a proof of the existence of arbitrarily large primes in which the only $\mathcal{E}^2$ function required is $\pi(x)$, that is that

$I\Delta_0(\pi) + \mathrm{def}(\pi) \models$ *there exist arbitrarily large primes*,

where $I\Delta_0(\pi)$ denotes $I\Delta_0$ with induction allowed on bounded formulas involving $\leq,+,\cdot,\pi$, and $\mathrm{def}(\pi)$ is an axiom asserting that $\pi(0) = 0$ and for every x,

$$\pi(x+1) = \begin{cases} \pi(x)+1, & \text{if } x+1 \text{ is prime,} \\ \pi(x), & \text{otherwise.} \end{cases}$$

Before proceeding further, a few words on the interdependence of the sections of this chapter would seem to be in order. §1 contains a resume of Sylvester's method and its history. This may or may not help the reader understand the proof of the existence of arbitrarily large primes in $I\Delta_0 + PHP(\Delta_0)$ given in §3, which technically depends only on some basic properties of $I\Delta_0$ dealt with in §2. It is therefore possible to omit §1 and go straight on to §2 and §3, coming back only to provide the background for the $I\Delta_0 + PHP(\Delta_0)$ proof of Sylvester's theorem sketched in §4, or "to see where it all comes from". §5 is devoted to a discussion of census functions with the theorem that $I\mathcal{E}^2_* \vdash PHP(\mathcal{E}^2_*)$ as a corollary. (§5 is independent of all other sections.) §6 contains the proof of a lemma about the sums which can be defined by Δ_0 formulas in models of $I\Delta_0$. This lemma is used in §4 but its proof is deferred in order to be able to make use of the results in §5.

§1. Historical perspective.

The number theoretic idea underlying the $I\Delta_0 + PHP(\Delta_0)$ proof of the existence of arbitrary large primes to be given in §3 comes from Sylvester [1891]. This paper (which admittedly is somewhat eccentric) seems to have been sadly neglected by most 20th century number theorists, even to the extent that this fundamental idea is often credited to a living mathematician. However the principle is clearly stated there in a footnote as follows:

"The author was wandering in an endless maze in his attempts at a general proof of his theorem, until in an auspicious hour when taking a walk on the Banbury road (which leads out of Oxford) the Law of Ademption flashed upon his brain: meaning thereby the law (the nerve, so to say, of the preceeding investigation) that *if all the terms of a natural arithmetical series be increased by the same quantity so as to form a second such series, no prime number can enter in a higher power as a factor of the product of the terms in this latter series, when a suitable term has been* taken away *from it, then the highest power which enters as a factor into the product of the terms of the original series*."

Sylvester goes on to explain:

"The whole matter is thus made to rest on ... (Tschebyscheff's) ... superior limit to the sum of the logarithms of the primes not exceeding a given number, from which ... a superior limit may be deduced to the number of such primes."

To understand this in the case of the arithmetic progressions $1,2,\ldots,x$ and $y+1, y+2,\ldots,y+x$, let $[z]_p$ denote the largest power of the prime p which divides z, and $\{z\}_p$ denote the exponent of p in z, so

$$z = \prod_{p|z} [z]_p = \prod_{p|z} p^{\{z\}_p} .$$

Now consider the products:

$$\prod_{1\leqslant s\leqslant x} s = \prod_{p\leqslant x}\ \prod_{1\leqslant s\leqslant x} [s]_p, \quad \prod_{1\leqslant i\leqslant x} (y+i) = \prod_{p\leqslant y+x}\ \prod_{1\leqslant i\leqslant x} [y+i]_p .$$

Sylvester's crucial observation was that for each prime p,

$$w_p \prod_{1\leqslant s\leqslant x} [s]_p \geqslant \prod_{1\leqslant i\leqslant x} [y+i]_p \tag{1}$$

where w_p is the largest single factor on the right hand side of this inequality (and thus will not occur if the corresponding term $y+m$ is deleted from the arithmetic progression).

This was proved using the fact that

$$\{ \prod_{1\leqslant s\leqslant x} s \}_p = \sum_{k\geqslant 1} |\{s\leqslant x\colon\ p^k|s\}| = \sum_{k\geqslant 1} [x/p^k]. \tag{2}$$

By the maximality of w_p, if $w_p|y+m$ then for all k,

$$p^k|y+m+j \iff p^k|j \quad (\text{for } 1 \leqslant j \leqslant x-m)$$

and $$p^k|y+m-j \iff p^k|j \quad (\text{for } 1 \leqslant j \leqslant m-1).$$

Thus,

$$\{ \prod_{m<i\leqslant x} (y+i)\}_p = \{ \prod_{1\leqslant j\leqslant x-m} j \}_p = \sum_{k\geqslant 1} [(x-m)/p^k]$$

$$\{ \prod_{1\leqslant i<m} (y+i)\}_p = \{ \prod_{1\leqslant j\leqslant m-1} j \}_p = \sum_{k\geqslant 1} [(m-1)/p^k]$$

and therefore

$$\{ \prod_{1\leqslant i\leqslant x} (y+i)\}_p = \{w_p\}_p + \sum_{k\geqslant 1} ([(m-1)/p^k] + [(x-m)/p^k])$$

$$\leqslant \{w_p\}_p + \sum_{k\geqslant 1} [\frac{x}{p^k}] = \{w_p\}_p + \{ \prod_{1\leqslant s\leqslant x} s\}_p$$

which implies (1).

From (1) it follows that if no prime divisor of the numbers $y+1, y+2, \ldots, y+x$ exceeds x then

$$(\prod_{p\leqslant x} w_p).(\prod_{1\leqslant s\leqslant x} s) \geqslant \prod_{1\leqslant i\leqslant x} (y+i) \tag{3}$$

and hence

$$\Big(\prod_{x-\xi(x)<i\leqslant x}(y+i)\Big)\cdot\Big(\prod_{1\leqslant s\leqslant x} s\Big) \geqslant \prod_{1\leqslant i\leqslant x}(y+i) \tag{4}$$

for any function $\xi(x) \geqslant \pi(x)$ (where $\pi(x) = |\{p \leqslant x\colon p \text{ prime}\}|$).

(4) can be rewritten as

$$x!y! \geqslant (x+y-\xi(x))!\ .$$

But Chebychev had already shown how to prove $\pi(x) \leqslant Ax(\log x)^{-1}$ via the bound $\sum\limits_{p\leqslant x} \log p \leqslant Bx$ (where A, B are constants and log denotes the logarithm to base e). Taking x sufficiently large and $\xi(x) = Ax(\log x)^{-1}$ with a "good" value of A $(1 < A < 2 \log 2)$ one can show that (4) fails for $y \geqslant x$ by applying Stirling's asymptotic formula for $x!$, or alternatively just the weak version $\log(x!) = x \log x - x + o(x)$. (The latter can easily be proved by comparing $\sum\limits_{1\leqslant n\leqslant x} \log n$ with $\int_1^x \log t\, dt$. The notation $g(x) = o(f(x))$ means $g(x)/f(x) \to 0$ as $x \to \infty$.) This proves Sylvester's theorem for all but finitely many values of x. (Actually Sylvester used a slightly different argument aimed at making the checking of these cases easier.)

If $y \geqslant x^m$ for m a sufficiently large constant then the same result can be obtained by means of much simpler bounds. For example suppose we take the trivial upper bound $\xi(x) = [x/2]+1 \geqslant \pi(x)$ and simplify (4) to:

$$y^{[x/2]+1}\ x^x \geqslant y^x \tag{5}$$

or in logarithmic form:

$$([x/2]+1)\log y + x \log x \geqslant x \log y. \tag{5a}$$

Clearly (5a) fails for all $y > x^m$, provided $m > 2$ and x is sufficiently large, and thus establishes Sylvester's theorem in this case.

Superficially (that is, as written) the above proofs appear to involve numbers greater than 2^x (as the products in (1) to (4) can have values of this order), and as was noted in the introduction the existence of numbers of this size for all x cannot be proved in $I\Delta_0 + PHP(\Delta_0)$.

However there is still some hope since the logarithmic version of (1) namely:

$$\log w_p + \sum_{1\leqslant s\leqslant x} \{s\}_p \log p \geqslant \sum_{1\leqslant i\leqslant x} \{y+i\}_p \log p \qquad (1a)$$

and also the logarithmic version of (4) involve "quantities" whose values can be bounded by a polynomial in x. This suggests using "*approximate logarithms*", that is functions taking values which are rational numbers in the sense of the model under consideration, and which behave in the way one would expect approximations to logarithms to behave. For the logician it should perhaps be mentioned that the making of approximations which are most naturally viewed at the logarithmic level (because the error can then be written as a term rather than as a factor) is crucial for many arguments in prime number theory. The built in tolerance to errors of the magnitude is the reason we will be able to make do with "approximations" which are not very good. This is important for our work with $I\Delta_0$ because it does not seem to be known even whether the predicate $y = [x \log 2]$ can be defined on N by a Δ_0 formula, although it is in $\mathcal{E}^2_*$. In other words, the approximations to the function $y = \log x$ which are known to be definable by Δ_0 formulas are rather inaccurate. For this and other reasons, the use of approximate logarithms (which is a fairly obvious technique for Peano Arithmetic - see for example Woods [1977]) appears to deserve some care when applied to $I\Delta_0$.

Fortunately however, we will be able to postpone these considerations to §4, since for the special case of Sylvester's theorem with $y > x^m$ (m sufficiently large) the crudeness of the bounds which suffice allows the use of integer valued approximate logarithms which can be obtained relatively easily (provided we use base 2 rather than e).

There is still one further difficulty. The logarithmic version of the proof sketched above involves establishing inequalities by manipulating sums having a large number of terms (interchanging double summations, etc.), and as yet there is no known way of defining the predicate $z = \sum_{i \leqslant x} f(i)$ on N by a Δ_0 formula, given an arbitrary function $y = f(i)$ defined by a Δ_0 formula, let alone a proof that this can be done using only the axioms of $I\Delta_0 + PHP(\Delta_0)$. (For N this is another special case of the $\Delta_0 = \mathcal{E}_*^2$ problem.) We will side step this problem by "unravelling" the inequalities to find the underlying "*comparison map*". This is possible because the inequalities were proved in the first place by comparing terms in some order, and is useful because in the cases considered below the comparison map is Δ_0 definable.

§2. Some functions available in $I\Delta_0$.

Although it is not possible to prove in $I\Delta_0$ that

$$\forall x\,\forall y\,\exists z(z = x^y),$$

it *is* possible to define the predicate $z = x^y$ by a Δ_0 formula which provably satisfies the usual inductive definition whenever x^y does exist.

PROPOSITION 2.1 *There is a* Δ_0 *formula* $z = x^y$ *such that*
$I\Delta_0 \vdash z = x^y$ *is a partial function,* *and*
$I\Delta_0 \vdash \forall x(1 = x^0 \wedge \forall y\,\forall z(z = x^y \rightarrow z.x = x^{y+1}))$.

For a proof see Dimitracopoulos [1980]. Note that any other Δ_0 formula with the same property is $I\Delta_0$ provably equivalent to $z = x^y$. The other properties of exponentiation which could reasonably be expected to hold in $I\Delta_0$ can fairly obviously also be proved so we will use these freely.

In particular we can define by a Δ_0 formula an "inverse" function $y = [\log_x z]$ (which is provably total) by taking y to be the largest number such that $x^y \leq z$. It is very important here to realise that [log] should be considered as a single symbol as it does *not* seem to be at all clear that even $\log_2 z$ can be given a reasonable meaning for an arbitrary model of $I\Delta_0$ (unlike the situation for, say, Peano Arithmetic). Thus notation such as $[\log_2 z]$, although intended to be suggestive, should not be taken too literally.

To make it easier to indicate what techniques are available in $I\Delta_0$ we will now consider a fixed model $M \models I\Delta_0$, although the procedures described are quite uniform and could be handled syntactically. The further down we go in M(that is, towards 0) the more the behaviour of the initial segments is compelled to approach that of the initial segments in models of Peano Arithmetic. (In fact, M will always have an initial segment which is a model of Peano Arithmetic - see for example Lessan

[1978], proposition 4.1.7.) Thus we will be concerned with the "top" section of M.

The ability to code "finite" sequences of numbers by a single number is rather limited in $I\Delta_0$ (since in general it is to be expected that giving a distinct code number to each sequence of 0's and 1's of length x will require code numbers up to at least $2^x - 1$). However the technique can be applied if the length of the sequence and the size of its elements are sufficiently small. By a Δ_0 *sequence* we will mean a function taking a value $u_i \in M$ for each $i \in [1,d]$ (for some $d \in M$) which is defined on M by a Δ_0 formula (possibly involving parameters). A Δ_0 *function* will be a function defined on M (or some obvious subset) by a Δ_0 formula (again possibly with parameters).

LEMMA 2.2 *For each* $n \in N$ *there is a* Δ_0 *function* $g_n(c,i,a)$ *with the property that if* $a \in M$ *and* $u_1,u_2,\ldots,u_d$ *is any* Δ_0 *sequence with* $d \leq [\log_2 a]/([\log_2[\log_2 a]] + 1)$ *and each* $u_i < [\log_2 a]^n$, *then there exists* $c \in M$ *such that*

$$\forall i \in [1,d]\ (u_i = g_n(c,i,a)).$$

Proof:

Let $b = [\log_2 a]^n$. c will be the number having $u_d u_{d-1} \ldots u_1$ as the digits of its representation to base b, so we take

$$g_n(c,i,a) = \left[\frac{c}{b^{i-1}}\right] - b\left[\frac{c}{b^i}\right] .$$

That c exists can be established by using induction on j to prove that for each $j \in [1,d]$ there is a number $c_j < b^j$ with digits $u_j u_{j-1} \ldots u_1$ to base b. Since

$$b^j \leq b^d \leq [\log_2 a]^{n[\log_2 a]/([\log_2[\log_2 a]] + 1)}$$
$$< 2^{n[\log_2 a]} \leq a^n$$

the quantifiers in the induction hypothesis can be bounded by a^n

(which exists in M since $n \in N$ and M is closed under multiplication). Thus the proof only requires $I\Delta_0$.

Of course coding is important not only as a means of quantifying over known Δ_0 sequences but also as a technique for constructing new Δ_0 sequences. In particular the coding technique used in lemma 2.2 enables us to define sums of terms from an arbitrary Δ_0 sequence having considerably more elements than the sequences considered in its statement. (Later in section 6 it will be shown that sums of sequences having much larger elements can also be handled.)

LEMMA 2.3 *If* $u_1,u_2,\ldots,u_d$ *is a* Δ_0 *sequence in* M *with* $d \leqslant [\log_2 a]^k$ *and each* $u_i < [\log_2 a]^n$ *for some* $n,k \in N$, $a \in M$, *then there is a* Δ_0 *sequence (with elements denoted by)* $\sum_{1\leqslant i\leqslant j} u_i$, $j \in [1,d]$, *such that* $\sum_{1\leqslant i\leqslant 1} u_i = u_1$, *and for all* $j \in [1,d-1]$,

$$\sum_{1\leqslant i\leqslant 1} u_i = \sum_{1\leqslant i\leqslant j} u_i + u_{j+1} .$$

Proof:

$$\text{Let } d_0 = [\log_2 a]/([\log_2 a]] + 1).$$

If $d \leqslant d_0$ we simply prove the existence of a number coding a sequence $v_1,v_2,\ldots,v_d$ with $v_1 = u_1$ and $v_{j+1} = v_j + u_{j+1}$ for all $j \in [1,d-1]$. This can be done since if each $u_i \leqslant [\log_2 a]^n$ then each $v_j < [\log_2 a]^{n+1}$.

The result is extended to the case $d \leqslant d_0^k$ for $k = 2^m \in N$ by induction on m using a "divide and conquer" argument. Suppose the lemma has been proved for all Δ_0 sequences $u_1,u_2,\ldots,u_d$ with $d \leqslant d_1$ (where $d_1 \leqslant [\log_2 a]^k$ for some $k \in N$). If $w_1,w_2,\ldots,w_d$ is a Δ_0 sequence of length $d \leqslant d_1^2$ and there is some $n \in N$ such that each $w_i < [\log_2 a]^n$, then $\sum_{1\leqslant i\leqslant j} w_i$ can be defined by:

$$\sum_{1\leqslant i\leqslant j} w_i = \sum_{0\leqslant s<[j/d_1]} \sum_{1\leqslant i\leqslant d_1} w_{sd_1+i} + \sum_{1\leqslant i\leqslant j-d_2} w_{d_2+i}$$

where $d_2 = [j/d_1]d_1$. Clearly the inner and final sums can be handled by the induction hypothesis. So also can the sum $\sum_{0\le s<[j/d_1]}$, since each of its terms is of the form:

$$\sum_{1\le i\le d_1} w_{sd_1+i} < d_1[\log_2 a]^n \le [\log_2 a]^{k+n} ,$$

and there are at most d_1 of them.

The properties $\sum_{1\le i\le 1} w_i = 1$ and $\sum_{1\le i\le j+1} w_i = \sum_{1\le i\le j} w_i + w_{j+1}$ are inherited from the similar properties possessed by the sums with at most d_1 terms.

A different sort of coding is used to give Δ_0 definitions of the functions:

$p_n(x)$ = *the* n *th prime divisor of* x *(in order of magnitude)*

$\nu(x)$ = *the number of (distinct) prime divisors of* x.

LEMMA 2.4 *For every* $x \in M$ *there is some* $c \in M$ *such that*

(i) *If* p *is the least prime divisor of* x *then* $c \equiv 1 \pmod p$.

(ii) *If* $p < q$ *are prime divisors of* x, $c \equiv n \pmod p$, *and there is no prime* $r|x$ *with* $p < r < q$, *then* $c \equiv n+1 \pmod q$.

Proof:

Develop enough of the Chinese remainder theorem in $I\Delta_0$ to prove the existence of c.

Clearly if $n < p$ and $p|x$ then $p = p_n(x) \iff c \equiv n \pmod p$. Also,

$$n = \nu(x) \iff p_n(x) \text{ is the largest prime divisor of } x.$$

It should be remarked that, subject to the limitations imposed by the fact that models need not be closed under functions of faster than polynomial growth, the development of the properties of congruences and primes (up until the problem of the existence of "infinity" many) proceeds quite smoothly in $I\Delta_0$. For example (essentially) the standard proofs that the usual definitions of the primes are equivalent, and that every $x > 1$ has at least one prime divisor, work. Some models of weaker axiom systems (notably the system with induction restricted to quantifier free formulas studied by Wilkie [1977] and Shepherdson [1965]) do not have these properties. (See for example the appendix to Woods [1977].)

§3. The pigeon hole principle and arbitrarily large primes.

We are now in a position to give an $I\Delta_0 + PHP(\Delta_0)$ proof of the existence of arbitrarily large primes. In fact the argument below shows in $I\Delta_0 + PHP(\Delta_0)$, that for all x,y at least one of the numbers $y+1, y+2, \ldots, y+x$ is divisible by some prime $p > x$, provided that $y > x^5$ and $x > C$ where C is some fixed (standard) natural number. The condition $x > C$ can be removed by:

LEMMA 3.1 *Suppose* $M \models I\Delta_0$ *and* $x \in N$. *Then for every* $y \in M$ *with* $y \geq x \geq 1$, *at least one of the numbers* $y+1, y+2, \ldots, y+x$ *has a prime divisor* $p > x$.

Proof:

If $y \in N$ this is just Sylvester's theorem for N, so suppose $y \in M \setminus N$ (that is, y is *nonstandard*). If $\nu(y+i) > x$ for some $i \in [1,x]$ then we are done, while if $\nu(y+i) \leq x$ for all $i \in [1,x]$, then assign to each $y+i$ the least prime p such that $p^e | y+i$ for some nonstandard power p^e. (Some such p must exist since the product of a standard number of standard numbers is standard.) The x primes p obtained in this way are distinct, since otherwise $p^e | (y+i_1) - (y+i_2) = i_1 - i_2 \in N$ for some $p^e \in M \setminus N$. Thus at least one of them satisfies $p > x$.

(This argument is adapted from Grimm [1969].)

THEOREM 3.2 $I\Delta_0 + PHP(\Delta_0) \vdash \forall x \exists p (p > x \wedge p$ *is prime*).

Proof:

We will construct a Δ_0 formula $\theta(u,v,x,y)$ such that it can be proved in $I\Delta_0$ that if $p \leq x$ for all prime divisors p of $y+1, y+2, \ldots, y+x$, then $\{<u,v>: \theta(u,v,x,y)\}$ is the graph of a one-to-one map from $[1,a]$ into $[1,b]$, where

$$a = x[\log_2 y]$$

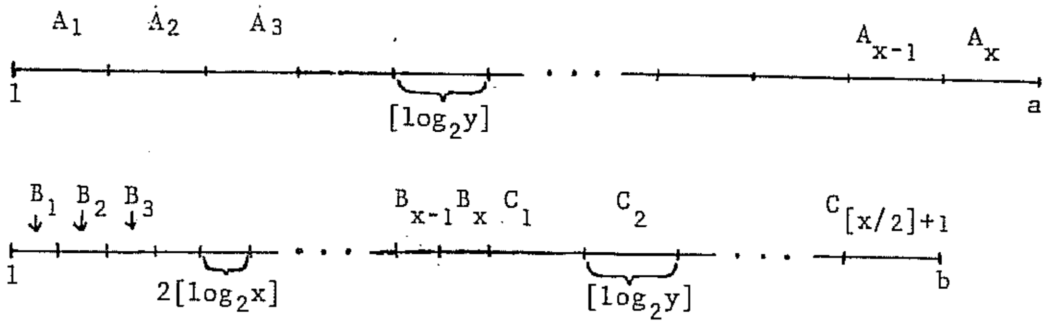
$$b = 2x[\log_2 x] + ([x/2] + 1)[\log_2 y].$$

If $a > b$, this is clearly contrary to $\mathrm{PHP}(\Delta_0)$. But for $y \geq x^5$,

$$(x - [x/2] - 1)[\log_2 y] \geq 5(x - [x/2] - 1)[\log_2 x]$$
$$> 2x[\log_2 x]$$

provided $x > C$, where $C \in N$ is a suitably chosen constant, and therefore $a > b$.

To construct the map from $[1,a]$ to $[1,b]$ we first subdivide $[1,a]$ into x (dijoint) intervals A_i, $i = 1,2,\ldots,x$, each of *length* $|A_i| = [\log_2 y]$, and $[1,b]$ into x intervals B_r, $r = 1,2,\ldots,x$, each of length $|B_r| = 2[\log_2 x]$, followed by $[x/2] + 1$ intervals C_i, $i = 1,2,\ldots,[x/2] + 1$, each of length $|C_i| = [\log_2 y]$.



Each A_i is now subdivided in a manner determined by the prime power decomposition $y + i = \prod_{1\leq j\leq\nu(y+i)} p_j^{e_j}$, the idea being to represent factors $p_j^{e_j}$ of this product by lengths corresponding roughly to their logarithms. (The indexing P_j, $j = 1,2,\ldots,\nu(y+i)$ of the prime divisors of $y + i$ is available in $I\Delta_0$ by lemma 2.4.) A_i is divided into nonempty subintervals A_{ij}, $j = 1,2,\ldots,h$, with

$$|A_{ij}| = \begin{cases} e_j([\log_2 p_j] + 1) & \text{for } j < h, \\ \leq e_h([\log_2 p_h] + 1) & \text{for } j = h. \end{cases}$$

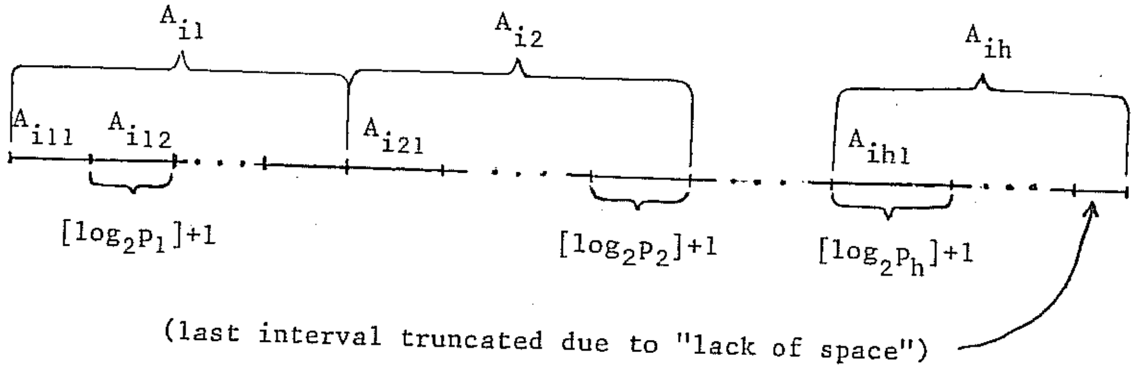
By lemma 2.3 the sum $\sum_{1\leq t\leq j} e_t([\log_2 p_t] + 1)$ makes sense in $I\Delta_0$, and the endpoints of the A_{ij}'s form a Δ_0 sequence. Also it can easily be proved in $I\Delta_0$ that

$$\sum_{1\leq j\leq\nu(y+i)} e_j([\log\ p_j] + 1) > [\log_2 y]$$

so it is possible to "carry out" this construction for some $h \leq \nu(y+i)$.

Each A_{ij} is then further subdivided into subintervals A_{ijk} with $|A_{ijk}| = [\log_2 p_j] + 1$, corresponding to the factors p_j of $p_j^{e_j}$. (In the case of A_{ih} we construct as many disjoint subintervals of length $[\log_2 p_j] + 1$ as there is space for, allowing the final A_{ihk} to have $|A_{ijk}| \leq [\log_2 p_h] + 1$.)

$\underline{A_i}$:



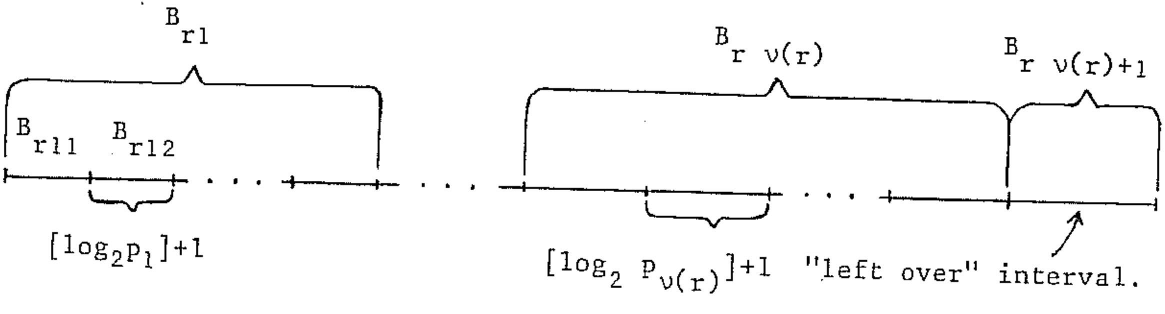
Similarly each B_r is subdivided using the prime power decomposition of $r = \prod_{1 \leq s \leq \nu(r)} p_s^{e_s}$, first into intervals B_{rs}, $s = 1,2,\ldots,\nu(r)$, with $|B_{rs}| = e_s([\log_2 p_s] + 1)$, and a "left over' interval $B_{r\nu(r)+1}$ (possibly empty), and then (for $s \leq \nu(r)$) into subintervals B_{rst}, $t = 1,2,\ldots,e_s$, with $|B_{rst}| = [\log_2 p_s] + 1$. The first of these subdivisions can always be carried out since

$$\sum_{1 \leq s \leq \nu(r)} e_s([\log_2 p_s] + 1) \leq \sum_{1 \leq s \leq \nu(r)} 2e_s[\log_2 p_s]$$

$$\leq 2[\log_2 r] \leq 2[\log_2 x].$$

The elements (if any) of the "left over" interval $B_{r\nu(r)+1}$ will not occur in the range of the function to be constructed.

$\underline{B_r}$:



Notice that each interval $A_{ijk}, B_{rst}(s \leqslant \nu(r))$ is associated with a prime by this construction, so these primes can be used as *labels* on the elements in these intervals. Give labels $2,3,5,\ldots,2i-1,\ldots,2[x/2]+1$ to the elements in the intervals $C_1,C_2,C_3,\ldots,C_i,\ldots,C_{[x/2]+1}$ respectively. Only elements having prime labels $\leqslant x$ will occur in the range and domain of the one-to-one function f which we will now construct. f will preserve labels and can therefore be defined separately for each label $p \leqslant x$.

Fix $p \leqslant x$ and consider the least number i^* with the property:

$$\forall i \in [1,x]\ \forall m(\ p^m|y+i \rightarrow p^m\ y+i^*\ ).$$

Define f on the interval A_{i^*j} with label p to be a one-to-one map into the interval C_i with label p. This can be done uniformly using a Δ_0 formula since we can define the endpoints of the intervals and then map consecutive elements to consecutive elements. There are enough of these in C_i because

$$|A_{i^*j}| \leqslant |A_{i^*}| = [\log_2 y] = |C_i|.$$

To define f on the remaining elements of $[1,a]$ with label p, it suffices to define a one-to-one map taking each $A_{ijk}(i \neq i^*)$ with label p to some B_{rst} with the same label. B_{rst} has enough elements because

$$|A_{ijk}| \leqslant [\log_2 p]+1 = |B_{rst}|.$$

In fact we can map A_{ijk} to some B_{rst} with $t=k$. To see this observe that for any given i,k, if there exists j such that A_{ijk} has label p, then $p^k|y+i$. But then $p^k\Big|y+i^*$ so $p^k\Big|\,|i-i^*|$, and therefore either $i = i^* + zp^k$ for some $z \leqslant [(x-i^*)/p^k]$, or $i = i^* - zp^k$ for some $z \leqslant [i^*/p^k]$. Also

$$\begin{aligned}\exists s(\ B_{rsk}\ \textit{has label}\ \ p) &\iff p^k|r \\ &\iff \exists z \leqslant [x/p^k]\ (\ r = zp^k)\ ,\end{aligned}$$

and $[(x-i^*)/p^k] + [i^*/p^k] \leqslant [x/p^k]$, so there is an obvious one-to-one map from $\{i: i \neq i^* \wedge \exists j(A_{ijk}$ *has label* $p)\}$ into $\{r: \exists s(B_{rsk}$ *has label* $p)\}$.

This completes the Δ_0 definition of a function f mapping those elements of $[1,a]$ which have labels $p \leqslant x$, one-to-one into $[1,b]$. But every element of $[1,a]$ has a label which is a prime divisor of some number among $y+1, y+2, \ldots, y+x$, so if these have no prime divisor greater than x, then f is defined on all of $[1,a]$, which is contrary to $\mathrm{PHP}(\Delta_0)$ if $a > b$.

§4. Approximate logarithms and Sylvester's theorem

To obtain a proof of the general case of Sylvester's theorem using only $I\Delta_0 + PHP(\Delta_0)$ we will need:

(I) more "accurate" approximate logarithms than those used in §3,

(II) some way of getting around the requirement (indicated in §1) for a good upper bound on $\pi(x)$.

Fix $M \models I\Delta_0$, $a \in M \setminus \{0,1,2\}$, $n \in N \setminus \{0\}$, and let $Q^+(M)$ denote the set of nonnegative rational numbers in the sense of M. We will now construct a Δ_0 function $\log^*: \{x \in Q^+(M): x \geqslant 1\} \to Q^+(M)$. More precisely, the construction gives a Δ_0 formula with parameter a, which defines a map taking each ordered pairs $<b,c>$ with $b,c \in M$, $b \geqslant c$, to some $<u,v> \in M \times M$ with $\log^* \frac{b}{c} = \frac{u}{v}$.

The intention is that $\log^* x$ should "behave like" an approximation to log x with error less than a fraction $K[\log_2 a]^{-n}$ of its value for x large, where $K \in N$ is a constant. The definition could be extended to all of $Q^+(M) \setminus \{0\}$ by taking $\log^* \frac{b}{c} = -\log^* \frac{c}{b}$ for $b < c$.

We first define a "Δ_0" function $\log^+ x$ for $x \in Q^+(M)$ with $1 \leqslant x < 4$, by considering a grid of squares "under the graph" of the function $y = \frac{1}{x}$, the sides of the squares being parallel to the axes and of length $h = 1/2^k$, where $k \in M$ satisfies $2^{k-1} < [\log_2 a]^n \leqslant 2^k$. (Recall that $\log x = \int_1^x \frac{1}{t}\, dt$.)

For $1 \leqslant x < 4$ we put:

$$\log^+ x = \left(\sum_{1 \leqslant j \leqslant [(x-1)/h]} \left[\frac{1}{(1+jh)h} \right] \right) . h^2 \ ,$$

that is, we count the number of complete squares of area h^2 which lie entirely in the region under the graph. Since $[(x-1)/h] < 6[\log_2 a]^n$ and $\left[\frac{1}{(1+jh)h}\right] \leqslant 2[\log a]^n$, the sum in brackets can be handled by lemma 2.3.

Heuristically we expect that $\log^+ x$ will behave like an approximation to "log x" accurate to within $C[\log_2 a]^{-n}$ for some $C \varepsilon N$, since the number of squares which lie partially in the region under the graph is $\leqslant C[\log_2 a]^n$ for some $C \varepsilon N$, and these therefore represent an area $\leqslant C[\log_2 a]^{-n}$. What we will actually show in $I\Delta_0$ is that for $1 \leqslant x \leqslant 2$, $1 \leqslant y < 2$,

$$\log^+(x.y) \doteqdot \log^+ x + \log^+ y, \qquad (1)$$

where $\doteqdot$ *means that the two sides differ by at most* $C[\log_2 a]^{-n}$ *for some constant* $C \varepsilon N$ *independent of* x,y,a,M.

Note that it is trivial to verify directly from the definition of $\log^+$ that if $x_1 = u/2^k \leqslant x < (u+1)/2^k$ with $u \varepsilon M$, then $\log^+ x = \log^+ x_1$ and $\log^+(x.y) \doteqdot \log^+(x_1.y)$ (for $1 \leqslant y < 2$). Therefore we may suppose that $x = u/2^k$ for some $u \varepsilon M$, where obviously $u \leqslant 2^{k+1}$ for $1 \leqslant x \leqslant 2$.

Since $\int_1^{xy} \frac{1}{t}\,dt = \int_1^{x} \frac{1}{t}\,dt + \int_x^{xy} \frac{1}{t}\,dt$, proving (1) corresponds to showing

$$\int_x^{xy} \frac{1}{t}\,dt = \int_1^{y} \frac{1}{t_1}\,dt_1$$

in the standard case, which is of course done by putting $t_1 = t/x$. The effect of this change of variable is to transform the square grid (with sides = h) under the graph of $\frac{1}{t}$ into a rectangular grid (with vertical sides = xh and horizontal sides = h/x) under the graph of $\frac{1}{t_1}$.

The area ($\doteqdot \log^+(x.y) - \log^+ x$) represented by the complete rectangles of this sort which lie entirely in the region corresponding to $\int_1^{y} \frac{1}{t_1}\,dt_1$ can be compared with the approximation $\log^+ y$ to $\int_1^{y} \frac{1}{t_1}\,dt_1$ obtained by counting complete squares in a square grid with sides = h, without going beyond $I\Delta_0$. To do this consider a *common refinement* of these square and rectangular grids consisting of a grid of squares with sides $h_1 = h/(u.2^k)$, where $x = u/2^k$. Since $h_1 \geqslant 1/2^{3k+1} > 1/(2^4[\log_2 a]^{3n})$, the number of squares in this common refinement is small enough to enable lemma 2.3 to be used to count the number of these "captured" when the square grid with side h is used and compare this with the number captured when using the rectangular grid. The details of the formal proof are more appropriate to a tedious first course in integral calculus so are omitted.

Now to define $\log^* x$ for all $x \in Q^+(M), x \geqslant 1$, write $x = 2^m.w$ where $m \in M$ and $w = x/2^m \in Q^+(M)$ with $1 \leqslant w < 2$, and put

$$\log x = m.\log^+ 2 + \log^+ w = [\log_2 x].\log^+ 2 + \log^+ w.$$

LEMMA 4.1 $\log^* 1 = 0$, *and for all* $x,y \in Q^+(M), x,y \geqslant 1$,

$$\log^* x + \log^* y \doteqdot \log^*(x.y).$$

(*That is,* $|\log^* x + \log^* y - \log^*(x.y)| < K[\log_2 a]^{-n}$ *for some constant* $K \in N$.)

Proof:

Suppose $x = 2^b.u$, $y = 2^c.v$ where $b,c \in M$ and $1 \leqslant u < 2$, $1 \leqslant v < 2$, so $x.y = 2^{b+c}.u.v$.

If $2 \leqslant u.v < 4$, then

$$\log^*(x.y) = (b+c+1).\log^+ 2 + \log^+ \frac{u.v}{2}.$$

But by (1), $\log^+ 2 + \log^+ \frac{u.v}{2} \doteqdot \log^+(u.v)$, so

$$\log^*(x.y) \doteqdot (b+c).\log^+ 2 + \log^+(u.v).$$

The same is trivially true if $1 \leqslant u.v < 2$, and therefore by (1),

$$\log^*(x.y) \doteqdot (b.\log^+ 2 + \log^+ u) + (c.\log^+ 2 + \log^+ v)$$

$$= \log^* x + \log^* y.$$

Notice also that for every $i \in N$ and every standard rational

$\varepsilon > 0$ there is some $j \in N$ (independent of M) such that $\log^* i$ can be made to approximate $\log k$ to within ε simply by taking $a > j$. (More precisely the cut in the standard rationals determined by $\log^* i$ can be brought to within ε of the real number $\log i$.) *We will assume in future that* a *is large enough to validate any standard arithmetic we do with* $\log^*$.

At this stage it is convenient to state the following improvement of lemma 2.3, the proof of which will be deferred until §6.

<u>LEMMA 4.2</u> *If* $u_1, u_2, \ldots, u_d$ *is a* Δ_0 *sequence of elements of* M *where* $d \leq [\log_2 a]^n$ *and each* $u_i \leq u$ *for some* $a, u \in M$, $n \in N$, *then there is a* Δ_0 *sequence* $\sum_{1 \leq i \leq j} u_i$, $j = 1,2,\ldots,d$ *such that* $\sum_{1 \leq i \leq 1} u_i = u_1$ *and*

$$\sum_{1 \leq i \leq j+1} u_i = \sum_{1 \leq i \leq j} u_i + u_{j+1} \quad \textit{for all} \quad j \in [1,d-1].$$

Lemma 4.2 will enable us to define $\sum_{1 \leq i \leq x} \log^* i$. Consider the points $x \in Q^+(M)$ at which there is a "jump" in the value of the "step function" $\log^* x$. From the definitions of $\log^*$ and $\log^+$ it can be seen that these jumps occur at:

$$1 + \frac{1}{2^k}, \ 1 + \frac{2}{2^k}, \ 1 + \frac{3}{2^k}, \ldots, 2, 2 + \frac{2}{2^k}, \ 2 + \frac{2.2}{2^k}, \ 2 + \frac{3.2}{2^k}, \ldots, 4,$$

$$\ldots, 2^m, \ 2^m + \frac{2^m}{2^k}, \ 2^m + \frac{2.2^m}{2^k}, \ 2^m + \frac{3.2^m}{2^k}, \ldots, 2^{m+1}, \ldots \tag{2}$$

where $m \in M$, and $k \in M$ satisfies $2^{k-1} < [\log_2 a]^n \leq 2^k$. Let $v_1, v_2, \ldots, v_d$ be a list of the points $v \in M$, $v < x$, at which $\log^*(v+1) > \log^* v$. From (2) it is easily seen that:

(i) $v_j = j$ for $j \leq 2^k$

(ii) The numbers $v_j + 1$, $2^k \leq j \leq d$ form a Δ_0 sequence comprised of that part of list (2) which lies between $2^k + 1$ and x.

Therefore $d \leq 2^k([\log_2 x] + 1) \leq 2[\log_2 a]^n.([\log_2 x] + 1)$.

Recalling that each $\log^* v_j = u_j/2^{2k}$ for some $u_j \in M$ we see that we can make the definition:

$$\sum_{1 \leq i \leq x} \log^* i = 2^{-2k} \sum_{1 \leq j \leq d} (v_j - v_{j-1})u_j + (x - v_d)\log^* x,$$

where $v_0 = 0$. The sum on the right can be handled by lemma 4.2, so it is obvious that $\sum_{1 \leq i \leq x} \log^* i$ will have the desired inductive properties.

Next, define a partial function $\exp^* w$ for $Q^+(M)$ by taking $\exp^* w$ to be the least $x \in M$ such that $\log^* x \geq w$. Observe that $\exp^* \log^* x = v_d + 1$.

Now suppose $x \geq 2^k$, $x \in M$, and $\log^*(x+1) > \log^* x$ (or equivalently, $\exp^* \log^*(x+1) = x+1$). Then the distance between $x+1 = v_d + 1$ and $\exp^* \log^* x = v_{d-1} + 1$ is

$$\exp^* \log^*(x+1) - \exp^* \log^* x = 2^{[\log_2 x] - k}, \qquad (3)$$

and from the definitions of $\log^*$ and $\log^+$ it can be seen that

$$\log^*(x+1) - \log^* x = \left[\frac{2^{[\log_2 x] + k}}{x+1}\right] . 2^{-2k},$$

so $2^{[\log_2 x] - k} - (x+1).2^{-2k} < (x+1).(\log\ (x+1) - \log\ x) \leq 2^{[\log_2 x] - k}$.

(4)

We are now able to prove a $\log^+$ analogue of the relation $\sum_{1 \leq i \leq x} \log\ i = x \log\ x - x + o(x)$. By $r = s + O(t)$ we will mean that the terms r,s,t satisfy $|r - s| \leq Kt$ for some constant $K \in N$ (independent of M and the variables occurring in r,s,t).

LEMMA 4.3 *For all* $x > [\log_2 a]^{2n}$,

$$\sum_{1 \leq i \leq x} \log^* i = x \log^* x - x + O\left(\frac{x[\log_2 x]}{[\log_2 a]^n}\right).$$

Proof:

Let $\sigma(x) = (x+1)\log^* x - \exp^* \log^* x$ and let

$$\gamma(x) = \sum_{1\leq i\leq x} \log^* i - \sigma(x).$$

If $x \geq 2^k$ then by (3) and (4),

$$\sigma(x+1) - \sigma(x) = \begin{cases} \log^*(x+1), \text{ if } \log^*(x+1) = \log^* x \\ \log^*(x+1) + O(x.2^{-2k}), \text{ if } \log^*(x+1) > \log^* x. \end{cases}$$

Using this it can be shown induction on $x \geq 2^k$ that

$$\gamma(x) - \gamma(2^k) = O(|\{j\colon 2^k \leq v_j < x\}|.x.2^{-2k})$$

$$= O\left(\frac{x[\log_2 x]}{[\log_2 a]^n}\right).$$

Therefore, $\gamma(x) = O\left(2^k \log^*(2^k) + \dfrac{x[\log_2 x]}{[\log_2 a]^n}\right)$

$$= O\left([\log_2 a]^n[\log_2[\log_2 a]] + \frac{x[\log_2 x]}{[\log_2 a]^n}\right)$$

$$= O\left(\frac{x[\log_2 x]}{[\log\ a]^n}\right) \text{ for } x \geq [\log_2 a]^{2n}.$$

But by (3),

$$\sigma(x) = x \log^* x - x + O([\log_2 x] + 2^{[\log_2 x]-k}),$$

so for all $x \geq [\log_2 a]^{2n}$,

$$\sum_{\leq i\leq x} \log^* i = \sigma(x) + \gamma(x) = x \log^* x - x + O\left(\frac{x[\log_2 x]}{[\log_2 a]^n}\right).$$

Our aim now will be to sketch an argument showing the existence of an $I\Delta_0 + PHP(\Delta_0)$ proof of:

<u>SYLVESTER'S THEOREM</u>: *If* $1 \leq x \leq y$ *then some number among* $y+1, y+2, \ldots, y+x$ *has a prime divisor* $p > x$.

In view of lemma 3.1 we may suppose that $x \geq K$ for any fixed $K \in N$. Also, the method of §3 extends readily to the case $y \geq x^{2-\varepsilon}$, where $\varepsilon > 0$ is any sufficiently small standard rational number. (This will be left to the reader. Hint: use a better, but still trivial, bound on $\pi(x)$, for example $\pi(x) < x/3$, together with more "accurate" approximate logarithms than those used in §3. Actually at the expense of doing some extra number theory for N, any $\varepsilon < 1$ can be shown to work.) *We can therefore assume that* $y + x \geq x^{2-\varepsilon}$.

We now turn to the question of the need for a "good" upper bound on $\pi(x)$ in the case $y + x < x^{2-\varepsilon}$. Before considering $I\Delta_0 + PHP(\Delta_0)$ we will investigate how good an upper bound on $\pi(x)$ is required for the standard proof of Sylvester's theorem for N described in §1, and how this can be obtained. Actually we will work with the closely related function $\theta(x) = \sum_{p \leq x} \log p$ rather than $\pi(x)$ since this will be more convenient later. (It is relatively easy to prove $\pi(x) = \frac{\theta(x)}{\log x} + o(\pi(x))$.)

Under the assumption that $y+1, y+2, \ldots, y+x$ all have no prime divisor $p > x$, inequality (4) of §1 asserts (in logarithmic form):

$$\sum_{p \leq x} \log w_p + \sum_{1 \leq i \leq x} \log i \geq \sum_{y < i \leq y+x} \log i \qquad (5)$$

where w_p is the largest power of p which divides some $y + i$. Since $w_p \leq y + x < x^{2-\varepsilon}$ we see that if $p > x^{1-\varepsilon/2} = x_0$ (say) then either $w_p = p$ or $w_p = 1$. Therefore

$$\sum_{x_0 < p \leq x} \log p + x^{1-\varepsilon/2} \log(x^{2-\varepsilon}) + \sum_{1 \leq i \leq x} \log i \geq \sum_{y < i \leq y+x} \log i$$

so $$\sum_{p \leq x} \log p + \sum_{1 \leq i \leq x} \log i + o(x) \geq \sum_{y < i \leq y+x} \log i \geq \sum_{x < i \leq 2x} \log i. \qquad (6)$$

Using $\sum_{1 \leq i \leq x} \log i = x \log x - x + o(x)$ it follows that:

$$\sum_{p\leqslant x} \log p \geqslant (2 \log 2)x + o(x),$$

so a contradiction will be obtained if it can be shown that

$\sum_{p\leqslant x} \log p \leqslant Ax + o(x)$ for some constant $A < 2 \log 2$.

<u>REMARK 4.4</u> If we do <u>not</u> make the assumption that no prime divisor of $y+1, y+2, \ldots, y+x$ exceeds x then instead of (5) we obtain:

$$\sum_{p\leqslant x+y} \log w_p + \sum_{1\leqslant i\leqslant x} \log i \geqslant \sum_{y<i\leqslant y+x} \log i.$$

Taking $y = x$ it follows from this (as above) that

$$\sum_{p\leqslant x} \log p \geqslant (\log 2)x + o(x).$$

This proves proposition 4.2 of chapter 2 of this thesis.

Sylvester's method can also be used to give an upper bound on $\sum_{p\leqslant x} \log p$. Our starting point is the method used to prove inequality (1) of §1. Recall that this asserted that

$$w_p \prod_{1\leqslant s\leqslant x} [s]_p \geqslant \prod_{1\leqslant i\leqslant x} [y+i]_p.$$

A close examination of the proof shows that we also have:

$$\prod_{1\leqslant s\leqslant x} [s]_p \leqslant \prod_{1\leqslant i\leqslant x} [y+i]_p.$$

Furthermore, if $p > x$ and $p \mid y+i$ for some $i \in [1,x]$, then

$$p \cdot \left(\prod_{1\leqslant s\leqslant x} [s]_p \right) \leqslant \prod_{1\leqslant i\leqslant x} [y+i]_p$$

since each $[s]_p = 1$. Taking $y = x$ we see that

$$\left(\prod_{x<p\leqslant 2x} p \right) \cdot \left(\prod_{1\leqslant s\leqslant x} s \right) \leqslant \prod_{1\leqslant i\leqslant x} (x+i),$$

or in logarithmic form:

$$\sum_{x<p\leqslant x} \log p + \sum_{1\leqslant i\leqslant x} \log i \leqslant \sum_{x<i\leqslant 2x} \log i,$$

from which it can easily be deduced that

$$\sum_{p\leqslant x} \log p \leqslant (2 \log 2)x + o(x).$$

This, of course, is not good enough for our purposes since we require a constant strictly less than $2 \log 2$.

There are several ways to produce the slight improvement required. One approach is to take heed of the implication in the quotation from Sylvester given in §1 that the method applies to arithmetic progressions other than $y+1, y+2, \ldots, y+x$. Indeed if y is odd then all the primes $p \in [y+1, y+2x]$ will lie in the arithmetic progression:

$$y+2, y+4, \ldots, y+2i, \ldots, y+2x,$$

while as before

$$\prod_{1\leqslant s\leqslant x} [s]_p \leqslant \prod_{1\leqslant i\leqslant x} [y+2i]_p \quad \underline{\text{for all primes } p \neq 2}.$$

Thus for $y = x$ (and odd) it follows that

$$\left(\prod_{x<p\leqslant 3x} p\right)\cdot\left(\prod_{1\leqslant s\leqslant x} s\right) \leqslant \left(\prod_{1\leqslant i\leqslant x} (x+2i)\right)\cdot\left[\prod_{1\leqslant s\leqslant x} s\right]_2 .$$

But from equation (2) of §1 we know that

$$\log\left[\prod_{1\leqslant s\leqslant x} s\right]_2 = \sum_{1\leqslant k\leqslant [\log_2 x]} \left[x/2^k\right] \log 2 \leqslant x \log 2,$$

so $$\sum_{x<p\leqslant 3x} \log p + \sum_{1\leqslant i\leqslant x} \log i \leqslant \sum_{1\leqslant i\leqslant x} \log(x+2i) + x \log 2 \qquad (7)$$

But $$\sum_{1\leqslant i\leqslant x} \log(2i-1) = \sum_{1\leqslant i\leqslant 2x} \log i - \sum_{1\leqslant i\leqslant x} \log 2i,$$

$$= \sum_{1\leqslant i\leqslant 2x} \log i - \sum_{1\leqslant i\leqslant x} \log i - x \log 2, \qquad (8)$$

so $$\sum_{1\leqslant i\leqslant x} \log(x+2i) = \sum_{x<i\leqslant 3x} \log i - \sum_{x/2<i\leqslant 3x/2} \log i - x \log 2,$$

Putting $\sigma_1(x) = \sum_{1\leqslant i\leqslant x/3} \log i$ and

$\sigma_2(x) = \sum_{x/3<i\leq x} \log i - \sum_{x/6<i\leq x/2} \log i$, it follows that

$$\sum_{x<p\leq 3x} \log p + \sigma_1(3x) \leq \sigma_2(3x).$$

Hence for arbitrary x,

$$\sum_{x/3^{j+1}<p\leq x/3^j} \log p + \sigma_1(x/3^j) \leq \sigma_2(x/3^j) + O(\log x),$$

and thus

$$\sum_{p\leq x} \log p + \sum_{j\leq [\log_3 x]} \sigma_1(x/3^j) \leq \sum_{j\leq [\log_3 x]} \sigma_2(x/3^j) + O(\log x) \qquad (9)$$

But since $\displaystyle\sum_{1\leq i\leq x} \log i = x \log x - x + o\left(\frac{x}{\log x}\right)$, (10)

$$\sigma_2(x) - \sigma_1(x) = (\tfrac{1}{2} \log 3 + \tfrac{1}{3} \log 2)x + o\left(\frac{x}{\log x}\right), \quad \text{so}$$

$$\sum_{j\leq [\log_3 x]} (\sigma_2(x/3^j) - \sigma_1(x/3^j)) = (\sum_j \frac{1}{3^j})\cdot(\tfrac{1}{2} \log 3 + \tfrac{1}{3} \log 2)x + o(x)$$

$$= \tfrac{3}{2} (\tfrac{1}{2} \log 3 + \tfrac{1}{3} \log 2)x + o(x)$$

Therefore, $\sum_{p\leq x} \log p \leq (\frac{3}{4} \log 3 + \frac{1}{2} \log 2)x + o(x)$, and as the reader may check, $\frac{3}{4} \log 3 + \frac{1}{2} \log 2 < 2 \log 2$. (Note the use of an error bound better than $o(x)$ in (10). This is of course possible since the actual error is $O(\log x)$ by Stirling's formula.)

THEOREM 4.5 $I\Delta_0 + PHP(\Delta_0) \vdash$ *Sylvester's theorem.*

Proof: (Sketch.)

Let $M \models I\Delta_0 + PHP(\Delta_0)$, $a \in M$, and $n \in N$. (We will choose a, n explicitly later.) Define $\log^*$ as above and let σ_1^*, σ_2^* be the $\log^*$ analogues of σ_1, σ_2. Then the "starred" version of (9) is

$$\sum_{p\leq x} \log^* p + \sum_{j\leq [\log_3 x]} \sigma_1^*(x/3^j) \leq \sum_{j\leq [\log_3 x]} \sigma_2^*(x/3^j) + o(x). \qquad (9)^*$$

There is a way of *interpreting* this inequality which is viable in $I\Delta_0$. Observe first that multiplying by 2^k, where as before $2^{k-1} < a \leq 2^k$, converts the inequality into one involving sums of terms of the form $2^{2k}\log^* i$ and $2^{2k}\log^* p$, and by the definition of $\log^*$ all such terms are elements of M (rather than just $Q^+(M)$). For the same reason the sums

$$b = \sum_{j \leqslant [\log_3 x]} 2^{2k}\sigma_1(x/3^j), \quad c = \sum_{j \leqslant [\log_3 x]} 2^{2k}\sigma_2(x/3^j)$$

can be defined using lemma 4.2. For suitably chosen b_1, c_1 approximately equal to b,c, the proof of $(9)^*$ (as indicated for (9) above) can be turned into the construction of *a one-to-one comparison map*:

$$f_1\colon\ [1,b_1] \longrightarrow [1,c_1] \setminus \bigcup_{p \leqslant x} A_p$$

where $\{A_p\colon p$ *prime* $\wedge\ p \leqslant x\}$ *is a set of disjoint subintervals of* $[1,c_1]$ *with lengths* $|A_p| = 2^k\log^* p$ *and* f_1 *is* Δ_0 *definable.*

The construction of f_1 is very similar to the construction of the map f used in §3. $[1,b_1]$ is first subdivided into intervals of length roughly $2^{2k}\sigma_1^*(x/3^j) = \sum_{1 \leqslant i \leqslant x/3^{j+1}} 2^{2k}\log^* i$, then into intervals of length approximately $2^{2k}\log^* i$, and these are then subdivided into intervals of length exactly $2^{2k}\log^* p_m$ (except where we "run out of space") corresponding to the prime factors of $i = \prod_{1 \leqslant m \leqslant \nu(i)} p_m^{e_m}$. Since the starred version of the equation $\log(w.z) = \log w + \log z$ is only approximate, namely $|\log^*(w.z) - \log^* w - \log^* z| < \frac{K}{[\log_2 a]^n}$ for some constant $K \in N$, we will actually start from intervals slightly smaller than $2^{2k}\log^* i$ so as to ensure that we always "use up" all the space available in the interval. Since $\sum_{1 \leqslant m \leqslant \nu(i)} e_m \leqslant [\log_2 x]$ we have

$$\sum_{1 \leqslant m \leqslant \nu(i)} e_m \log^* p_m \geqslant \log^* i - K[\log_2 x].[\log_2 a]^{-n},$$

so it suffices to reduce the length from $2^{2k}\log^* i$ to $2^{2k}\log^* i - 2^{k+1}K[\log_2 x]$, (that is, to use approximately $\log^* i - K[\log_2 x].[\log_2 a]^{-n}$ in place of $\log^* i$.) This reduction corresponds to taking $b_1 = b - 2^{k+1}K[\log_2 x] \sum_{j\leq[\log_3 x]} [x/3^{j+1}]$. Clearly the reduced length subintervals can be"marked out" using a Δ_0 function, and

$$b_1 = 2^{2k}\left(\sum_{j\leq[\log_3 x]} \sigma_1^*(x/3^j) + O\left(\frac{x[\log_2 x]}{[\log_2 a]^n}\right)\right).$$

Similarly $[1,c_1]$ is subdivided into intervals of length either $2^{2k}\log 2$ or approximately $2^{2k}\log^*(x_j + 2i)$, where x_j is an odd number divisible by 3 approximating $x/3^{j+1}$. (These can be marked off by a Δ_0 function since the function $\sum_{1\leq i\leq z}\log^*(2i-1)$ can be defined, either directly in the same way as $\sum_{1\leq i\leq x}\log^* i$, or alternatively via (8) at the expense of introducing an $O(x.[\log_2 a]^{-n})$ error.) This time the intervals of length $2^{2k}\log^*(x_j + 2i)$ will have to be expanded to ensure that there is "enough space" for subintervals of length $2^{2k}\log^* p$ corresponding to <u>all</u> of the prime factors of $x_j + 2i$. However for $x \geq [\log_2 a]^n$ we can still choose:

$$c_1 = 2^{2k}\left(\sum_{j\leq[\log_3 x]} \sigma_2^*(x/3^j) + O\left(\frac{x[\log_2 x]}{[\log_2 a]^n}\right)\right).$$

The function f_1 maps subintervals of $[1,b_1]$ with length $2^{2k}\log^* p$ to subintervals of $[1,c_1]$ with the same length in an almost completely analogous fashion to the function constructed in §3 (with intervals of length $[\log_2 p] + 1$).

In a similar way, under the assumption that no prime $p > x$ is a divisor of $x+1, x+2, \ldots, x+y$, the proof of a starred version of (6), namely:

$$\sum_{1\leq i\leq x}\log^*(x+i) \leq \sum_{1\leq i\leq x}\log^*(y+i) \leq \sum_{p\leq x}\log^* p + \sum_{1\leq i\leq x}\log^* i + O\left(\frac{x}{[\log_2 x]}\right) \tag{6*}$$

can be interpreted as defining a one-to-one Δ_0 comparison map

$$f_2\colon\ [1,b_2] \longrightarrow [1,c_2] \setminus \bigcup_{p\leq x} B_p,$$

where
$$b_2 = 2^{2k}\left(\sum_{1\leq i\leq x} \log^*(x+i) + O\left(\frac{x[\log_2 x]}{[\log_2 a]^n}\right)\right),$$

$$c_1 = 2^{2k}\left(\sum_{1\leq i\leq x} \log^* i + O\left(\frac{x}{[\log_2 x]} + \frac{x[\log_2 x]}{[\log_2 a]^n}\right)\right),$$

and $\{B_p\colon p\ \mathit{prime}\ \wedge\ p \leq x\}$ is a set of disjoint intervals with $B_p \cap [1,c_1] = \phi$ and lengths $|B_p| = 2^{2k}\log^* p$. As with the A_p's there are Δ_0 functions mapping $\{p \leq x\colon p\ \mathit{prime}\}$ to the end points of the B_p's, but note that we are <u>not</u> assuming that the A_p's or B_p's can be Δ_0 indexed by any <u>interval</u> $[1,m]$. Obviously however there <u>is</u> a one-to-one Δ_0 function

$$g\colon \bigcup_{p\leq x} B_p \longrightarrow \bigcup_{p\leq x} A_p$$

with $g\colon B_p \longrightarrow A_p$ for each $p \leq x$,

Now it is possible, in effect, to add inequalities $(6)^*$ and $(9)^*$. More precisely, the inequality:

$$\sum_{j\leq[\log_3 x]} \sigma_1^*(x/3^j) + \sum_{1\leq i\leq x} \log^*(x+i) \tag{11}$$
$$\leq \sum_{j\leq[\log_3 x]} \sigma_2^*(x/3^j) + \sum_{1\leq i\leq x} \log^* i + O\left(\frac{x}{[\log_2 x]} + \frac{x[\log_2 x]}{[\log_2 a]^n}\right)$$

follows by PHP(Δ_0) once we construct a one-to-one Δ_0 function $f\colon [1,b_1+b_2] \longrightarrow [1,c_1+c_2]$ by taking

$$f(x) = \begin{cases} f_1(x) & \text{if } x \in [1,b_1], \\ c_1 + f_2(x-b_1) & \text{if } x \in [b_1+1,\ b_1+b_2] \wedge f_2(x-b_1) \in [1,c_2], \\ g(f_2(x-b_1)) & \text{if } x \in [b_1+1,\ b_1+b_2] \wedge f_2(x-b_1) \in \bigcup_{p\leq x} B_p. \end{cases}$$

But now choose $a = x$, $n = 2$, so that the error term in (11) is

$$O\left(\frac{x[\log_2 x]}{[\log_2 a]^n}\right) = O\left(\frac{x}{[\log_2 x]}\right)$$

which is neglible compared with x. Estimating both sides of (11) using the relation

$$\sum_{1\leq i\leq x} \log^* i = x \log^* x - x + O\left(\frac{x}{[\log_2 x]}\right)$$

obtained from lemma 4.3 (for the same values of a,n) now shows (by arithmetic already done above) that

$(2 \log^* 2 - \frac{3}{4} \log^* 3 - \frac{1}{2} \log^* 2)x \leq C \frac{x}{[\log_2 x]}$ for some $C \in N$

But this is false for all $x > C_0$, where $C_0 \in N$ is a constant, contrary to our earlier observation that x can be assumed larger than any such C_0.

§5. Counting the pigeons - census functions and the pigeon hole principle.

The problem of proving the pigeon hole principle in $I\Delta_0$ suggests introducing a new (undefined) function symbol f and then considering the axiom system $I\Delta_0(f)$ obtained from $I\Delta_0$ by allowing induction on $\Delta_0(f)$ formulas, that is, on bounded formulas in which f as well as $+$ and $\cdot$ may occur. Specifically, let $PHP(f)$ be a sentence saying:

$$\forall n(f \textit{ does \underline{not} map } [1,n] \textit{ one-to-one into } [1,n-1]).$$

Does $I\Delta_0(f) \vdash PHP(f)$?

Obviously an affirmative answer here would imply $I\Delta_0 \vdash PHP(\Delta_0)$, however a negative answer seems more likely in view of a conditional result due to Alex Wilkie based on a conjecture which appears (implicitly) in Cook and Reckhow [1979]. Consider one of the standard axiomatisations of the propositional calculus and a doubly indexed set of propositional variables $A_{ij}, i,j \in N$.

<u>CONJECTURE 5.1</u> *For each* $m \in N$ *there exist arbitrarily large numbers* $n \in N$ *such that every proof in proposition calculus of the tautology*

$$\bigwedge_{i\leq n} \bigvee_{j\leq n-1} A_{ij} \rightarrow \bigvee_{k\leq n-1} \bigvee_{i<j\leq n} (A_{ik} \wedge A_{jk})$$

has more than n^m *symbols.*

In other words, the conjecture states that the time required to write out the shortest proof of this formula for a given value of n cannot be bounded by a polynomial in n (or alternatively by a polynomial in the length of the formula). Note that the formula embodies the pigeon hole principle for $[0,n]$ and is therefore a tautology by virtue of the truth of the principle for N.

<u>PROPOSITION 5.2</u> (Wilkie) *The above conjecture implies*

$$I\Delta_0(f) \nvdash PHP(f).$$

Now consider any predicate $\theta(x,\vec{y})$ in some language for N. The *census function* c_θ for $\theta(x,\vec{y})$ (with respect to x) is defined by $c_\theta(x,\vec{y}) = |\{j \in [1,x]: \theta(j,\vec{y})\}|$, or alternatively (with models other than N in mind) by:

$$c_\theta(0,\vec{y}) = 0$$

$$c_\theta(x+1,\vec{y}) = \begin{cases} c_\theta(x,\vec{y}) + 1 & \text{if } \theta(x+1,\vec{y}) \text{ holds,} \\ c_\theta(x,\vec{y}) & \text{otherwise.} \end{cases}$$

Let $\mathrm{def}(c_\theta)$ be an axiom asserting that this definition is satisfied. In what follows, bounded formulas involving $f,g,+,\cdot,\leq$ will be called $\Delta_0(f,g)$ formulas, and $I\Delta_0(f,g)$ will denote the axiom system similar to $I\Delta_0$ but with induction allowed on $\Delta_0(f,g)$ formulas.

THEOREM 5.3 *Let* $\theta(x,y)$ *be the* $\Delta_0(f)$ *formula* $f(x) \leq y$ *and let* $c_\theta(x,y)$ *be a function symbol for the corresponding census function with respect to* x). *Then*

$I\Delta_0(f,c_\theta) + \mathrm{def}(c_\theta) \vdash \mathrm{PHP}(f)$.

Proof:

Working in $I\Delta_0(f,c_\theta) + \mathrm{def}(c_\theta)$, suppose to the contrary that $f: [1,n] \to [1,n-1]$ and is one-to-one. The idea is to remove $f(1), f(2), \ldots$ from $[1,n-1]$ one step at a time, closing up the gaps so caused as we go, until after $n-2$ steps we are left with a one-to-one map from $[n-1,n]$ into $\{1\}$ which can trivially be proved to be impossible. At the ith step we will have the function f_i^* defined on $[i+1,n]$ by:

$$f_i^*(x) = f(x) - c_\theta(i,f(x)).$$

(Intuitively $f_i^*(x) = f(x) - |\{j \in [1,i]: f(j) \leq f(x)\}|$.)

Clearly the predicate $y = f_i^*(x)$ can be defined by a $\Delta_0(f,c_\theta)$ formula with variables x,y,i, provided that we can show $c_\theta(i,f(x)) \leq f(x)$ for all $i \leq n-2$, $x \in [i+1,n]$. Intuitively we should have $c_\theta(i,y) \leq y$ for all y, since as f is one-to-one we would expect that

$$c_\theta(x,y) = |\{j \in [1,x]: f(j) \leq y\}| = |\{j \in [1,y]: \exists z \in [1,x](f(z) = j)\}|,$$

in other words, that $c_\theta(x,y)$ is the census function with respect to y of the formula $\exists z\, \varepsilon[1,x](f(z) = y)$. We now show that $c_\theta(x,y)$ has the inductive properties defining that census function.

Firstly, $c_\theta(x,0) = 0$ by a trivial induction on x. We will also use induction on x to prove:

$$c_\theta(x,y+1) = c_\theta(x,y) + \begin{cases} 1 & \text{if } \exists z\ \varepsilon\ [1,x](f(z) = y+1), \\ 0 & \text{if } \exists z\ \varepsilon\ [1,x](f(z) \neq y+1). \end{cases}$$

This is trivially true for $x = 0$ since both sides are zero. Suppose the equation holds for x. We want to show

$$c_\theta(x+1,y+1) - c_\theta(x+1,y) = \begin{cases} 1 & \text{if } \exists z\ \varepsilon[1,x+1](f(z) = y+1), \\ 0 & \text{if } \forall z\ \varepsilon[1,x+1\ (f(z) \neq y+1). \end{cases}$$

From the definition of c_θ we know that

$$c_\theta(x+1,y) = c_\theta(x,y) + \begin{cases} 1 & \text{if } f(x+1) \leq y, \\ 0 & \text{if } f(x+1) > y, \end{cases}$$

and

$$c_\theta(x+1,y+1) = c_\theta(x,y+1) + \begin{cases} 1 & \text{if } f(x+1) \leq y+1, \\ 0 & \text{if } f(x+1) > y+1. \end{cases}$$

Therefore

$$c_\theta(x+1,y+1) - c_\theta(x+1,y) = c_\theta(x,y+1) - c_\theta(x,y) + \begin{cases} 1 & \text{if } f(x+1) = y+1, \\ 0 & \text{if } f(x+1) \neq y+1. \end{cases}$$

and hence by the induction hypothesis,

$$c_\theta(x+1,y+1) - c_\theta(x+1,y) = \begin{cases} 1 & \text{if } \exists z\ \varepsilon[1,x](f(z) = y+1) \\ 0 & \text{if } \forall z\ \varepsilon[1,x](f(z) \neq y+1) \end{cases}$$

$$+ \begin{cases} 1 & \text{if } f(x+1) = y+1 \\ 0 & \text{if } f(x+1) \neq y+1. \end{cases}$$

Since f is one-to-one the terms on the right cannot both be nonzero, so

$$c_\theta(x+1,y+1) - c_\theta(x+1,y) \doteq \begin{cases} 1 & \text{if } \exists z \in [1,x+1](f(z) = y+1), \\ 0 & \text{if } \forall z \in [1,x+1](f(z) \neq y+1), \end{cases}$$

as required. Thus $c_\theta(x,y)$ is the census function with respect to y for the formula $\exists z \leq x(f(z) = y)$.

That $c_\theta(x,y) \leq y$ for all x,y now follows by a trivial induction on y. More generally we can show that for all x,y_1,y_2,

$$y_1 \leq y_2 \rightarrow c_\theta(x,y_2) - c_\theta(x,y_1) \leq y_2 - y_1. \qquad (1)$$

(Let $y_2 = y_1 + k$ and use induction on k. Equality holds for $k = 0$, so suppose $c_\theta(x,y_1+k) - c_\theta(x,y_1) \leq k$. Then there $c_\theta(x,y_1+k+1) \leq c_\theta(x,y_1+k) + 1$, it follows that $c_\theta(x,y_1+k+1) - c_\theta(x,y_1) \leq k + 1$.)

Similarly

$$y_1 < y_2 \wedge \forall z \in [1,x](f(z) \neq y_2) \rightarrow c_\theta(x,y_2) - c_\theta(x,y_1) \leq y_2 - y_1.$$

Since f is one-to-one, $\forall z \in [1,i](f(z) \neq f(i+1))$, so it follows that for all i,y,

$$f(i+1) > y \rightarrow f(i+1) - c_\theta(i,f(i+1)) > y - c_\theta(i,y).$$

But by (1),

$$f(i+1) \leq y \rightarrow f(i+1) - c_\theta(i,f(i+1)) \leq y - c_\theta(i,y),$$

so

$$f(i+1) \leq y \iff f(i+1) - c_\theta(i,f(i+1)) \leq y - c_\theta(i,y). \qquad (2)$$

Using this we can now show that f_i^* can be "produced" in the way stated at the beginning, namely that $f_0^* = f$, and on $[i+2,n]$,

$$f_{i+1}^*(x) = \begin{cases} f_i^*(x) - 1 & \text{if } f_i^*(x) \geq f_i^*(i+1), \\ f_i^*(x) & \text{if } f_i^*(x) < f_i^*(i+1). \end{cases} \qquad (3)$$

From the way f_i^* was <u>actually</u> defined, this is equivalent to

$$c_\theta(i+1,f(x)) = \begin{cases} c_\theta(i,f(x))+1 \text{ if } f(x)-c_\theta(i,f(x)) \geq f(i+1)-c_\theta(i,f(i+1)), \\ c_\theta(i,f(x)) \quad \text{otherwise.} \end{cases} \tag{4}$$

But by definition

$$c_\theta(i+1,f(x)) = \begin{cases} c_\theta(i,f(x))+1 \text{ if } f_i(i+1) \leq f(x), \\ c_\theta(i,f(x)) \quad \text{otherwise,} \end{cases}$$

so (4) follows immediately from (2).

To complete the proof we will prove the following hypothesis by induction on $i \leq n-2$:

IH(i): f_i^* *is one-to-one and its range is contained in* $[1,n-i-1]$.

(Recall that the domain of definition of f_i^* is $[i+1,n]$.)

As $f_0^* = f$, IH(0) holds, so suppose IH(i) is satisfied. Since f_i^* is one-to-one, we know from (3) that for all $x \in [i+2,n]$,

$$f_{i+1}^*(x) = \begin{cases} f_i^*(x) & \text{if } f_i^*(x) < f_i^*(i+1), \\ f_i^*(x)-1 & \text{if } f_i^*(x) > f_i^*(i+1), \end{cases}$$

so we see immediately that for $x \in [i+2,n]$, either

$$f_{i+1}^*(x) = f_i^*(x) \geq 1 \quad \text{or} \quad f_{i+1}^*(x) = f_i^*(x)-1 \geq f_i^*(i+1) \geq 1.$$

Also, either $f_{i+1}^*(x) = f_i^*(x) \leq f_i^*(i+1)-1 \leq n-(i+1)-1$, or $f_{i+1}^*(x) = f_i^*(x)-1 \leq n-(i+1)-1$. Thus the range of f_{i+1}^* is contained in $[1,n-(i+1)-1]$.

Now suppose $f_{i+1}^*(x) = f_{i+1}(y)$ for some $x,y \in [i+2,n]$ with $x \neq y$. Since f_i^* is one-to-one we may assume without loss of generality that $f_i^*(x) < f_i^*(y)$. Clearly we must then have $f_{i+1}^*(x) = f_i^*(x) < f_i^*(i+1) \leq f_i^*(y)-1 = f_{i+1}^*(y)$, so $f_{i+1}^*(x) < f_{i+1}^*(y)$ contrary to the assumption that $f_{i+1}^*(x) = f_{i+1}^*(y)$. Thus f_{i+1}^* is one-to-one and IH(i+1) holds.

By IH(n-2), f_{n-2}^* has range $\{1\}$ and f_{n-2}^* is one-to-one. But the domain of f^*_{n-2} is $[n-1,n]$ so this is impossible, completing the proof of PHP(f).

REMARK: A similar argument shows that the theorem remains true if $\theta(x,y)$ is taken to be $\exists z \in [1,y](f(z) = x)$, instead of $f(x) \leq y$.

COROLLARY 5.4 $\quad I\mathcal{E}^2_* \vdash \mathrm{PHP}(\mathcal{E}^2_*)$

Proof:

Using essentially the standard argument (see Grzegorczyk [1953]) it can be shown that for each bounded formula $\phi(\vec{x})$ involving (function symbols for) $\mathcal{E}^2$ functions there is an $\mathcal{E}^2$ function χ_ϕ which can be proved in $I\mathcal{E}^2_*$ to be the characteristic function for ϕ. Therefore we will refer to such formulas ϕ as $\mathcal{E}^2_*$ formulas.

Recall that $\mathrm{PHP}(\mathcal{E}^2_*)$ states:

$$\forall x\, \forall y(\forall u_1 \leq y\, \forall u_2 \leq y\, \forall v \leq y(\theta(\vec{x},y,u_1,v) \wedge \theta(x,y,u_2,v) \rightarrow u_1 = u_2)$$
$$\wedge\, \forall u \leq y\, \exists v \leq y\ \theta(\vec{x},y,u,v) \rightarrow \forall v \leq y\, \exists u \leq y\ \theta(\vec{x},y,u,v)),$$

for each $\mathcal{E}^2_*$ formula $\theta(\vec{x},y,u,v)$.

Working in $I\mathcal{E}^2_*$, suppose $\theta(\vec{x},y,u,v)$ is an $\mathcal{E}^2_*$ formula and that, contrary to $\mathrm{PHP}(\mathcal{E}^2_*)$, for some values of the parameters $\vec{x},y$,

$$g(u) = \min\{v\colon\ \theta(\vec{x},y,u,v)\}$$

defines a function $g\colon [0,y] \rightarrow [0,y]$ which is one-to-one but not onto. Let $f(u) = g(u)+1$ and $n = y+1$ so $f\colon [1,n] \rightarrow [1,n]$ is also one-to-one but not onto. We may suppose (by redefining f at $f^{-1}(n)$ if this exists) that $f\colon [1,n] \rightarrow [1,n-1]$.

Let $\phi(\vec{x},y,u,v)$ be an $\mathcal{E}^2_*$ formula defining the predicate $f(u) \leq v$. The census function $c_\theta(\vec{x},y,u,v)$ for ϕ with respect to u has the $\mathcal{E}^2$ definition:

$$c_\phi(\vec{x},y,0,v) = 0$$

$$c_\phi(x,y,u+1,v) = c_\phi(x,y,u,v) + \chi_\phi(x,y,u+1,v)$$

where X_ϕ is the $\mathcal{E}^2$ characteristic function of ϕ. Thus in $I\mathcal{E}^2_*$ we can apply induction to $\Delta_0(f,c_\phi)$ formulas and deduce by theorem 5.3 that PHP(f) holds, contradicting our assumption that PHP($\mathcal{E}^2_*$) fails.

In $I\mathcal{E}^2_*$ we can also prove the following version of the pigeon hole principle provided A,B and f are defined by $\mathcal{E}^2_*$ formulas (with parameters allowed):

If $B \subset A \subseteq [1,n]$, $B \neq A$, *and* $f\colon A \to B$, *then* f *is not one-to-one.*

This is because we have $\mathcal{E}^2$ census functions $c_A(x)$, $c_B(x)$ for A and B, so taking c_A^{-1} to be the one-to-one function with range A defined by:

$$c_A^{-1}(y) = \min\{x\colon c_A(x) = y\},$$

we can prove in $I\mathcal{E}^2_*$ that if f is one-to-one then for some k,m with $k > m$, the function $c_B\, f\, c_A^{-1}\colon [1,k] \to [1,m]$ (formed by composition of functions) is one-to-one (and onto) which is contrary to corollary 5.4.

Similarly the bijection extension principle, namely:

If $A,B \subseteq [1,n]$ *and* $f\colon A \leftrightarrow B$ *then there is an extension* f^* *of* f *such that* $f^*\colon [1,n] \leftrightarrow [1,n]$, is available in $I\mathcal{E}^2_*$ for A,B,f with $\mathcal{E}^2_*$ definitions. For (working in $I\mathcal{E}^2_*$) suppose there is some bijection $f\colon A \leftrightarrow B$ where $A,B \subset [1,n]$. Then (as above) there is some k such that

$$c_B\, f\, c_A^{-1}\colon [1,k] \leftrightarrow [1,k].$$

(Intuitively $k = |A| = |B|$.) Denoting the complements of A,B by A',B', it can be proved that the $\mathcal{E}^2$ functions

$$c_{A'}(x) = x - c_A(x), \quad c_{B'}(x) = x - c_B(x)$$

have the property that $c_{A'}^{-1}\colon [1,n-k] \leftrightarrow [1,n]\setminus A$ and $c_{B'}^{-1}\colon [1,n-k] \leftrightarrow [1,n]\setminus B$, and therefore

$$c_{B'}^{-1} c_{A'}\colon [1,n]\setminus A \leftrightarrow [1,n]\setminus B.$$

Thus we can take

$$f^*(x) = \begin{cases} f(x) & \text{if } x \in A, \\ c_{B'}^{-1} c_{A'}(x) & \text{if } x \in [1,u] \setminus A. \end{cases}$$

As indicated in the introduction, a special case of the $\Delta_0 = \mathcal{E}_*^2$ problem asks whether for each predicate $\theta(\vec{x},y)$ defined on N by a Δ_0 formula, the corresponding census function $c_\theta(x,\vec{y})$ has the property that the predicate $z = c_\theta(x,\vec{y})$ can also be defined by a Δ_0 formula. One can "strengthen" this question to asking whether there is some <u>uniform</u> way of defining the census function. For example is there some <u>bounded</u> formula $\phi(x,y,A)$ in a new unary predicate variable A (as well as $+,\cdot,\leq$) such that for all $A \subseteq N, x, z \in N$,

$$z = c_A(x) \iff \phi(x,z,A)?$$

The following corollary gives a conditional answer to an axiomatic version of this problem (in obvious notation).

<u>COROLLARY 5.5</u> *If conjecture 5.1 is true then*

$$I\Delta_0(A) \nvdash \phi(0,0,A) \wedge \forall x \forall z(\phi(x,z,A) \to$$
$$(\phi(x+1,z+1,A) \iff A(x+1)) \wedge (\phi(x+1,z,A) \iff \neg A(x+1)))$$

for any $\Delta_0(A)$ *formula* $\phi(x,z,A)$.

<u>Proof</u>:

Suppose the contrary held. Then replacing $A(v)$ in $\phi(x,z,A)$ by the formula $f(v) \leq y$ would yield a $\Delta_0(f)$ definition of $z = c_\theta(x,y)$ (where $\theta(x,y)$ is $f(x) \leq y$), and we would have $I\Delta_0(f) \vdash \mathrm{def}(c_\theta)$. Therefore by theorem 5.3, $I\Delta_0(f) \vdash \mathrm{PHP}(f)$ in contradiction to proposition 5.2.

§6. On summing sequences

In this section we will consider the problem of summing a Δ_0 sequence $u_1,u_2,\ldots,u_d$ in a model $M \models I\Delta_0$ with $d \leq [\log_2 a]^n$ for some $a \in M$, $n \in N$. But first we will prove the following number theoretic lemma:

LEMMA 6.1 *There is some* $C \in N$ *such that*

$$I\Delta_0 \vdash \forall x \exists y \,(y = \prod_{p \leq C[\log_2 x]} p \quad \wedge y > x).$$

Proof:

Let $x_0 \in M \models I\Delta_0$. (The constant C produced by the proof will clearly be independent of the choice of M.) In order to prove the lemma for $x = x_0$ we may suppose that $x_0 \in M \setminus N$, since if $x_0 \in N$ then the result follows immediately by standard number theory (for example, remark 4.4).

Define an approximate logarithm function $\log^*$ as in §4 using $a = [\log_2 x_0]^2$ and $n = 2$. Then for $x \leq a$ there is a Δ_0 definition of the sum $\sum_{p \leq x} \log^* p$ by lemma 2.3, and using the method described in the proof of theorem 4.4 we can prove a * analogue of the result of remark 4.4, namely that $\sum_{p \leq x} \log^* p \geq C_0 x$ for all $x \in [2,a]$, where C_0 is some fixed standard rational number. Similarly a * analogue of inequality (9) of §4 can be proved leading to the result:

$$\sum_{p \leq x} \log^* p < C_1 x \quad \text{for some fixed } C_1 \in N, \text{ and all } x \leq a.$$

In both cases the pigeon hole principle is used to establish the relevant inequality (for example, inequality $(9)^*$ of §4 in the second case) but each of these applications of the pigeon hole principle applies to a Δ_0 function

$$f: [1,m_1] \longrightarrow [1,m_2] \text{ with } m_1, m_2 \leq a^n \leq [\log_2 x_0]^{2n}$$

for some $n \in N$, and can therefore be handled by theorem 5.3, since the census function for the formula $f(x) \le y$ can obviously be defined using lemma 2.3. Similarly all uses of lemma 4.4 in the original arguments can be replaced by uses of lemma 2.3, since taking $a \le [\log_2 x_0]^2$ ensures that the size of the terms to be summed will be sufficiently small.

Now choose $C \varepsilon N$ large enough so that $\sum_{p \le Cx} \log^* p > (1+\varepsilon)x$ for all $x \le a/C$, where $\varepsilon > 0$ is some standard rational number. In particular this inequality will be satisfied for $x = [\log_2 x_0]$. Thus

$$(1+\varepsilon)[\log_2 x_0] < \sum_{p \le C[\log_2 x_0]} \log^* p < D \log^* x_0$$

where $D > C\, C_1/\log 2$, $D \varepsilon N$.

Using the fact that $|\log^*(u.v) - \log^* u - \log^* v| \le \dfrac{A}{[\log_2 a]^2}$ for some constant $A \varepsilon N$, it follows by induction on $i \le C[\log x_0]$ that:

$$\exists y \le x_0^{D+\varepsilon} \;(\; y \textit{ is squarefree} \wedge \forall p(\; p \textit{ is prime} \to (\; p|y \iff p \le i))$$

$$\wedge\; (1+\varepsilon)[\log_2 x_0] - \frac{Ai}{[\log_2 a]^2} \le \log^* y + \sum_{i < p \le C[\log_2 x_0]} \log^* p < D \log^* x_0 + \frac{Ai}{[\log a]^2}$$

Taking $i = [\log_2 x_0]$, it follows that a number $y = \prod_{p \le C[\log_2 x]} p$ exists, and that $[\log_2 x_0] < \log^* y$ so $x_0 < y$.

REMARK: Notice that the product version of Sylvester's method (described in §1) does not seem to be applicable, since it would involve numbers of size $[\log_2 x_0]!$ (that is roughly $x_0^{[\log_2[\log_2 x]]}$) which need not exist in M.

We can now prove:

LEMMA 4.2 *If* $M \models I\Delta_0$ *and* $u_1, u_2, \ldots, u_d$ *is a* Δ_0 *sequence of elements* *of* M *with* $d \leqslant [\log_2 a]^n$ *and each* $u_i \leqslant u$ *for some* $a, u \in M$, $n \in N$, *then there is a* Δ_0 *sequence* $\sum_{1\leqslant i\leqslant j} u_i$, $j = i,2,\ldots,d$, *such that* $\sum_{1\leqslant i\leqslant 1} u_i = u_1$ *and for all* $j \in [1,d-1]$,

$$\sum_{1\leqslant i\leqslant j+1} u_i = \sum_{1\leqslant i\leqslant j} u_i + u_{j+1}.$$

Proof:

Clearly if the sum exists its value will be $\leqslant [\log_2 a]^n u = v$ (say). By lemma 6.1 there is some $C \in N$ such that there exists $y \in M$ with $y = \prod_{p\leqslant C[\log_2 v]} p > v$. The idea is to do the summation mod p for each of these primes p. Let $(x)_p$ denote the least number w such that $x \equiv w \pmod p$. Then for each $p \leqslant C[\log_2 v]$, the numbers $(u_1)_p, (u_2)_p, \ldots, (u_d)_p$ form a Δ_0 sequence having p as a parameter in its definition. Since each $(u_i)_p \leqslant C[\log_2 v]$ the sum $\sum_{1\leqslant i\leqslant j} (u_i)_p$ can be defined for $j \in [1,d]$ using lemma 2.3. Now write:

$$z = \sum_{1\leqslant i\leqslant j} u_i \iff$$

$$z \leqslant v \wedge \forall p\,(\, p \textit{ is prime} \wedge p|y \rightarrow \sum_{1\leqslant i\leqslant j} (u_i)_p \equiv z \pmod p).$$

More precisely a unique z satisfying the right hand side of this equivalence can be proved to exist by induction on j, and a similar induction shows

$$\sum_{1\leqslant i\leqslant j+1} u_i = \sum_{1\leqslant i\leqslant j} u_i + u_{j+1}.$$

The lemma fails if the condition that there be some upper bound u on the elements of the Δ_0 sequence is omitted.

Obviously if the Δ_0 sequence $u_1, u_2, \ldots, u_d$ has no maximum element (that is, if $M = \{x \in M: \exists i \in [1,d]\ (\ x \leqslant u_i)\})$ then $\sum_{1\leqslant i\leqslant j} u_j$ cannot exist for all j, since if it did we could prove by a

Δ_0 induction on $d-j$ that $u_j \leqslant \sum_{1\leqslant i\leqslant j} u_i \leqslant \sum_{1\leqslant i\leqslant d} u_i$ for all $j \in [1,d]$, which is impossible. Therefore the problem amounts to showing that such Δ_0 sequences do exist in some model M (and that it is possible to have $d \leqslant [\log_2 a]^n$ for some $a \in M$, $n \in N$). This has been done by:

PARIS, J.B.; KIRBY, L.A.S., Σ_n - *Collection schemas in arithmetic.* Logic Colloquium '77, North Holland, 1978.

§7. The future?

Wilkie's question of whether

$I\Delta_0 \vdash$ *there exist arbitrarily large prime numbers*

remains an intrigueing problem. On the one hand there is the theorem of Alex. Wilkie that the existence of arbitrarily large primes cannot be proved by induction on open formulas (together with the usual "algebraic" axioms) which has recently been extended by Zofia Adamowicz (unpublished) to induction on certain formulas with simple quantifier prefixes. On the other hand there is the possibility that the number theoretic properties of $I\Delta_0$ (which as we have seen are quite strong in some ways) may be powerful enough to prove the theorem. However even if this does transpire, the story cannot end there. For consider Linnik's theorem (see for example Prachar [1958]) that there is some constant C such that for all $a,b \in N$ with a coprime to b, there exists a prime $p \equiv a \pmod b$, $p < b^C$. In view of this theorem it is just as sensible to ask whether

$I\Delta_0 \vdash$ *Dirichlet's theorem*

that is, the theorem which states that if a,b are coprime then there exist arbitrarily large primes $p \equiv a \pmod b$.

Of course there are numerous other examples of number theoretic properties which hold in every structure of the form $\{b \in M\colon \exists n \in N(b \leqslant a^n)\}$, where M is a model of (say) Peano arithmetic and $a \in M$, but for which no proof in $I\Delta_0$ is known. *The really important problem would seem to be to find a proof that some "natural" property of this sort cannot be proved in* $I\Delta_0$. Other possible examples include additive basis theorems such as that of Schnirelmann [1933] which states that there is some constant C such that every $x \in N$, $x \geqslant 2$, can be written as a sum of primes $x = p_1 + p_2 + \ldots + p_m$ with $n \leqslant C$, or even the analogous result with squarefree numbers instead of primes. (There do exist arbitrarily large squarefree numbers in any model of $I\Delta_0$.)

The proof that $I\Delta_0 \vdash$ *Sylvester's theorem* also suggests some lines of further enquiry. As the reader may have noticed, much of the proof hinged on avoiding explicit use of the bijection extension principle.

PROBLEM: *Is the bijection extension principle a conservative extension of* $I\Delta_0 + PHP(\Delta_0)$ *in the sense that adding new function symbols* $f_1, f_2, \ldots$ *for the functions whose existence is demanded by the principle, and allowing induction on* $\Delta_0(f_1, f_2, \ldots)$ *formulas, cannot prove any extra theorems in the original language?*

Of course this would follow from a proof of $\Delta_0 = \mathcal{E}^2_*$ which did not go beyond the capabilities of $I\Delta_0 + PHP(\Delta_0)$, but perhaps this is more than is needed.

The $I\Delta_0 + PHP(\Delta_0)$ proofs of the existence of arbitrarily large primes also seem to make essential use of argument by contradiction, since if we do not assume that $y+1, y+2, \ldots, y+x$ have no prime divisor greater than x then we have no way of proving that the function to which we wish to apply the pigeon hole principle is properly defined. This motivates the following question:

PROBLEM: *Can classical* $I\Delta_0 + PHP(\Delta_0)$ *be intepreted in intuitionistic* $I\Delta_0 + PHP(\Delta_0)$?

(Where intuitionistic $I\Delta_0 + PHP(\Delta_0)$ is defined in an obvious manner.) Of course we have the usual (Gödel) interpretation of classical $I\Delta_0$ in intuitionistic $I\Delta_0$, so this question may have to await the solution of the $I\Delta_0 \vdash PHP(\Delta_0)$ problem. On the other hand it might conceivably provide valuable insight into that problem.

CHAPTER 2 Definability properties of the coprimeness predicate.

This chapter describes some of the logical properties of the two place predicate $\perp$ defined for natural numbers x,y by:

$x \perp y \iff x$ *and* y *have no common prime divisor.*

Throughout $L_{\alpha,\beta,\ldots,\gamma}$ will denote the language consisting of all formulas which can be constructed (according to the usual rules) from the logical symbols $\forall,\exists,\neg,\wedge,\vee,\rightarrow$; variables $x,y,z,\ldots$ intended to range over the natural numbers N ; and the predicates or operations $\alpha,\beta,\ldots,\gamma$. For example, $L_{=,{}',|}$ is the language with equality $(x=y)$ and divisibility $(x\,|\,y)$ predicates, and the successor operation $(x'=x+1)$.

Given a language L for N, a basic "first question" is whether addition and multiplication are L definable, that is, are there L formulas $\phi_A(x,y,z)$, $\phi_M(x,y,z)$ such that for all $x,y,z \in N$, $\phi_A(x,y,z) \iff x+y=z$ and $\phi_M(x,y,z) \iff x.y=z$? If so, then a second basic question is: what sort of defining formulas are possible? Julia Robinson [1949] made a fundamental contribution to this subject when she investigated these questions for various languages, and in particular proved that addition and multiplication are $L_{{}',|}$ definable. If we regard $'$ and $|$ as weak primitives of an additive and multiplicative kind respectively, then we see that in a certain sense this theorem is stronger than the corresponding results for $L_{\leqslant,|}$, $L_{=,+,|}$, $L_{=,{}',\cdot}$ and $L_{\leqslant,\cdot}$. For successor can be defined from $\leqslant$ which in turn is $L_{=,+}$ definable, but $\leqslant$ is not $L_{=,{}'}$ definable, nor is addition $L_{\leqslant}$ definable. Similarly divisibility is $L_{=,\cdot}$ definable, but multiplication is not $L_{|}$ definable. Thus a natural way to attempt to strengthen the theorem is to replace $|$ by a new "weaker" primitive which is $L_{|}$ definable, but which, by itself, is not strong enough to allow $|$ to be defined. The comprimeness predicate being an obvious candidate, Julia Robinson asked in her paper whether multiplication (and therefore addition) can be defined in $L_{=,{}',\perp}$ or even whether multiplication is $L_{=,+,\perp}$ definable. The present chapter is primarily devoted to these questions and the intermediate problem of defining multiplication in $L_{\leq,\perp}$.

In §1 two simple proofs of the $L_{=,+,\perp}$ definability of multiplication are given. Correspondence with Professor Robinson has revealed that the second of these was discovered independently (and much earlier) by her but has not previously been published. A stronger theorem, namely that addition and multiplication can be defined by <u>bounded</u> $L_{\leq,\perp}$ formulas, is proved in §4 by a more complicated argument. In fact, a relation on N can be defined by a bounded $L_{\leq,+,\cdot}$ formula if and only if it can be defined by a bounded $L_{\overset{*}{\leq}}$ formula, where $\overset{*}{\leq}$ is the preordering defined by:

$$x \overset{*}{\leq} y \iff x = y \vee (x \leq y \wedge x \perp y) .$$

(A *preordering* is a binary relation which can be extended to a linear ordering. Note that $\overset{*}{\leq}$ is *not* transitive.) As the relations on N which can be defined by bounded $L_{\leq,+,\cdot}$ formulas were shown by Bennett [1962] to be precisely the *rudimentary predicates* of computational complexity theory, this theorem yields a new characterisation of these. In particular it follows that every rudimentary set of positive natural numbers is the spectrum of a sentence ϕ of the predicate calculus with a single binary predicate symbol γ (say). (The *spectrum* of a sentence ϕ is taken here to be the set $\{|M| : M \models \phi\}$ comprised of the cardinalities of the domains of all finite normal models M of ϕ.) Furthermore it is shown that ϕ may be chosen in such a way that all finite normal models of ϕ are graphs (that is, have γ symmetric and antireflexive) and that similar results hold for partial orderings and quasiorderings.

Concerning the original problem about the properties of $L_{\prime,\perp}$ it is proved in §2 that there cannot be any algorithm for deciding whether arbitrary sentences of this language are true. Also, Julia Robinson's question about the $L_{=,\prime,\perp}$ definability of multiplication turns out to be equivalent to an open problem in number theory. Indeed all of the following statements are equivalent:

(i) $z = x.y$ is $L_{=,\prime,\perp}$ (or $L_{\prime,\perp}$) definable.

(ii) $z = x+y$ is $L_{=,\prime,\perp}$ (or $L_{\prime,\perp}$) definable.

(iii) $x \leqslant y$ is $L_{=,\prime,\perp}$ (or $L_{\prime,\perp}$) definable.

(iv) $x = y$ is $L_{\prime,\perp}$ definable.

(v) *There is some* $k \in N$ *such that every natural number* x *is determined uniquely by the sequence* $S_0, S_1, \ldots, S_k$ *of sets of (distinct) prime numbers defined by* $S_i = \{p : p \mid x+i\}$.

In view of the classical theorem of Størmer [1897] which states that for any x there are at most finitely many numbers which produce the same sets S_0, S_1 as x, and the (admittedly limited) numerical evidence which can be extracted from the tables compiled by Lehmer [1964], it seems plausible to conjecture that (v) may be true for quite small values of k. This conviction is strengthened by the fact that statement (v) (with $k = 20$ for x sufficiently large) is a consequence of a general conjecture of Hall and Schinzel about the magnitude of the solutions of certain diophantine equations. The truth of (i) - (v) would also follow from an affirmative answer to a question posed recently by Erdös [1980], who asked whether there is some k such that every natural number x is determined uniquely by $\bigcup_{i \leqslant k} S_i$.

§0. Two facts about langauges for number theory.

As a preliminary we will list two "pieces of folklaw" which have consequences for many languages which describe the natural numbers. The first is a corollary of Matijasevič's negative solution to Hilbert's "tenth problem" about the existence of an algorithm for deciding whether a polynomial equation with integer coefficients has a solution in N. (An account of this work may be found in Davis [1973].)

An *existential formula* is one of the form $\exists\vec{w}\,Q(\vec{w})$ where $Q(\vec{w})$ is quantifier free. For any language L, the *existential* L *theory* of N consists of all existential sentences of L which are true of N. A theory T is *decidable* if there is an algorithm for deciding whether an arbitrary sentence is in T.

PROPOSITION 0.1 *If the existential* $L_{=,\alpha,\beta,\ldots,\gamma}$ *theory of* N *is decidable then at least one of the predicates* $z = x + y$, $z = x.y$ *cannot be defined by an existential* $L_{=,\alpha,\beta,\ldots,\gamma}$ *formula.*

Proof:

Set $A(x,y,z) \iff z = x + y$ and $M(x,y,z) \iff z = x.y$. For each $L_{=,+,\cdot,\prime}$ sentence ψ of the form

$$\exists\vec{w}\,(P(\vec{w}) = Q(\vec{w})) \qquad (*)$$

there is an effectively found equivalent $L_{=,A,M}$ sentence of the form $\exists\vec{y} \bigwedge_{h=1}^{m} \psi_h(\vec{y})$ where each $\psi_h(\vec{y})$ is of one of the forms $A(y_i, y_j, y_k)$, $M(y_i, y_j, y_k)$, $y_i = y_j$. (The reason for including this last form is that $x = 1 \iff \exists u \exists v (\neg u = v \wedge x.u = u \wedge x.v = v)$.)

Now if $\phi_A(x,y,z)$ and $\phi_M(x,y,z)$ are existential $L_{=,\alpha,\beta,\ldots,\gamma}$ definitions for $A(x,y,z)$ and $M(x,y,z)$ respectively, then each $\psi_h(\vec{y})$ of the form $A(y_i, y_j, y_k)$ can be replaced by $\phi_A(y_i, y_j, y_k)$, and similarly each $\psi_h(\vec{y})$ of the form $M(y_i, y_j, y_k)$ can be replaced by $\phi_M(y_i, y_j, y_k)$. The resulting sentence is clearly equivalent both to ψ and to an effectively found existential $L_{=,\alpha,\beta,\ldots,\gamma}$ sentence. Thus if the existential $L_{=,\alpha,\beta,\ldots,\gamma}$ theory of N is decidable then there is a decision procedure for $L_{=,+,\cdot,}$ sentences of form (*). But since the terms $P(\vec{w})$, $Q(\vec{w})$ are (essentially) arbitrary polynomials, this contradicts Matijasevič's theorem.

Consider any language with a symbol ($\leq$, say) intended to denote an ordering or preordering. A *quantified variable* $\forall x$ or $\exists x$ is *bounded* by y (with respect to $\leq$) in a formula ψ if it begins a subformula of ψ of the form $\forall x(x \leq y \rightarrow \phi)$ or $\exists x(x \leq y \wedge \phi)$. (These will be abbreviated to $\forall x \leq y\ \phi$ and $\exists x \leq y\ \phi$.) A *formula* ψ is *bounded* if every quantified variable in ψ is bounded by some variable. If only one of the primitive symbols of the language denotes an ordering or preordering, "bounded" will, of course, mean bounded with respect to that symbol.

The basic idea behind the next proposition already appears in the literature, for example in work of Paris and Dimitracopoulos [????]. However the proof will be sketched here as we will later make use of the technique and it is desirable to emphasise one of its subtleties.

PROPOSITION 0.2 *If the predicates* $z = x + y$, $z = x.y$ *are definable by bounded* $L_{\leq,\alpha,\beta,\ldots,\gamma}$ *formulas then every bounded* $L_{\leq,+,\cdot}$ *formula is equivalent to an effectively found bounded* $L_{\leq,\alpha,\beta,\ldots,\gamma}$ *formula.*

Proof:

Taking $A(x,y,z)$, $M(x,y,z)$ as in the previous proof, it suffices to show that each bounded $L_{\leq,+,\cdot}$ formula $\psi(\vec{w})$ is equivalent to an effectively found bounded $L_{\leq,A,M}$ formula $\psi^*(\vec{w})$.

Note that $\psi(w_1, w_2, \ldots, w_m) \iff \bigvee_{i=1}^{m} \theta_i(\vec{w})$ where

$$\theta_i(\vec{w}) \iff \bigwedge_{\substack{j=1\\ j\neq i}}^{m} (w_j \leq w_i) \wedge \psi(w_1, w_2, \ldots, w_m),$$ so we may consider each $\theta_i(\vec{w})$ separately. Clearly all terms occuring in $\theta_i(\vec{w})$ can be bounded by a polynomial in w_i, and indeed (at the expense of treating the first few values of w_i separately) it can be assumed that the bound is w_i^n for some constant $n \in N$.

Replacing the operations $+$ and $.$ in $\theta_i(\vec{w})$ by the predicates A, M using quantified variables bounded by a new variable c, yields a bounded $L_{\leqslant,A,M}$ formula $\rho_i(\vec{w}, c)$ such that

$$\forall\vec{w}\ \forall c \geq w_i^n\ (\rho_i(\vec{w}, c) \Longleftrightarrow \theta_i(\vec{w})).$$

But each number less than $([\sqrt{w_i}] + 1)^{2n}$ has a representation $a_1\, a_2 \ldots a_{2n}$ to base $[\sqrt{w_i}] + 1$, the digits of which are numbers $a_k \leqslant [\sqrt{w_i}]$, and it is easy to see that there are bounded $L_{\leqslant,A,M}$ formulas $\phi_A(\vec{x},\vec{y},\vec{z},w_i)$, $\phi_M(\vec{x},\vec{y},\vec{z},w_i)$, $\phi_{\leqslant}(\vec{x},\vec{y},w_i)$ such that for any three numbers less than $([\sqrt{w_i}] + 1)^{2n}$ with representations $x_1\, x_2 \ldots x_{2n}$, $y_1\, y_2 \ldots y_{2n}$, $z_1\, z_2 \ldots z_{2n}$,

$$\phi_A(x_1, x_2, \ldots, x_{2n}, y_1, y_2, \ldots, y_{2n}, z_1, z_2, \ldots, z_{2n})$$

$$\Longleftrightarrow x_1\, x_2 \ldots x_{2n} + y_1\, y_2 \ldots y_{2n} = z_1\, z_2 \ldots z_{2n}, \text{ etc.. }$$

Also (and this point is <u>very</u> important) there is a bounded $L_{\leqslant,A,M}$ formula $\eta(x, \vec{a}, w_i)$ such that

$$\eta(x, \vec{a}, w_i) \Longleftrightarrow x \leqslant w_i \wedge x = a_1\, a_2 \ldots a_{2n}.$$

By quantifying over 2n-tuples $a_1, a_2, \ldots, a_{2n}$ in place of numbers less than $([\sqrt{w_i}] + 1)^n$ and using $\phi_A, \phi_M, \phi_{\leqslant}$ instead of $A, M, \leqslant$ where appropriate, we can now obviously construct a bounded $L_{\leqslant,A,M}$ formula equivalent to $\rho_i(\vec{w}, ([\sqrt{w_i}] + 1)^{2n} - 1)$, and therefore equivalent to $\theta_i(\vec{w})$.

§1. Addition and coprimeness.

We now give two different ways of defining multiplication by an $L_{=,+,\perp}$ formula. The observation underlying the first of these is:

LEMMA 1.1 *Given a sequence of numbers* $v_0 < v_1 < \ldots < v_n$ *there exist numbers* x,c *such that for all* y, *if* $x + v_0 \leq y \leq x + v_n$ *then*

$$y \perp c \iff \exists i(y = x + v_i) .$$

Proof:

Let $w_0, w_1, \ldots, w_k$ be the numbers between v_0 and v_n which are not v_i's, and choose a sequence of $k+1$ distinct primes p_j such that $p_j \nmid \prod_{i \leq n} (v_i - w_j)$. By the Chinese remainder theorem there exists a number $x \equiv -w_j (\mathrm{mod}\, p_j)$ for all j. But it is easy to check that if $c = \prod_{j \leq k} p_j$ then x,c have the required property.

THEOREM 1.2 $z = x.y$ *is* $L_{=,+,\perp}$ *definable*.

Proof:

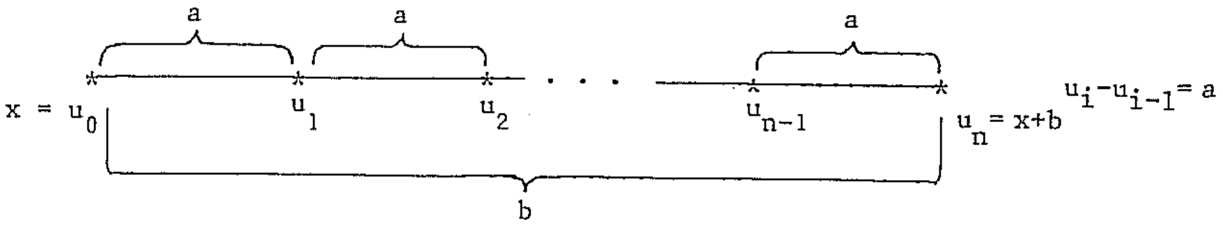
Since $z = x.y$ is $L_{=,+,|}$ definable (see Tarski [1949] or Robinson [1949]) it suffices to show that the divisibility predicate $a|b$ is $L_{=,+,\perp}$ definable.

It is claimed that $a|b$ if and only if there exist x,c such that

(i) $x \perp c \wedge (x + b) \perp c$

(ii) *If* $x \leq u < u_* \leq x + b$ *and* u,u_* *are consecutive numbers coprime to* c *then* $u_* = u + a$.

To see this note that (i) and (ii) describes the pattern:



where the numbers between x and $x + b$ which are coprime to c are $x = u_0 < u_1 < u_2 < \ldots < u_n = x + b$. Obviously if such a pattern exists then $a|b$. Conversely, if $a|b$ then lemma 1.1 ensures that such a pattern does exist.

Since $\leqslant$ is $L_{=,+}$ definable (i) and (ii) can clearly be expressed by $L_{=,+,\perp}$ formulas so the theorem follows.

Analysing the above argument shows that $x.y = z \iff \exists w\, \psi(x, y, z, w)$ for some bounded $L_{\leqslant,+,\perp}$ formula $\psi(x, y, z, w)$. Our second proof of theorem 1.2 will do better than this.

<u>THEOREM 1.3</u> $z = x.y$ *is definable by a bounded* $L_{\leqslant,+,\perp}$ *formula.*

<u>Proof</u> (also found by J. Robinson):

Note first that

$x = 0 \iff \forall y \leqslant x\ (y \geqslant x)$

$x = 1 \iff \neg x = 0 \wedge \forall y \leqslant x\ (y = 0 \vee y = x)$

x is a prime $\iff \neg x = 0 \wedge \neg x = 1 \wedge \forall y \leqslant x(y = 0 \vee y \perp x \vee y = x)$.

By Schnirelmann's theorem [1933] there is a natural number n such that every number $x \geqslant 2$ is the sum of fewer than n primes. Thus the product of any two numbers $x, y \geqslant 2$ can be expressed in the form

$$x.y = (p_1 + \dots + p_m).(q_1 + \dots + q_k) = \sum_{i,j} p_i.q_j$$

where $x = p_1 + \dots + p_m$, $y = q_1 + \dots + q_k$, the p_i's and q_j's are primes, and $m, k < n$. Since the sum $\sum_{i,j} p_i.q_j$ has fewer than n^2 terms the theorem will follow if it can be shown that there is a bounded $L_{\leqslant,\perp}$ formula which defines $z = p.q$ for all primes p, q. But $z = p.q \iff$ z *is the least number greater than both* p *and* q *such that*

$$\neg z \perp p \wedge \neg z \perp q \wedge \forall y \leqslant z\, (y \perp p \wedge y \perp q \rightarrow y \perp z)\ .$$

Recalling proposition 0.2 and noting that $x \perp y$ can be defined by a bounded $L_{\leqslant,+,\cdot}$ formula we have:

<u>COROLLARY 1.3</u> *A relation on* N *can be defined by a bounded* $L_{\leqslant,+,\cdot}$ *formula if and only if it can be defined by a bounded* $L_{\leqslant,+,\perp}$ *formula.*

In the course of settling Hilbert's tenth problem Matijasevič proved the following theorem. (See, for example, Davis [1973].)

<u>PROPOSITION 1.4</u> (Matijasevič) *Every bounded* $L_{\leqslant,+,\cdot}$ *formula is equivalent to an existential* $L_{\leqslant,+,\cdot}$ *formula.*

However the analogue of this fails for $L_{\leqslant,+,\perp}$ since $z = x.y$ cannot be defined by an existential $L_{\leqslant,+,\perp}$ formula. This follows from a result of Bel'tyukov [1976] and Lipshitz [1978]. (See also Wilkie [????].)

PROPOSITION 1.5 (Bel'tyukov, Lipshitz) *The existential* $L_{\leqslant,+,',|}$ *theory of* N *is decidable.*

COROLLARY 1.6 *The existential* $L_{\leqslant,+,\perp}$ *theory of* N *is decidable.*

Proof:

We transform the decision problem for existential $L_{\leqslant,+,\perp}$ sentences into a decision problem for certain existential $L_{\leqslant,+,',|}$ sentences. Given any existential $L_{\leqslant,+,\perp}$ sentence ϕ, there is an effectively found equivalent sentence of the form $\exists\vec{y} \bigvee_{i=1}^{k} \psi_i(\vec{y})$, where each $\psi_i(\vec{y})$ is a finite conjunction of atomic and negated atomic formulas. If all unnegated atomic formulas of the form $s \perp t$ (s and t terms) are replaced by $\exists a(s \mid a \wedge t \mid a')$ and all negated atomic formulas of the form $\neg s \perp t$ are replaced by $\exists a \exists b(\neg a \mid b \wedge a \mid s \wedge a \mid t)$, then the resulting $L_{\leqslant,+,',|}$ sentence (equivalent to ϕ) is clearly equivalent to an effectively found existential $L_{\leqslant,+,',|}$ sentence.

Applying proposition 0.1 gives immediately:

COROLLARY 1.7 $z = x.y$ *cannot be defined by an existential* $L_{\leqslant,+,\perp}$ *formula.*

§2. Successor and coprimeness.

We begin by defining some notation which will be used throughout this section.

P is the set of all prime numbers. (p and q will always denote elements of P.)

$\nu(x)$ is the number of distinct prime divisors of x.

$x \approx y \iff \nu(x) = \nu(y)$

$x \sim y \iff x$ *and* y *have the same (distinct) prime divisors.*

$x \underset{n}{\sim} y \iff x \sim y \wedge x+1 \sim y+1 \wedge \ldots \wedge x+n \sim y+n$.

(Thus $x = y \rightarrow \ldots \rightarrow x \underset{n+1}{\sim} y \rightarrow x \underset{n}{\sim} y \rightarrow \ldots \rightarrow x \underset{1}{\sim} y \rightarrow x \sim y \rightarrow x \approx y$.)

For $x \neq 0$, $\bar{x}, \bar{x}^n$ and $\bar{\bar{x}}$ will denote, respectively, the $\sim, \underset{n}{\sim}$ and $\approx$ equivalence classes of which x is a representative, and we will write $\bar{x} \mid y$ or $\bar{x} \mid \bar{y}$ if $\forall p \in P\,(p \mid x \rightarrow p \mid y)$. Let $\bar{P} = \{\bar{p} : p \in P\}$. Elements of $\bar{P}$ will be denoted by the capital letters U,V,W,X,Y,Z. Thus, for example, $Z = \{p^m : m \in N \setminus \{0\}\}$ for some prime p, and $Z \mid x \iff \bar{p} \mid x \iff p \mid x$.

The reason for introducing these concepts is that with the exception of $\approx, \bar{\bar{x}}, \nu(x)$ and P, they are all easy to define in $L_{\prime,\perp}$. In particular we have:

(i) $x \sim y \iff \forall z(z \perp x \iff z \perp y)$

(ii) $\bar{x} \mid y \iff \forall z(z \perp y \rightarrow z \perp x)$

(iii) $\bar{x} \in \bar{P} \iff \forall y\, \forall z(y \perp z \rightarrow y \perp x \vee z \perp x) \wedge \exists y(\neg y \perp x)$

(iv) $x = 0 \iff \forall y(\neg y \perp x \vee \forall z(y \perp z))$.

Consequently when constructing "$L_{\prime,\perp}$ formulas" we may use the expressions on the left of (i), (ii), (iii), terms $0, 1, 2, \ldots$ (that is, $0, 0', 0'', \ldots$), the capital letter notation for elements of $\bar{P}$, and, for each fixed $n \in N$, the predicates $x \underset{n}{\sim} y$ and $\bar{x}^n = \bar{y}^n$.

Let $\bar{\bar{N}} = \{\bar{\bar{x}} : x \in N \setminus \{0\}\}$, $\bar{N} = \{\bar{x} : x \in N \setminus \{0\}\}$, $\bar{N}^n = \{\bar{x}^n : x \in N \setminus \{0\}\}$. Notice that the map $\nu^{-1} : N \rightarrow \bar{\bar{N}}$ defined by $\nu^{-1}(\nu(x)) = \bar{\bar{x}}$ is a bijection and may thus be used to induce operations $+, \cdot$ on $\bar{\bar{N}}$ so that $\langle N, =, +, \cdot \rangle \cong \langle \bar{\bar{N}}, =, +, \cdot \rangle$ under ν^{-1}. This is the basic observation underlying the proofs in this section. It will be shown below that when considered as predicates on N, the expressions $\bar{\bar{x}} = \bar{\bar{y}}$, $\bar{\bar{x}} + \bar{\bar{y}} = \bar{\bar{z}}$ and $\bar{\bar{x}} . \bar{\bar{y}} = \bar{\bar{z}}$ can be defined by $L_{',\perp}$ formulas, and hence there is an effective way of transforming any sentence ψ of $L_{=,+,\cdot}$ into an equivalent $L_{',\perp}$ sentence ψ^* asserting $<\bar{\bar{N}}, =, +, \cdot> \models \psi$. Thus we will be able to deduce the undecidability of the $L_{',\perp}$ theory of N from the known undecidability of the $L_{=,+,\cdot}$ theory. Also, the existence of these formulas reduces the problem of the $L_{',\perp}$ definability of $z = x.y$, etc., to the question of whether the isomorphism ν^{-1} (or equivalently the predicate $y = \nu(x)$) is $L_{',\perp}$ definable.

To ease notational problems we will adopt the following conventions:

(i) If $\theta(y_1, \ldots, y_n)$ is an $L_{=,+,\cdot}$ definable predicate and $\bar{\bar{x}}_1, \ldots, \bar{\bar{x}}_n \varepsilon \bar{\bar{N}}$ then $\bar{\bar{N}} \models \theta(\bar{\bar{x}}_1, \ldots, \bar{\bar{x}}_n)$ will be used as an abbreviation for

$$< \bar{\bar{N}}, =, +, \cdot> \models \theta(y_1, \ldots, y_n) [\bar{\bar{x}}_1, \ldots, \bar{\bar{x}}_n] .$$

(ii) If $y = \gamma(x_1, \ldots, x_n)$ is an $L_{=,+,\cdot}$ definable function, then $\gamma(\bar{\bar{x}}_1, \ldots, \bar{\bar{x}}_n)$ will denote the element $\bar{\bar{y}} \varepsilon \bar{\bar{N}}$ which satisfies $\bar{\bar{N}} \models \bar{\bar{y}} = \gamma(\bar{\bar{x}}_1, \ldots, \bar{\bar{x}}_n)$. (For example, $\bar{\bar{x}} + 1$ is the element $\bar{\bar{y}}$ for which $\bar{\bar{N}} \models \bar{\bar{y}} = \bar{\bar{x}} + 1$).

(iii) We will write $\bar{\bar{x}} \leq \bar{\bar{y}}$ instead of $\bar{\bar{N}} \models \bar{\bar{x}} \leq \bar{\bar{y}}$.

LEMMA 2.1 $\bar{\bar{x}} \leq \bar{\bar{y}}$ *is* $L_{',\perp}$ *definable.*

Proof:

Write $x \overset{*}{\leq} y$ if there exists u such that for

$$f_u = \{<X, Y> : X|x \wedge Y|y \wedge X \neq \bar{2} \wedge Y \neq \bar{2} \wedge \exists z(\bar{z} \varepsilon \bar{P} \wedge \bar{z}|u \wedge X | z' \wedge Y | z')\}$$

one of the following is satisfied:

(i) $\neg \bar{2}|x \vee (\bar{2}|x \wedge \bar{2}|y)$ *and* f_u *is the graph of a one-to-one function mapping* $\{X : X|x \wedge X \neq \bar{2}\}$ *into* $\{Y : Y|y \wedge Y \neq \bar{2}\}$.

(ii) $\bar{2}|x \wedge \neg \bar{2}|y$ *and* f_u *is the graph of a one-to-one function mapping* $\{X : X|x \wedge X \neq \bar{2}\}$ *onto a proper subset of* $\{Y : Y|y\}$.

(iii) $\neg \bar{2}|x \wedge \bar{2}|y$ *and* f_u *is the graph of a one-to-one function with range contained in* $\{Y : Y|y \wedge Y \neq \bar{2}\}$ *and domain equal to* $\{X : X|x\} \smallsetminus \{W\}$

for some $W|x$.

Obviously $x \overset{*}{\leqslant} y$ can be defined by an $L_{\prime,\perp}$ formula. Also, for any numbers x,y if $x \overset{*}{\leqslant} y$ then $x \neq 0$, $y \neq 0$ and there is a one-to-one map from $\{\bar{q} : q|x\}$ into $\{\bar{q} : q|y\}$, so $\nu(x) \leqslant \nu(y)$ and therefore $\bar{\bar{x}} \leqslant \bar{\bar{y}}$.

On the other hand, if $\bar{\bar{x}} \leqslant \bar{\bar{y}}$ then $x \neq 0$, $y \neq 0$ and $\nu(x) \leqslant \nu(y)$, so it is possible to choose a one-to-one map taking each prime $q|x$ to a prime $r_q|y$, and having the additional property that if $q|y$ then $r_q = q$. Now choose for each $q|x$ a prime p_q satisfying

$$\begin{aligned} p_q &\equiv -1 \pmod{q} \\ &\equiv -1 \pmod{r_q} \\ &\equiv \;\; 1 \pmod{s} \quad \text{for all other primes } s|x.y\,. \end{aligned}$$

Such primes p_q exist by virtue of the Chinese remainder theorem (which is used to combine these congruences into a single congruence of the form $p_q \equiv a (\mathrm{mod}\ b)$ with $a \perp b$) and Dirichlet's theorem that if $a \perp b$ then there are infinitely many primes $p \equiv a (\mathrm{mod}\ b)$. (A proof of Dirichlet's theorem may be found in Shapiro [1950].) Noting that $p \equiv 1 (\mathrm{mod}\ s) \Rightarrow p^m \not\equiv -1 (\mathrm{mod}\ s)$ for all m and all $s > 2$, it is easy to check that $u = \prod_{q|x} p_q$ has the property required to make $x \overset{*}{\leqslant} y$.

Thus, $\bar{\bar{x}} \leqslant \bar{\bar{y}} \iff x \overset{*}{\leqslant} y$.

LEMMA 2.2 $\bar{\bar{x}} + \bar{\bar{y}} = \bar{\bar{z}}$ *and* $\bar{\bar{x}}.\bar{\bar{y}} = \bar{\bar{z}}$ *are* $L_{\prime,\perp}$ *definable.*

Proof:

We simply extend the trick of coding one-to-one mappings of prime divisors developed in lemma 2.1 to define $+,\cdot$ on $\bar{\bar{N}}$ in the same way that one defines $+,\cdot$ for cardinal arithmetic in set theory. For example, $\bar{\bar{x}}.\bar{\bar{y}} = \bar{\bar{z}}$ if and only if there exist some representatives x, y, z of these equivalence classes such that

$$\{<X,Y,Z> : X|x \wedge Y|y \wedge Z|z \wedge \exists w(\bar{w} = Z \wedge X|w' \wedge Y|w')\}$$

is the graph of a one-to-one function from $\{<X,Y> : X|x \wedge Y|y\}$ onto $\{Z : Z|z\}$.

In fact, since $\bar{\bar{x}} \leqslant \bar{\bar{y}}$ is $L_{\prime,\perp}$ definable, and addition is $L_{\leqslant,\cdot}$

definable, it is sufficient just to define $\bar{\bar{x}}.\bar{\bar{y}} = \bar{\bar{z}}$.

It follows immediately from these lemmas that for each $L_{=,+,\cdot}$ formula $\psi(x_1, \ldots, x_n)$ there is an effectively found $L_{\prime,\perp}$ formula $\psi^*(x_1, \ldots, x_n)$ such that

$$\psi^*(x_1, \ldots, x_n) \iff \bar{\bar{N}} \models \psi(\bar{\bar{x}}_1, \ldots, \bar{\bar{x}}_n) .$$

In particular, since $\langle \bar{\bar{N}}, =, +, \cdot \rangle \cong \langle N, =, +, \cdot \rangle$, we have thus proved:

<u>THEOREM 2.3</u> *For every sentence* ψ *of* $L_{=,+,\cdot}$ *there is an effectively found sentence* ψ^* *of* $L_{\prime,\perp}$ *such that* $\psi \iff \psi^*$.

<u>COROLLARY 2.4</u> *The* $L_{\prime,\perp}$ *theory of* N *is undecidable.*

It seems prudent at this stage to eliminate any further "coding" difficulties by proving a general coding lemma (lemma 2.6) which will be more than adequate for the purposes of this paper. As a preliminary we define $\delta : \bar{P} \to \bar{N}$ by taking $\delta(\bar{p}) = \overline{p+1}$ for all primes p, and prove:

<u>LEMMA 2.5</u> $\bar{y} = \delta(Z)$ *is* $L_{\prime,\perp}$ *definable. (More precisely, the predicate* $\bar{x} \in \bar{P} \wedge \bar{y} = \delta(\bar{x})$ *is* $L_{\prime,\perp}$ *definable.)*

<u>Proof</u>:

Let $\zeta(x, y, Z)$ be the formula:

$$Z = \overline{x'} \wedge \bar{y} = \overline{x''} \wedge \forall w(Z = \overline{w'} \to \bar{x}|w \wedge (\bar{w}|x \to \overline{x''}|w'')).$$

It is claimed that $\bar{y} = \delta(Z) \iff \exists x\, \zeta(x, y, Z)$.

To check this, suppose first that for some Z, $\zeta(x, y, Z)$ is satisfied by $\langle x, y\rangle = \langle x_1, y_1\rangle, \langle x_2, y_2\rangle$. Then (taking $w = x_2$ in $\zeta(x_1, y_1, Z)$) we see that $\bar{x}_1 \mid x_2$ and (taking $w = x_1$ in $\zeta(x_2, y_2, Z)$) it follows that $\overline{x_2''}|x_1''$. Similarly, $\overline{x_1''}|x_2''$ so $\overline{x_1''} = \overline{x_2''}$, that is, $y_1 \sim y_2$.

Thus if $\exists x\, \zeta(x, y_1, Z)$ holds then $\{y : \exists x\, \zeta(x, y, Z)\} = \bar{y}_1$, so it suffices to show $\exists x\, \zeta(x, p+1, Z)$ for the prime p with

$$Z = \bar{p} = \{p^m : m > 0\}.$$

But $x = p-1$ has this property. For using the fact that $\overline{w'} = Z$ if and only if $w = p^m - 1$ for some $m > 0$, it can be seen that $\zeta(p-1, p+1, Z)$ is equivalent to $\forall m > 0(\overline{p-1}|p^m - 1 \wedge (\overline{p^m - 1}|p - 1 \to \overline{p+1}|p^m + 1))$, and as

$p-1|p^m-1$, this in turn is equivalent to the formula:

$$\forall m>0\,(\forall q \in P\,(\,p^m \equiv 1 (\mathrm{mod}\ q) \rightarrow p \equiv 1 (\mathrm{mod}\ q)) \longrightarrow$$
$$\forall q \in P\,(\,p \equiv -1\,(\mathrm{mod}\ q) \rightarrow p^m \equiv -1\,(\mathrm{mod}\ q))),$$

which is true.

LEMMA 2.6 *For each* $n \in N$ *there is an* $L_{',\perp}$ *formula* $\theta_n(x_1,\ldots,x_n,u)$ *with the property that for every finite set* $R \subset \bar{N} \times \bar{N} \times \ldots \times \bar{N}$ (n *copies of* $\bar{N}$) *there is some* $\bar{u} \in \bar{N}$ *such that for every* $v \in \bar{u}$,

$$\forall x_1 \ldots \forall x_n\,(<\bar{x}_1,\ldots,\bar{x}_n> \in R \iff \theta_n(x_1,\ldots,x_n,v)).$$

Proof:

It suffices to prove the lemma for $n = 2$ since the general case then follows by induction, defining θ_{n+1} by:

$$\theta_{n+1}(x_1,\ldots,x_{n+1},u) \iff \exists v(\theta_n(x_1,\ldots,x_n,v) \wedge \theta_2(x_{n+1},v,u)).$$

The proof has two parts. The first describes how to "decode" an arbitrary $\bar{u} \in \bar{N}$ to obtain a finite set $R_{\bar{u}} \subset \bar{N} \times \bar{N}$ in such a way that when considered as a predicate in x_1, x_2, u, the expression $<\bar{x}_1,\bar{x}_2> \in R_{\bar{u}}$ is $L_{',\perp}$ definable. In the second part it is shown that if $R \subset \bar{N} \times \bar{N}$ is finite then a code $\bar{u}$ with $R_{\bar{u}} = R$ can always be found.

As an aid to decoding $\bar{u}$, define $\bar{a} \cap \bar{b}$ for $\bar{a},\bar{b} \in \bar{N}$ by $\bar{c} = \bar{a} \cap \bar{b} \iff \forall Z \in \bar{P}\,(\,Z|c \iff Z|\bar{a} \wedge Z|\bar{b})$. (If $\bar{a},\bar{b}$ are regarded as denoting subsets of $\bar{P}$ then $\bar{a} \cap \bar{b}$ denotes their intersection). For $\bar{z} \in \bar{N}$, $W \in \bar{P}$ define $(\bar{z})_W \in \bar{N}$ by:

$$\forall Z \in \bar{P}\,(\,Z|(\bar{z})_W \iff (Z \neq \bar{2} \wedge Z \neq W \wedge Z|\bar{z}) \vee (Z = \bar{2} \wedge W|\bar{z})).$$

($(\bar{z})_W$ corresponds to the subset of $\bar{P}$ obtained by replacing W by $\bar{2}$ in the set denoted by $\bar{z}$, if W occurs there, and by suppressing any occurrence of $\bar{2}$ otherwise.)

We also introduce an $L_{',\perp}$ definable partial ordering $\overset{*}{<}$ on $\bar{P}$ by taking

$$Z_1 \overset{*}{<} Z_2 \iff \overline{\overline{\delta(Z_1)}} < \overline{\overline{\delta(Z_2)}}\ ,$$

where $\bar{a} = \overline{\overline{\delta(Z)}} \iff \exists b(a \approx b \wedge \bar{b} = \delta(Z))$. In other words, if $Z_1 = \bar{p}_1$, $Z_2 = \bar{p}_2$

with p_1, p_2 prime, then $Z_1 \overset{*}{<} Z_2$ if and only if $p_1 + 1$ has fewer (distinct) prime divisors than $p_2 + 1$.

The first step in decoding an arbitrary $\bar{u} \in \bar{N}$ is to check whether there exist $W, X \in \bar{P}$ and $\bar{a}, \bar{b} \in \bar{N}$ such that $\bar{\bar{a}} = \bar{\bar{b}}$,

$$\{Z \in \bar{P} : Z|\bar{u}\} = \{W, X\} \cup \{(Y \in \bar{P} : Y|\bar{a}\} \cup \{Z \in \bar{P} : Z|\bar{b}\} ,$$

and this set is totally ordered by $\overset{*}{<}$ with $W \overset{*}{<} X \overset{*}{<} Y \overset{*}{<} Z$ for all $Y|\bar{a}$, $Z|\bar{b}$. If not, put $R_{\bar{u}} = \phi$. If so, then regard W as a surrogate for $\bar{2}$ (this is needed because, as in the proof of lemma 2.1, it is necessary to treat $\bar{2}$ as a special case) and define $R_{\bar{u}}$ by:

$<\bar{x}_1, \bar{x}_2> \in R_{\bar{u}} \iff$ *there exist* $Y|\bar{a}$, $Z|\bar{b}$ *such that*

$$Y|\delta(Z) \wedge \bar{x}_1 = (\delta(X) \cap \delta(Y))_W \wedge \bar{x}_2 = (\delta(X) \cap \delta(Z))_W .$$

In view of lemma 2.5 it is obvious that there is an $L_{\prime,\perp}$ formula $\theta_2(x_1, x_2, u)$ such that

$$\forall x_1 \ \forall x_2 (<\bar{x}_1, \bar{x}_2> \ \in R_{\bar{u}} \iff \theta_2(x_1, x_2, u)) .$$

To complete the proof it suffices to show that for every finite $R \subset \bar{N} \times \bar{N}$ there is some $\bar{u} \in \bar{N}$ such that $R = R_{\bar{u}}$. If elements of $\bar{N}$ are thought of as denoting sets of primes then any such R may be regarded as a binary relation between subsets of some finite set $S \subset P$. Choose a prime $w \notin S$ as the surrogate for 2 and form a new relation R^* by replacing 2 by w at each occurrence of 2 in the subsets related by R. Thus R^* is a relation between "subsets" $\bar{y}$ of $S^* = S \cup \{w\} \setminus \{2\}$. Enumerate all of these as $\bar{y}_1, \bar{y}_2, \ldots, \bar{y}_m$. Now using the Chinese remainder theorem and Dirichlet's theorem choose, <u>in turn</u>, odd primes

$r < s_1 < s_2 < \ldots < s_m < t_1 < t_2 < \ldots < t_m$ satisfying:

(i) $r \equiv -1 \pmod{q}$ for all primes $q \in S^*$.

(ii) For $i = 1, 2, \ldots, m$,

$s_i \equiv -1 \pmod{q}$ for all primes $q|y_i$,

$\not\equiv -1 \pmod{q}$ for all other primes $q|r+1$.

(iii) For $j = 1, 2, \ldots, m$,

$t_j \equiv -1 \pmod{q}$ for all primes $q|y_j$,

$t_j \not\equiv -1 \pmod q$ for all other primes $q | r+1$,

$\equiv -1 \pmod{s_i}$ for all i such that $< \bar{y}_i, \bar{y}_j > \in R$,

$\not\equiv -1 \pmod{s_i}$ for all i such that $< \bar{y}_i, \bar{y}_j > \notin R$.

(iv) $\overline{w} \overset{*}{<} \overline{r} \overset{*}{<} \overline{s_1} \overset{*}{<} \overline{s_2} \overset{*}{<} \dots \overset{*}{<} \overline{s_m}$

$\overset{*}{<} \overline{t_1} \overset{*}{<} \overline{t_2} \overset{*}{<} \dots \overset{*}{<} \overline{t_m}$.

This last condition can be satisfied by adding, <u>at each stage</u>, sufficiently many extra congruences of the form:

$$r, s_i, \text{ or } t_j \equiv -1 \pmod q, \ q \text{ prime} .$$

Taking $W = \bar{w}$, $X = \bar{r}$, $a = \prod_{1 \leq i \leq m} s_i$, $b = \prod_{1 \leq j \leq m} t_j$ it is now easy to check that $R = R_{\bar{u}}$ for $u = w.r.a.b$.

<u>LEMMA 2.7</u> *For each* $n \in N$ *there is an* $L_{\prime,\perp}$ *formula* $\chi_n(x, y, u)$ *with the property that for any finite set* $R \subset \bar{\bar{N}} \times \bar{N}^n$ *there is some* $u \in N$ *such that* $\forall x \forall y(<\bar{\bar{x}}, \bar{y}^n> \in R \iff \chi_n(x, y, u))$.

<u>Proof:</u>

With θ_{n+2} as in lemma 2.6, take $\chi_n(x, y, u)$ to be the formula:

$$\exists z \ \exists \vec{w}(x \approx z \wedge \bigwedge_{i=0}^{n} (w_i \sim y + i) \wedge \theta_{n+2}(z, w_0, w_1, \dots, w_n, u)) \ .$$

This works because of the obvious correspondence between $\bar{y}^n$ and $<\bar{y}, \overline{y+1}, \dots, \overline{y+n}>$, and the fact that $\approx$ equivalence classes are unions of $\sim$ equivalence classes so $\bar{x}$ can be used as a "representative" for $\bar{\bar{x}}$.

To complete the preparations for our final theorem about definability in $L_{\prime,\perp}$ we require a simple number theoretic lemma.

<u>LEMMA 2.8</u> *If* $a \underset{k}{\sim} b$ *and* $a \neq b$ *then*

$$\prod_{p \leq k+1} p \ \leq \prod_{p | a(a+1)\dots(a+k)} p \ \leq \ |a-b| \ .$$

<u>Proof:</u>

In fact $\leq$ can be replaced by $|$. If $p \leq k+1$ then p divides at least one of the $k+1$ numbers $a, a+1, \dots, a+k$. Also if $p | a+i$ for some $i \leq k$, then since $a \underset{k}{\sim} b$, it follows that $p | b+i$ and hence $p \ \big| \ |a-b|$.

THEOREM 2.9 *The following statements are equivalent:*

(i) $x = y$ *is* $L_{\prime,\perp}$ *definable.*

(ii) $z = x.y$ *is* $L_{\prime,\perp}$ *(or* $L_{=,\prime,\perp}$*) definable.*

(iii) $z = x + y$ *is* $L_{\prime,\perp}$ (or $L_{=,\prime,\perp}$) *definable.*

(iv) $x \leqslant y$ *is* $L_{\prime,\perp}$ *(or* $L_{=,\prime,\perp}$*) definable.*

(v) $\exists k\ \forall x \forall y(x \underset{k}{\sim} y \rightarrow x = y)$.

Proof:

Obviously (v) → (i), while (ii) → (iii) follows from the $L_{=,\prime,\cdot}$ definability of addition (Robinson [1949]), and (iii) → (iv) is a consequence of the $L_{=,+}$ definability of $\leqslant$. Thus it will suffice to show (i) → (v) and (iv) → (v) → (ii).

(i) → (v):

Suppose $\phi(x, y)$ is an $L_{\prime,\perp}$ definition of equality, and let k be the largest number for which the variable y followed by k successor symbols occurs in $\phi(x, y)$. Since all atomic subformulas of $\phi(x, y)$ (and in particular all those containing y) are of the form $u^{\prime\prime\cdots\prime} \perp v^{\prime\prime\cdots\prime}$ (u and v variables) it is clear that

$$\forall y \forall z(y \underset{k}{\sim} z \rightarrow \forall x(\phi(x, y) \Longleftrightarrow \phi(x, z))).$$

Replacing ϕ by $=$ in this formula we see that

$$\forall y \forall z(y \underset{k}{\sim} z \rightarrow y = z).$$

(iv) → (v):

This can be proved by an extension of the argument used to show (i) → (v), however it seems easier to consider a nonstandard model $\langle M, =, +, \cdot\rangle$ of the $L_{=,+,\cdot}$ theory of N.

Suppose (v) fails. Then

$$M \models \forall k\ \exists x\ \exists y(x < y \wedge x \underset{k}{\sim} y).$$

Take a nonstandard number $c \in M$ and let $a, b \in M$ have the property that

$$M \models a < b \wedge a \underset{2c}{\sim} b\ .$$

By lemma 2.8, $b - a$ is nonstandard. This allows us to define a bijection $f : M \to M$ by:

$$\left.\begin{aligned} f(a+c+j) &= b+c+j \\ f(b+c+j) &= a+c+j \end{aligned}\right\} \text{ for all standard (positive, negative, or zero) integers } j,$$

$$f(d) = d\,, \text{ otherwise.}$$

Clearly f is an automorphism of $<M, =, ', \perp>$, but f does not preserve $\leq$, so this relation cannot be $L_{=,',\perp}$ definable.

(v) → (ii):

Suppose that $\forall y \forall z(y \underset{k}{\sim} z \to y = z)$. Then for every $y \neq 0$ the $\underset{k}{\sim}$ equivalence class $\bar{y}^k$ contains only the single element y, so by lemma 2.7 the $L_{',\perp}$ formula $\chi_k(x, y, u)$ has the property that for any finite set $R \subset \bar{\bar{N}} \times (N \smallsetminus \{0\})$ there is some $u \in N$ such that

$$\forall x \forall y(<\bar{\bar{x}}, y> \ \varepsilon \ R \Longleftrightarrow \chi_k(x, y, u)).$$

It follows that the isomorphism ν^{-1}, or more precisely, the predicate $\bar{\bar{x}} = \nu^{-1}(y)$ (which is the same as $\nu(x) = y$) is $L_{',\perp}$ definable. For $\bar{\bar{x}} = \nu^{-1}(y)$ if and only if $x = 1 \wedge y = 0$ or there is some finite set $R \subset \bar{\bar{N}} \times N$ such that

(1) R *is the graph of a one-to-one function,*

(2) $<\bar{\bar{2}}, 1> \ \varepsilon \ R$ (that is, $<\nu^{-1}(1), 1> \ \varepsilon \ R$)

(3) $\forall \bar{\bar{z}} < \bar{\bar{x}} \ \forall w \ (<\bar{\bar{z}}, w> \varepsilon \ R \to <\bar{\bar{z}}', w'> \ \varepsilon \ R)$

(4) $<\bar{\bar{x}}, y> \ \varepsilon \ R$,

Appealing to lemma 2.2 where necessary, it is clear that the existence of an R satisfying (1) - (4) can be asserted by an $L_{',\perp}$ formula.

But $z = x.y \Longleftrightarrow$

$$\exists u \exists v \exists w(\bar{\bar{u}} = \nu^{-1}(x) \wedge \bar{\bar{v}} = \nu^{-1}(y) \wedge \bar{\bar{w}} = \nu^{-1}(z) \wedge \bar{\bar{w}} = \bar{\bar{u}}.\bar{\bar{v}})$$

so multiplication is $L_{',\perp}$ definable.

This last theorem shows that settling the question of the $L_{',\perp}$ definability of multiplication is tantamount to establishing the truth or otherwise of the number theoretic problem (v). It seems unreasonable to expect that the latter problem (and therefore the former) can be handled without the use of nontrivial number theoretic methods.

POSTCRIPT: Recently the author has learned that Denis Richard has independently obtained a different proof of the undecidability of the $L_{=,\ ,\perp}$ theory of N (cf. corollary 2.4). His proof is based on the following number theoretic fact:

PROPOSITION (Birkhoff and Vandiver [1904]) *Suppose* $a > b$, $a \perp b$, *and* $n > 2$. *Then except for the single case* $a = 2$, $b = 1$, $n = 6$, *there is always some prime* $p \mid a^n - b^n$ *such that* $p \nmid a^m - b^m$ *for all* $m \in [1,n-1]$.

This also has the interesting consequence that all numbers x of the form $x = p^n - 1$, $p \in P$ have the property:

$$\forall y(x \underset{2}{\sim} y \rightarrow x = y).$$

§3. A number theoretic interlude(†).

It is timely now to survey what is currently known (for k fixed) about how large $|a-b|$ must be if $a \underset{k}{\sim} b$ with $a \neq b$. If a, b have this property then from lemma 2.8,

$$|a-b| \geqslant \prod_{p|a(a+1)\dots(a+k)} p \quad , \tag{1}$$

which suggests seeking good lower bounds on the product

$$\prod_{p|a(a+1)\dots(a+k)} p \quad . \tag{2}$$

As will be seen below, it follows from a conjecture of Hall and Schinzel that this approach should actually succeed in proving the conjecture

$$\exists k\ \forall a\ \forall b (a \underset{k}{\sim} b \rightarrow a = b).$$

For $k > 1$, the bounds on (2) discussed here are all due to Langevin [1975a], [1975b] and [1979], however as it seems desirable to collect the relevant details in one place, their derivation will be described.

The following lemmas will enable us to concentrate on bounding $\prod_{p|a(a+1)} p$ and $\prod_{p|a(a+2)} p$. For any $A \subseteq N$ let $S(A) = \{p : \exists a \varepsilon A (p|a)\}$.

LEMMA 3.1 *Suppose* $A = \bigcup_i A_i \subseteq \{a, a+1, \dots, a+k\}$ *where the* A_i*'s are all pairwise disjoint. Then*

$$\prod_{p\,\varepsilon\, S(A)} p \;\geqslant\; k^{-k+1} \prod_i \prod_{p\,\varepsilon\, S(A_i)} p \quad .$$

Proof:

Any prime p divides at most $[k/p] + 1$ of the numbers $a, a+1, \dots, a+k$ and at least $[k/p]$ of the numbers $2, 3, 4, \dots, k$, so

$$\prod_i \prod_{p\,\varepsilon\, S(A_i)} p \;\leqslant\; \left(\prod_{p \leqslant k} p^{[k/p]} \right) . \left(\prod_{p\,\varepsilon\, S(A)} p \right)$$

$$\leqslant\; k! \prod_{p\,\varepsilon\, S(A)} p \;\leqslant\; k^{k-1} \prod_{p\,\varepsilon\, S(A)} p \quad .$$

LEMMA 3.2 *Suppose* $g(a)$ *is a nondecreasing function such that for all* a,

(†) The only subsequent use made of the material in this section is in the proof of the $L_{\leqslant,\perp}$ definability of addition and multiplication given in §4. An alternative proof of that theorem which does not rely on the present section may be found in appendix I.

$$\prod_{p|a(a+1)} p > (g(a))^2 \quad \textit{and} \quad \prod_{p|a(a+2)} p > (g(a))^2 .$$

Then for all $k \geq 1$,

$$\prod_{p|a(a+1)\ldots(a+k)} p > k^{-k} (g(a))^{k+1} , \tag{3}$$

and for $k = 2$ *or* $k \geq 4$, $i \leq k$,

$$\prod_{p|\frac{a(a+1)\ldots(a+k)}{(a+i)}} p > k^{-k} (g(a))^{k} . \tag{4}$$

Proof:

Since 2 is the only prime which can divide more than one of the numbers $a, a+1, a+2$,

$$2\left(\prod_{p|a(a+1)(a+2)} p\right)^2 \geq \left(\prod_{p|a(a+1)} p\right).\left(\prod_{p|(a+1)(a+2)} p\right).\left(\prod_{p|a(a+2)} p\right),$$

so

$$\prod_{p|a(a+1)(a+2)} p > \frac{1}{\sqrt{2}} (g(a))^3 \geq \frac{1}{\sqrt{k}} (g(a))^3 \quad \text{for} \quad k > 1.$$

Writing

$$\{a,a+1,\ldots,a+k\} = \begin{cases} \bigcup_{i\leq m} \{a+2i,\ a+2i+1\}, & \text{if } k = 2m+1, \\ \bigcup_{i<m} \{a+2i,\ a+2i+1\} \cup \{a+2m,\ a+2m+1,\ a+2m+2\}, & \\ & \text{if } k = 2m+2, \end{cases}$$

(3) follows by lemma 3.1. The proof of (4) is similar.

For $n = 1,2$, bounds on $\prod_{p|a(a+n)} p$ are given by theorems 7 and 8 of Lehmer [1964]. These yield:

$$\sum_{p|a(a+n)} \log p > \frac{2}{3} (1 - o(1)) \log\log a, \quad \text{for} \quad n = 1,2. \tag{5}$$

(Lehmer's work also gives the corresponding result for $n = 4$.) $o(1)$ here denotes some real valued function of a which tends to zero as $a \to \infty$. Similarly, $o_k(1)$ will denote some function of a,k which, for each fixed k, tends to zero as $a \to \infty$.

By lemma 3.2 we obtain for all $k \geq 1$, Langevin's result:

$$\sum_{p|a(a+1)\dots(a+k)} \log p > \frac{k+1}{3}(1-o(1))\log\log a - k\log k$$
$$> \frac{k+1}{3}(1-o_k(1))\log\log a, \tag{6}$$

and for $i \leqslant k$ with $k = 2$ or $k \geqslant 4$, the analogous inequality:

$$\sum_{p|\frac{a(a+1)\dots(a+k)}{(a+i)}} \log p > \frac{k}{3}(1-o(1))\log\log a - k\log k$$
$$> \frac{k}{3}(1-o_k(1))\log\log a. \tag{7}$$

Combining (1) and (6) gives:

PROPOSITION 3.3 *For every real number* $\varepsilon > 0$, *and every* $k \geqslant 1$, *there exists* $a_k(\varepsilon)$ *such that for all* $a > a_k(\varepsilon)$ *and all* $b \neq a$,

$a \underset{k}{\sim} b \to \log|a-b| > \frac{k+1}{3}(1-\varepsilon)\log\log a.$

Lehmer deduced his inequalities (5) from properties of the Pell equation:

$$x^2 - dy^2 = 1 \tag{8}$$

using Størmer's observation that if $x = 2a+1$ then for some y, and some $d|\prod_{p|a(a+1)} p$,

$$(x-1)(x+1) = 4a(a+1) = dy^2$$

so x,y,d satisfy (8). (For $\prod_{p|a(a+2)} p$ take $x = a+1$.) An alternative approach used by Langevin is to consider the Mordell equation:

$$y^2 = x^3 + m . \tag{9}$$

Let v be the least number such that for some w,

$$4a(a+n) = (2a+n)^2 - n^2 = vw^3 \tag{10}$$

Then $v|(2\prod_{p|a(a+n)} p)^2$ and multiplying (10) by v^2 yields:

$$(v(2a+n))^2 = (vw)^3 + (vn)^2. \tag{11}$$

That is, $x = vw$, $y = v(2a+n)$ is a solution of (9) with $m = (vn)^2 \neq 0$.

THEOREM 3.4 *If there exist positive real valued constants* C_0, D *such that*

$$\forall y \forall z \neq 0(\exists x(y^2 - z^2 = x^3) \to z^D \geqslant C_0 y) \tag{12}$$

then there is a constant a_0 *such that for* $k > 4D-5$, $\forall a > a_0\ \forall b(a \underset{k}{\sim} b \rightarrow a = b)$.

Proof:

Applying (12) to (11) shows $v^{D-1} \geq C_0(2a+n)n^{-D}$,

so $$\prod_{p|a(a+n)} p \geq \frac{1}{2}\,(C_n(2a+n))^{1/(2D-2)}$$

where $C_n > 0$ depends only on n. Taking $n = 1,2$ it follows by lemma 3.2 that for each $k \geq 1$ there is a constant $C_k^* > 0$ such that

$$\prod_{p|a(a+1)\ldots(a+k)} p \geq C_k^* a^{(k+1)/(4D-4)}$$

(This inequality is implicit in Langevin [1979].) Thus there is some a_0 such that for all $a > a_0$,

$$\prod_{p|a(a+1)\ldots(a+k)} p > a \quad \text{for} \quad k = [4D-5] + 1 \, .$$

But if $a \underset{k}{\sim} b$ with $b \neq a$ and $a > a_0$ then we may suppose without loss of generality that $a > b$ and therefore $a - b > \prod_{p|a(a+1)\ldots(a+k)} p > a$ which is impossible.

The hypothesis of theorem 3.4 (with $D = 6$) is a consequence of the following conjecture attributed by Langevin [1975b] to Hall and Schinzel:

CONJECTURE (Hall and Schinzel) *For any pair of natural numbers* $n > 1$, $m > 1$, *there is a constant* C, *such that for all integers* x, y *with* $x^n \neq y^m$, $|x^n - y^m|^6 \geq C \max\{|x|^n, |y|^m\}$.

Thus it follows from this conjecture (with $n = 3$, $m = 2$) that if a is sufficiently large then a is determined uniquely by the sequence $S_0, S_1, \ldots, S_{20}$ where $S_i = \{p : p|a+i\}$. Although at present the hypothesis of theorem 3.4 remains unproved, Stark [1973] has shown that for every real number $\varepsilon > 0$ there is a constant $C(\varepsilon)$ such that all integer solutions of (9) with $m \neq 0$ satisfy

$$\log|m| + C(\varepsilon) > (1-\varepsilon) \log\log \max\{|x|, |y|\}.$$

Applying this to (11) yields:

$$2\log(vn) + C(\varepsilon) > (1-\varepsilon) \log\log(v(2a+n)) \, ,$$

from which it follows that

$$\sum_{p|a(a+n)} \log p > \frac{1}{4}\,(1-o(1))\log\log(2a+n) - \frac{1}{2}\log n\,. \tag{13}$$

As noted by Langevin [1975b], if $b > a$ and we take $n = b - a$ then (13) becomes

$$\sum_{p|ab} \log p + \frac{1}{2}\log(b-a) > \frac{1}{4}\,(1-o(1))\log\log(a+b)$$

and hence by (1),

$$a \sim b \rightarrow \log(b-a) > \frac{1}{6}\,(1-o(1))\log\log(a+b)\ ,$$

giving a bound for the case $k = 0$ which was omitted from proposition 3.3. (The existence of a constant $C > 0$ such that for $a \neq b$,

$$a \sim b \rightarrow \log|b-a| > C\log\log(a+b)$$

was also proved by Erdös and Shorey [1976] using a different method.)

(13) also provides an alternative starting point for obtaining weaker versions (with $\frac{1}{3}$ replaced by $\frac{1}{8}$) of the inequalities (6) and (7). This version of (7) is, nevertheless, quite suitable for use in the proofs of the definability results given in the next section. However the reader is warned that the $L_{\leqslant,\perp}$ defining formulas for addition and multiplication produced in this way may be much longer, not only because larger values of k must be used, but more importantly because it would seem a must be taken much larger in order to reduce the $o(1)$ term to a reasonable size if one starts with Stark's work rather than Lehmer's.

§4. Order and comprimeness.

We now turn to the problem of giving $L_{\leqslant,\perp}$ defintions of addition and multiplication. The basic idea is to exploit the fact that if $|a-b|$ is small with respect to a suitably chosen function of a and k then we certainly do have:

$$a \underset{k}{\sim} b \to a = b \,.$$

For technical reasons it will be desirable to exclude certain "troublesome" primes from the discussion by using a slightly weaker equivalence relation than $\sim$, namely that defined between integers a and b (for a given q) by:

$$a \overset{q}{\sim} b \iff \{p : p|a \wedge p \nmid q\} = \{p : p|b \wedge p \nmid q\}.$$

We will also want an analogue of $\underset{k}{\sim}$ which applies to arithmetic progressions other than $a, a+1, \dots, a+k$. (To make it easier to obtain bounded formulas this will be defined using $a, a-d, a-2d, \dots, a-kd$ rather than $a, a+d, a+2d, \dots, a+kd$.) Put:

$$a_{k,\tilde{d},q}\, b \iff a \overset{q}{\sim} b \wedge a-d \overset{q}{\sim} b-d \wedge \dots \wedge a-kd \overset{q}{\sim} b-kd \,.$$

LEMMA 4.1 *For* $k=2$ or $k \geqslant 4$ *and any real number* $\varepsilon > 0$, *there is some* $a_k(\varepsilon)$ *such that for all primes* q, *and all* a, b *with* $a > a_k(\varepsilon)$,

$$a_{k,\tilde{1},q}\, b \to \log|a-b| > \frac{k}{3}(1-\varepsilon)\log\log a \ \vee\ a=b \,.$$

Proof:

Suppose $a_{k,\tilde{1},q}\, b$ with q prime. The argument used to prove lemma 2.8 shows that if $a \neq b$ then

$$\log|a-b| \geqslant \sum_{\substack{p|a(a-1)\dots(a-k)\\ p\neq q}} \log p \,,$$

$$\geqslant \sum_{p \,\big|\, \frac{a(a-1)\dots(a-k)}{(a-i)}} \log p - \log k$$

for some $i \leqslant k$, since either $q \leqslant k$ or q divides at most one of the numbers $a+i$, $i \leqslant k$. The result now follows by inequality (7) of §3.

On several occassions we will have cause to use Chebychev's weak version of the Prime Number theorem (see for example, Hardy and Wright [1979] or §4 of chapter 1 of this thesis).

PROPOSITION 4.2 (Chebychev) *There is a constant* $A > 0$ *such that for all* $x \geqslant 2$, $\sum_{p \leqslant x} \log p > Ax$.

LEMMA 4.3 *There is a constant* B *such that for all primes* q *and all* a,b,k *with* $k > B \log a$,

$$a \geqslant b \wedge a \underset{k,1,q}{\sim} b \to a = b.$$

Proof:

Suppose that $a > b$ and $a \underset{k,1,q}{\sim} b$ for $k = [B \log a] + 1$. Then

$$\begin{aligned}\log(a-b) &\geqslant \sum_{\substack{p \mid a(a-1)\ldots(a-k) \\ p \neq q}} \log p \\ &\geqslant \sum_{p \leqslant k} \log p - \log k \quad \text{(as in lemma 2.8)} \\ &> Ak - \log k > (A/2)k\end{aligned}$$

for B (and therefore k) sufficiently large, where A is the constant in proposition 4.2. Thus

$$(A/2)k < \log(a-b) < \log a < B^{-1}k$$

which is false for $B > 2/A$.

REMARK: The statement of lemma 4.3 remains true if $\log a$ is replaced by $(\log a).(\log \log a)^{-1}$. (The extra details may be found in Langevin [1979], theorem 11.)

Let $\bar{w}^{k,d,q}$ denote the $\underset{k,d,q}{\sim}$ equivalence class to which w belongs.

LEMMA 4.4 *There exists* $k \in N$ *such that for any prime* q *and any number* y, *if* u *is the smallest number for which the elements of* $W = \{w : u \leqslant w \leqslant y\}$ *belong to distinct* $\underset{k,1,q}{\sim}$ *equivalence classes, and* $\leqslant$ *denotes the ordering induced on* $Y = \{\bar{w}^{k,1,q} : w \in W\}$ *by the ordering* $\leqslant$ *on* W, *then* *the triple* $\langle Y, \leqslant, q\rangle$ *determines* y *uniquely, that is, no other number* y *can give rise to the same triple in this way.*

Proof:

We will assume throughout that y is sufficiently large, since if the lemma is true for all $y > y_0$ with $k = k_0$, then it is true for all y with $k = \max\{k_0, y_0\}$. Under this assumption, it will be shown that the lemma holds for $k \geqslant 4$.

Suppose $y_1 < y_2$ both give rise to $\langle Y, \leqslant, q\rangle$ for some prime q, with u_1, W_2 and u_2, W_2 corresponding to y_1, y_2 respectively. Since $|W_1| = |Y| = |W_2|$, it follows that $y_1 - u_1 = y_2 - u_2 = m$ (say) and thus $W_1 = \{y_1, y_1 - 1, \dots, y_1 - m\}$, $W_2 = \{y_2, y_2 - 1, \dots, y_2 - m\}$. As the orderings on W_1, W_2 induce the same ordering on Y, we see that for each $i \leqslant m$, $\overline{y_1 - i}^{k,1,q} = \overline{y_2 - i}^{k,1,q}$ so $y_1 - i \overset{q}{\sim} y_2 - i$. Hence $y_1 \underset{m,1,q}{\sim} y_2$.

To obtain a contradiction by lemma 4.3 we need only show $m > B \log y_2$. This is the case for $k \geqslant 4$, since if $m < y_2$ then there is some $w \in W_2$ such that $w \underset{k,1,q}{\sim} u_2 - 1$, and therefore by lemma 4.1 (with $\varepsilon < \frac{1}{4}$),

$$u_2 - 1 < w - (\log w)^{4(1-\varepsilon)/3} \leqslant y_2 - (\log y_2)^{4(1-\varepsilon)/3}$$

so $$m = y_2 - u_2 > (\log y_2)^{4(1-\varepsilon)/3} - 1 > B \log y_2 .$$

We are now ready to investigate the definability properties of $L_{\leqslant,\perp}$. Clearly (cf. §1 and §2) each of the following predicates can be defined by bounded $L_{\leqslant,\perp}$ formula:

x *is squarefree*

$x \in P$ (that is, x *is prime*)

x *is squarefree* $\wedge$ $x|y$.

Observe also that if $d|a$ and $kd \leqslant a$ then $\{a, a-d, a-2d, \dots, a-kd\}$ consists of the largest $k+1$ elements of $\{w : w \leqslant a \wedge d|w\}$. Using this fact it is easily seen that for each fixed $k \in N$ there is a bounded $L_{\leqslant,\perp}$ formula which defines the predicate:

d *is squarefree* $\wedge$ $d|a \wedge a \underset{k,d,q}{\sim} b$.

Similarly, for each fixed $k \in N$ we can construct a bounded $L_{\leqslant,\perp}$ formula which defines the predicate:

$$d \textit{ is squarefree } \wedge d|a \wedge f_{k,d,q}(a) = \langle v_0, v_1, \ldots, v_k \rangle,$$

where $f_{k,d,q}$ is the mapping which "identifies" the equivalence class $\overline{a}^{k,d,q}$ with the $k+1$ tuple

$$\langle v_0, v_1, \ldots, v_k \rangle = \langle \prod_{\substack{p|a \\ p/q}} p, \prod_{\substack{p|a-d \\ p\nmid q}} p, \ldots, \prod_{\substack{p|a-kd \\ p\nmid q}} p \rangle$$

the elements of which are either squarefree or zero (adopting the convention that $\pi_{p\nmid q}\, p = 0$).

LEMMA 4.5 *The predicate* $q \in P \wedge q.y = z$ *can be defined by a bounded* $L_{\leqslant,\perp}$ *formula.*

Proof:

Fix k and suppose the prime $q|z$. Consider the triple $\langle Z, \leqslant, q\rangle$ constructed by taking u_* to be the least multiple of q such that all elements of

$$W_* = \{w: u_* \leqslant w \leqslant z \wedge q|w\} = \{z, z-q, z-2q, \ldots, u_*\}$$

belong to distinct $\underset{k,\tilde{q},q}{}$ equivalence classes, setting $Z = \{\overline{w}^{k,q,q}: w \in W_*\}$, and letting $\leqslant$ denote the ordering induced on Z by the natural ordering $\leqslant$ on W_*.

Now if $z = q.y$ and $\langle Y, \leqslant, q\rangle$ is the triple constructed from y as described in lemma 4.4, then identifying Y and Z with the corresponding sets of $k+1$ tuples of squarefree (or zero) numbers, it is clear that $\langle Y, \leqslant, q\rangle = \langle Z, \leqslant, q\rangle$. Also for each fixed k, there is obviously a bounded $L_{\leqslant,\perp}$ formula $\psi_k(q,y,z)$ which holds if and only if $q \in P$, $q|z$, and the triples $\langle Y, \leqslant, q\rangle$, $\langle Z, \leqslant, q\rangle$ constructed from y and z in the ways described above are equal.

But by lemma 4.4, y is uniquely determined by $\langle Y, \leqslant, q\rangle$ provided

k was chosen sufficiently large. Thus for some $k \in N$,

$$\forall q\,\forall y\,\forall z(\psi_k(q,y,z) \Leftrightarrow q \in P \wedge q.y = z)\ .$$

Recall that our goal is to prove:

THEOREM 4.6 *$z = x.y$ and $z = x+y$ can be defined by bounded $L_{\leqslant,\perp}$ formulas.*

In view of lemma 4.5 this is easily seen to be equivalent to the following theorem (the proof of which does not depend on the previous lemmas).

Let $\alpha(q,y,z)$ be the predicate defined by

$$\alpha(q,y,z) \Leftrightarrow q \in P \wedge q.y = z.$$

THEOREM 4.7 *$z = x.y$ and $z = x + y$ can be defined by bounded $L_{\leqslant,\alpha}$ formulas.*

Theorem 4.7 will be proved via several lemmas. The first two of these show that there are $L_{\leqslant,\alpha}$ formulas having all variables bounded by a parameter w, which define $x.y = z$ and $x + y = z$ provided w is somewhat larger than z, and which cannot be satisfied by any triple x,y,z not possessing the properties $x.y = z$, $x + y = z$ respectively.

LEMMA 4.8 *There is a bounded $L_{\leqslant,\alpha}$ formula $\phi_M(x,y,z,w)$ and constants $m, C > 0$ such that*

$$\forall w\,\forall x\,\forall y\,\forall z(\phi_M(w,x,y,z) \rightarrow x.y = z) \qquad (1)$$

$$\forall w\,\forall x\,\forall y\,\forall z(x.y = z \wedge z \leqslant w.(C \log w)^{-2m} \rightarrow \phi_m(x,y,z,w)). \qquad (2)$$

Proof:

Clearly there is a bounded $L_{\leqslant,\alpha}$ formula $\phi_M(x,y,z,w)$ which is satisfied if and only if either

$$((x = 0 \vee y = 0) \wedge z = 0) \vee (x = 1 \wedge y = 1 \wedge z = 1)$$

or $x \leqslant z$, $y \leqslant z$ and there exist squarefree numbers $d_1,d_2 \leqslant w$ such that

(i) $z < d_1 \wedge z < d_2 \wedge d_1 \perp d_2$

(ii) *For each prime* $p | d_1 d_2$ *there exist primes* q, r *with*

$q.x \leqslant w \wedge q.x \equiv 1 \pmod p$,

$r.y \leqslant w \wedge r.y \equiv 1 \pmod p$,

and $q.r.z \leqslant w \wedge q.r.z \equiv 1 \pmod p$.

For example, for p, q, r primes,

$q.r.z \leqslant w \wedge q.r.z \equiv 1 \pmod p \iff$

$\exists u \leqslant w \, \exists v \leqslant w \, (p | u-1 \wedge \alpha(q,v,u) \wedge \alpha(r,z,v))$.

Now suppose $\phi_M(x,y,z,w)$ holds. Then for each prime $p | d_1 d_2$,

$$q.r.x.y \equiv 1 \equiv q.r.z \pmod p$$

and therefore $x.y \equiv z \pmod p$ (since $q.r \perp p$). Hence $x.y \equiv z \pmod{d_1 d_2}$ and as both sides of this congruence are less than $d_1 d_2$ it follows that $x.y = z$. This establishes (1).

On the other hand, suppose $x.y = z > 1$. By proposition 4.2, if C is sufficiently large then

$$Cz^3 \log z < \prod_{p < C \log z} p \quad \text{and hence} \quad Cz^2 \log z < \prod_{\substack{p < C \log z \\ p \perp z}} p \, .$$

Therefore it is possible to choose coprime squarefree numbers d_1, d_2 such that $d_1 d_2 \perp z$, $z < d_i < Cz \log z$ (for $i = 1,2$) and all primes $p | d_1 d_2$ satisfy $p < C \log z$.

But by a theorem of Linnik (see, for example, Prachar [1957]) there is a constant $m > 1$ such that for any pair of numbers a, b with $a \perp b$ there is a prime $q < b^m$ satisfying $q \equiv a \pmod b$. Thus for each $p | d_1 d_2$ there exist primes q, $r < p^m < (C \log z)^m$ such that $q \equiv x^{-1} \pmod p$ and $r \equiv y^{-1} \pmod p$ where x^{-1} and y^{-1} denote natural numbers satisfying $x^{-1}.x \equiv 1 \pmod p$, $y^{-1}.y \equiv 1 \pmod p$, and hence $q.x \equiv 1 \pmod p$, $r.y \equiv 1 \pmod p$, and $q.r.z \equiv 1 \pmod p$.

Clearly if $z \leqslant w.(C \log w)^{-2m}$ then (provided C is large enough) $q.x$, $r.y$ and $q.r.z$ are all less than w, so $\phi_M(x,y,z,w)$ is satisfied.

Readers familiar with Robinson [1949] will notice that the proof of the above lemma is an adaption of the proof of the $L_{',|}$ definability of $x.y = z$ found in that paper. The essential difference is that here q and r are chosen to be primes so that only the predicate $\alpha(q,y,z)$ is required. A second difference is that to get a good upper bound on the least w which will work, the definition has been used modulo p for a sufficient number of small primes p. (Although it does simplify matters later, this second step is not absolutely necessary - without it theorems 4.6 and 4.7 could still be proved by iterating the construction used in lemma 4.10 below.)

Notice also that $z = x.y \Longleftrightarrow \exists w\, \phi_M(x,y,z,w)$. In view of Julia Robinson's observation that $z = x+y$ is $L_{=,',.}$ definable since for $z \neq 0$,

$$x+y = z \Longleftrightarrow (x.z+1).(y.z+1) = (x.y+1).z^2 + 1,$$

we have thus already proved (by a somewhat longer route than is necessary) that both $z = x.y$ and $z = x+y$ are $L_{\leqslant,\alpha}$ and hence $L_{\leqslant,\perp}$ definable. In order to proceed towards obtaining bounded definitions, we now adapt the $L_{=,',.}$ definition of addition just mentioned using much the same techniques as for the previous lemma.

LEMMA 4.9 *There is a bounded* $L_{\leqslant,\alpha}$ *formula* $\phi_A(x,y,z,w)$ *and constants* C *and* m (*the same as in lemma* 4.8) *such that*

$$\forall w\, \forall x\, \forall y\, \forall z\, (\phi_A(x,y,z,w) \rightarrow x+y = z) \tag{3}$$

$$\forall w\, \forall x\, \forall y\, \forall z\, (x+y = z \wedge z \leqslant w.(C \log w)^{-2m-1} \rightarrow \phi_A(x,y,z,w)). \tag{4}$$

Proof:

Note first that there is a bounded $L_{\leqslant,\alpha}$ formula $\psi(s,t,p,w)$

such that $\forall p \in P\ \forall s\ \forall t\ \forall w(\ \psi(s,t,p,w) \rightarrow s \equiv t \pmod p)$ (5)

and if $s,t \leqslant w.(C \log w)^{-2m-1}$ then for all primes $p \leqslant C \log w$,

$$s \equiv t \pmod p \rightarrow \psi(s,t,p,w). \tag{6}$$

For if p is prime then

$$s \equiv t \pmod p \iff (p|s \wedge p|t) \vee \exists v < p\ (p|v.s-1 \wedge p|v.t-1),$$

and replacing the two occurrences of $.$ in this second formula by the use of $\phi_M(\ldots,w)$ yields a predicate $\psi(s,t,p,w)$ which can obviously be defined by a bounded $L_{\leqslant,\alpha}$ formula. Applying lemma 4.8 shows that $\psi(s,t,p,w)$ has properties (5) and (6).

Now consider the expression:

$$\exists a < p\ \exists b < p\ \exists c < p\ (\ a \equiv x \pmod p) \wedge b \equiv y \pmod p) \wedge c \equiv z \pmod p)$$
$$\wedge\ (a.c+1).(b.c+1) \equiv (a.b+1).c.c+1 \pmod p)). \tag{7}$$

Let $\theta(x,y,z,p,w)$ be the bounded $L_{\leqslant,\alpha}$ formula obtained from this by replacing $.$ by $\phi_M(\ldots,w)$, and $\equiv$ by $\psi(\ldots,w)$. Take $\phi_A(x,y,z,w)$ to be a bounded $L_{\leqslant,\alpha}$ formula which asserts that there exists a squarefree number $d < w$ such that $d \perp z$, $2.x < d$, $2.y < d$, $z < d$, and $\forall p \in P\ (p|d \rightarrow \theta(x,y,z,p,w))$.

If $\phi_A(x,y,z,w)$ holds then for every prime $p|q$, $(x.z+1).(y.z+1) \equiv (x.y+1).z^2+1 \pmod p$ and therefore since $z \perp p$, $x+y \equiv z \pmod p$. Hence $x+y \equiv z \pmod d$ and thus $x+y = z$.

On the other hand if $x+y = z$ then (7) is clearly satisfied for all primes p, and thus $\theta(x,y,z,p,w)$ holds for all w,p with $p < C \log z$, $z < w(C \log w)^{-2m-1}$. Also (as in the proof of the previous lemma) there is a squarefree number $d \Big| \prod_{\substack{p<C\log z \\ p\perp z}} p$ such that $2.z < d < w$, so $\phi_A(x,y,z,w)$ is satisfied.

Let $\underset{w}{+}$, $\underset{w}{.}$ denote the partial functions (extendible to $+,.$) defined by

$$x \underset{w}{+} y = z \iff \phi_A(x,y,z,w)$$

$$x \underset{w}{.} y = z \iff \phi_M(x,y,z,w)$$

where ϕ_A, ϕ_M are the formulas constructed in lemmas 4.8 and 4.9. Denote by $h(w)$ the largest number $h \leqslant w$ such that for all $x,y,z \leqslant h$,

$$(x+y = z \to x \underset{w}{+} y = z) \wedge (x.t = z \to x \underset{w}{.} y = z).$$

For each w, the interval $[0,h(w)]$ is the largest initial segment of $[0,w]$ on which ϕ_A, ϕ_M correctly define the addition and multiplication predicates A,M. Our problem is to "extend" these definitions to the whole of $[0,w]$.

The predicate $v = h(w)$ can be defined by a bounded $L_{\leqslant,\alpha}$ formula. Indeed if $\underset{w}{+}, \underset{w}{\cdot}$ are commutative and $0 \underset{w}{.} y = 0$ for all y (as is the case with the definitions ϕ_A, ϕ_M actually constructed in the proofs above), then it is obvious that $h(w)$ is the largest number $h \leqslant w$ such that

(i) $\forall z \leqslant h\ \forall x \leqslant z\ \exists y \leqslant z\ (z = x \underset{w}{+} y)$

(ii) $\forall z \leqslant h\ \forall x \leqslant z\ (x \neq 0 \to \exists q \leqslant z\ \exists r < x\ (z = x \underset{w}{.} q \underset{w}{+} r))$.

((i) and (ii) say that subtraction and division work.)

Also by lemmas 4.8 and 4.9, $h(w) \geqslant w(C \log w)^{-2m-1}$. Applying Bertrand's postulate (see for example Hardy and Wright [1979] or chapter 1 of this thesis) that for every $x \geqslant 2$ there is a prime p with $x < p < 2x$, it follows that if w is sufficiently large then there is a prime p such that

$$h(w) \geqslant w(C \log w)^{-2m-1} \geqslant 2\sqrt{w} > p > \sqrt{w} \quad .$$

But for w less than any fixed bound, it is clearly possible to "write out" the addition and multiplication tables up to w as bounded $L_{\leqslant}$ formulas and incorporate these in ϕ_A and ϕ_M, so it can be assumed that for every $w > 1$ there is a prime p with $\sqrt{w} < p \leqslant h(w)$.

The technique used in the proof of proposition 0.2 allows us to define the addition and multiplication relations on an isomorphic copy $[0,a^2]^*$ of the interval $[0,a^2]$ by means of $L_{\leqslant,A,M}$ formulas in which all quantified variables are bounded by a. (The elements of $[0,a^2]^*$ are 4 tuples of numbers $\leqslant [\sqrt{a}]$ treated as the digits of the base $[\sqrt{a}]+1$ representation of a single number.) Replacing A,M by ϕ_A,ϕ_M we see that there are bounded $L_{\leqslant,\alpha}$ formulas $\phi_A^*(X,Y,Z,w,a)$, $\phi_M^*(X,Y,Z,w,a)$ (where X,Y,Z denote 4 tuples) such that

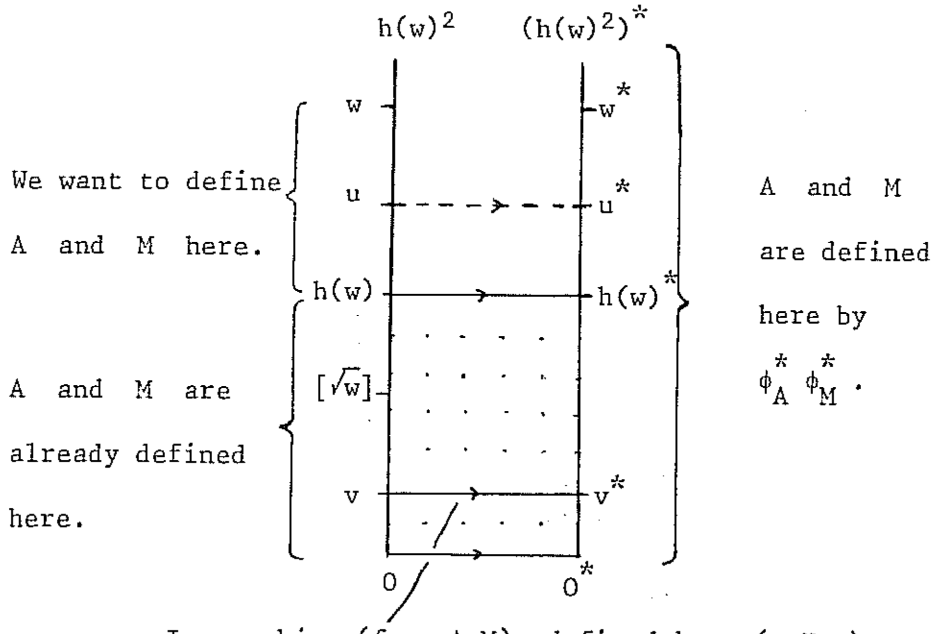
$$\phi_A^*(X,Y,Z,w,h(w)) \iff [0,h(w)^2]^* \models X+Y = Z$$

$$\phi_M^*(X,Y,Z,w,h(w)) \iff [0,h(w)^2]^* \models X\,.\,Y = Z.$$

Similarly, if for $u \in [0,w]$ we use u^* to denote the corresponding element of $[0,h(w)^2]^*$, then there is a bounded $L_{\leqslant,\alpha}$ formula $\eta(u,U,w)$ such that for all $u \in [0,h(w)^2]^*$,

$$\eta(u,U,w) \iff U = u^*.$$

Thus η defines the isomorphism $u \to u^*$ for $u \leqslant h(w)$.


h(w)²
(h(w)²)*
w
w*
We want to define A and M here.
u
u*
A and M are defined here by φ*A φ*M .
h(w)
h(w)*
A and M are already defined here.
[√w]
v
v*
0
0*

Isomorphism (for A,M) defined by $\eta(u,U,w)$.

Our strategy is to "extend" the definition given by $\eta(u,U,w)$ by constructing a new bounded $L_{\leqslant,\alpha}$ formula $\eta^*(u,U,w)$ which defines the isomorphism $u \to u^*$ for all $u \leqslant w$.

LEMMA 4.10 *There is a bounded* $L_{\leqslant,\alpha}$ *formula* $\eta^*(u,U,w)$ *such that*

$$\forall w\ \forall u \in [0,w]\ \forall U \in [0,h(w)^2]^*\ (\ U = u^* \Leftrightarrow \eta^*(u,U,w)).$$

Proof:

Essentially what we will do is to use bounded $L_{\leqslant,\alpha}$ formulas to associate certain elements and subsets of $[0,h(w)]$ with each $u \in [0,w]$ in such a way that these elements and subsets of $[0,h(w)]$ determine u uniquely in the interval $[0,h(w)^2]$, and thus their images under $*$ (which is already defined on $[0,h(w)]$ by $\eta(u,U,w)$) will completely determine u^* in $[0,h(w)^2]^*$. (The novelty of the argument lies in the use of sets, rather than just a fixed finite number of elements, to determine u.)

It is claimed that for all $u \in [0,w]$, $U \in [0,h(w)^2]^*$, $U = u^*$ if and only if U is the least element of $[0,h(w)^2]^*$ possessing the following three properties:

(I) *If* p *is the largest prime* $\leqslant h(w)$ *and* t *is the largest number such that* $p.t < u$, *then* t^* *is the largest element of* $[0,h(w)^2]^*$ *for which* $[0,h(w)^2]^* \models p^*.t^* < U$.

(II) $\exists s \leqslant u\ (2.s = u) \Leftrightarrow [0,h(w)^2]^* \models \exists s \leqslant U\ (2.S = U)$.

(III) *For* $i = 0,1,2$ *and all primes* $q \leqslant h(w)$,

$$\exists s \leqslant u\ (q.(3.s+1) < u \wedge \neg q.(3.s+i+1) < u) \Leftrightarrow$$
$$[0,h(w)^2]^* \models \exists s \leqslant U\ (q^*.(3^*.S + i^*) < U \wedge \neg q^*.(3^*.S + i^* + 1^*) < U).$$

Taking $U = u^*$ obviously satisfies (I), (II), (II), so it suffices to show that no $U < u^*$ has these three properties. Suppose $U < u^*$ satisfies (I), (II) and (III). Since $U < u^*$, we must have $U = v^*$ for some $v < u$. By (II), $2 \leqslant u - v$, while by (I), $pt < v < u \leqslant p(t+1)$ so $u - v < p \leqslant h(w)$. Hence by Bertrand's postulate there is some prime $q \leqslant h(w)$ such that $q \leqslant u - v < 2q$.

But now suppose $qs_1 < u \leqslant q(s_1 + 1)$ and $qs_2 < v \leqslant q(s_2 + 1)$. Then

$$qs_2 + q < v + (u - v) < q(s_2 + 1) + 2q$$

so $q(s_2 + 1) < u < q(s_2 + 3)$ and thus $s_1 = s_2 + 1$ or $s_1 = s_2 + 2$. In either case $s_1 \not\equiv s_2 (\text{mod } 3)$ so (III) fails.

Note that in (I), (II), (III) all terms of the form x^* are restricted to $x \leqslant h(w)$. (In this case of (I) to get $t \leqslant h(w)$ we make use of the assumption, justified above, that $\sqrt{w} < p \leqslant h(w)$ and therefore $t < \sqrt{w} < h(w)$.) Consequently since $U = u^*$ is defined on $[0,h(w)]$ by the formula $\eta(u,U,w)$, and addition and multiplication predicates are defined on $[0,h(w)^2]^*$ by ϕ_A^*, ϕ_M^*, all expressions in (I), (II), (III) beginning with $[0,h(w)^2]^* \models$ can be replaced by bounded $L_{\leqslant,\alpha}$ formulas. Since all occurrences of $.$ in (I), (II), (III) involve multiplication by a prime, and there is no nontrivial use of $+$, it is now clear that a suitable bounded $L_{\leqslant,\alpha}$ formula $\eta^*(u,U,w)$ can be constructed.

Completion of the proof of theorems 4.6 and 4.7:

To complete the proof that $x + y = z$ and $x.y = z$ can be defined by bounded $L_{\leqslant,\alpha}$ formulas (and therefore also by bounded $L_{\leqslant,\perp}$ formulas) it is only necessary to observe that for all $x,y \leqslant z$,

$x + y = z \iff$

$$\exists X \leqslant x \; \exists Y \leqslant y \; \exists Z \leqslant z \; (X = x^* \wedge Y = y^* \wedge Z = z^* \wedge \phi_A^*(X,Y,Z,z))$$

$x . y = z \iff$

$$\exists X \leqslant x \; \exists Y \leqslant y \; \exists Z \leqslant z \; (X = x^* \wedge Y = y^* \wedge Z = z^* \wedge \phi_M(X,Y,Z,z))$$

where $U = u^*$ is being used as an abbreviation for $\eta^*(u,U,z)$, and X,Y,Z are 4 tuples of variables each of which is bounded by z.

REMARK: The use of Linnik's (or some weaker) bound on the least prime $p \equiv a \pmod{b}$ in proving the above theorems can be avoided by appealing to the analogous result for squarefree numbers, which, at least at present, is easier to prove. (See Pracher [1958] or Erdös [1960].) The argument is modified as follows: having first proved lemma 4.5, that lemma is used to give a similar proof that the predicate

$$q \textit{ is squarefree} \wedge q.y = z$$

is definable by a bounded $L_{\leqslant,\perp}$ formula. The proofs of theorems 4.7 and 4.8 can then be reworked using this predicate instead of α, and squarefree numbers instead of primes at the appropriate places. (The original theorem 4.8 can of course be recovered as a corollary of theorem 4.7.)

Applying proposition 0.2 to theorem 4.7 yields:

<u>COROLLARY 4.11</u> *A relation on* N *can be defined by a bounded* $L_{\leqslant,+,\cdot}$ *formula if and only if it can be defined by a bounded* $L_{\leqslant,\perp}$ *formula.*

In order to deduce the same result with a single binary relation in place of $\leqslant,\perp$ we will use the following approximation to Goldbach's conjecture obtained by the use of "sieve" methods. (a readable introduction to these can be found in Gel'fond and Linnik [1962].)

<u>PROPOSITION 4.12</u> (Brun [1920]) *There is a number* $k \in N$ *with the property that for every* $n \in N$ *there exist numbers* q,r *with* $2n = q + r$, *such that every prime* $p|qr$ *satisfies* $p \geqslant (2n)^{1/k}$.

Define the preordering $\overset{*}{<}$ by:

$$x \overset{*}{<} y \iff x < y \wedge x \perp y \quad .$$

The reader is reminded that $\overset{*}{<}$ is <u>not</u> transitive. ($x \overset{*}{<} y \overset{*}{<} z$ should be read as $x \overset{*}{<} y \wedge y \overset{*}{<} z$.)

<u>LEMMA 4.13</u> *There is a number* $g \in N$ *such that for all* x,y *with* $x \overset{*}{<} y$ *there exist numbers* $z_0,z_1,\ldots,z_m$ *with* $m \leqslant g$ *satisfying:*

$$x = z_0 \overset{*}{<} z_1 \overset{*}{<} z_2 \overset{*}{<} \ldots \overset{*}{<} z_m = y.$$

Furthermore, the numbers z_i *may be chosen to be alternately even and odd.*

Proof:

Take $g = (2k-2)^k + 2k$ where k satisfies proposition 4.12. If $y - x = h \leqslant g$ we have $x \overset{*}{<} x+1 \overset{*}{<} x+2 \overset{*}{<} \ldots \overset{*}{<} x+h = y$ so we may suppose $y - x \geqslant g = (2k-2)^k + 2k$. Let

$$2n = \begin{cases} (y-x) - (2k-2) & \text{if this is even,} \\ (y-x) - (2k-2) - 1, & \text{otherwise,} \end{cases}$$

so $2n > (2k-2)^k$. By proposition 4.12 there exist q,r such that $2n = q + r$ and all primes $p | qr$ satisfy $p \geqslant (2n)^{1/k} > 2k - 2$. (In particular, since $k \geqslant 2$ we have $p > 2$, so q,r are both odd.) Observe that each prime $p | q$ can divide at most one of the numbers $x, x+1, \ldots, x+2k-2$, and that q has fewer than k prime divisors. It follows that $x + i \perp q$ for at least k of the numbers $i = 0,1,\ldots,2k-2$. Similarly, $x + 2n + i \perp r$ for at least k of these i's. Hence there is some $i \leqslant 2k-2$ such that $x + i \perp q$ <u>and</u> $x + 2n + i \perp r$. Taking $z = x + i + q = x + 2n + i - r$ exhibits a number satisfying $x + 1 \overset{*}{<} z \overset{*}{<} x + 2n + i$ and having the opposite parity to $x + i$. Thus the sequence

$$x \overset{*}{<} x+1 \overset{*}{<} x+2 \overset{*}{<} \ldots \overset{*}{<} x+i \overset{*}{<} z \overset{*}{<} x+2n+i \overset{*}{<} x+2n+i+1 \overset{*}{<} \ldots$$
$$\ldots \overset{*}{<} y-2 \overset{*}{<} y-1 \overset{*}{<} y$$

of at most $2k + 2$ elements has the required properties.

PROBLEM: *What is the smallest value of* g *for which this lemma is true*? (The author conjectures: *If* $y - x > 1$ *then some* z *in the interval* $x < z < y$ *is coprime to both* x *and* y.)

<u>THEOREM 4.14</u> *A relation on* N *can be defined by a bounded* $L_{\leqslant,+,\cdot}$

formula if and only if it can be defined by a bounded $L_{\overset{*}{\leq}}$ *formula, where* $\overset{*}{\leq}$ *is the preordering defined by:*

$$x \overset{*}{\leq} y \iff x = y \vee (x \leq y \wedge x \perp y)$$

Proof:

Obviously, $x = y \iff x \overset{*}{\leq} y \wedge y \overset{*}{\leq} x$

$$x \perp y \iff \neg x = y \wedge (x \overset{*}{\leq} y \vee y \overset{*}{\leq} x)$$

and for g satisfying lemma 4.13,

$$x \leq y \iff \exists z_{g-1} \overset{*}{\leq} y \; \exists z_{g-2} \overset{*}{\leq} z_{g-1} \ldots \exists z_1 \overset{*}{\leq} z_2 \; (x \overset{*}{\leq} z_1)$$

$$\forall x \leq y \; \phi \iff \forall z_{g-1} \overset{*}{\leq} y \; \forall z_{g-2} \overset{*}{\leq} z_{g-1} \ldots \forall z_1 \overset{*}{\leq} z_2 \; \forall x \overset{*}{\leq} z_1 \; \phi$$

$$\exists x \leq y \; \phi \iff \exists z_{g-1} \overset{*}{\leq} y \; \exists z_{g-2} \leq z_{g-1} \ldots \exists z_1 \overset{*}{\leq} z_2 \; \exists x \overset{*}{\leq} z_1 \; \phi$$

where the variables $z_1, z_2, \ldots, z_{q-1}$ are chosen from those which do not occur in ϕ.

Hence the theorem follows from corollary 4.11 by induction on the complexity of bounded $L_{\leq,\perp}$ formulas.

As mentioned in the introduction, a relation on N is *rudimentary* if and only if it can be defined by a bounded $L_{\leq,+,\cdot}$ formula. By a *rudimentary graph* G we will mean a structure $\langle N, \Upsilon \rangle$ where Υ is a two place, symmetric, antireflexive, rudimentary relation. (Vertices $x, y \in N$ are joined by an edge if and only if $x \Upsilon y$.) Let G_n denote the restriction of G to $[0,n]$.

Now consider the rudimentary graph $G = \langle N, \Upsilon \rangle$ where Υ is the symmetric relation defined by requiring $x \Upsilon y$ and $y \Upsilon x$ to hold between numbers x, y <u>with $x \leq y$</u> if and only if one of the following is satisfied:

(i) x *and* y *are both odd and* $x \perp y$.

(ii) x *is odd,* y *is even, and* $x \perp y$.

(iii) $x = 0$, $y \neq 0$, *and* y *is even.*

THEOREM 4.15 *The rudimentary graph* G *defined above has the property that for every rudimentary predicate* $\rho(x_1,\ldots,x_k)$ *there is an effectively found* L_γ *formula* $\phi(x_1,\ldots,x_k)$ *such that for all* $n \geq 5$, *and all* $x_1,\ldots,x_k \in [0,n]$,

$$\rho(x_1,\ldots,x_k) \iff G_n \models \phi(x_1,\ldots,x_k).$$

Proof:

By corollary 4.11, a predicate $\rho(x_1,\ldots,x_k)$ is rudimentary if and only if it can be defined by a bounded $L_{\leqslant,\perp}$ formula, so the theorem will follow by induction on the complexity of such formulas if it can be shown that there are L_γ formulas $\phi_\leqslant(x,y)$, $\phi_\perp(x,y)$ such that for all $n \geq 5$, and all $x,y \in [0,n]$,

$$x \leqslant y \iff G_n \models \phi_\leqslant (x,y)$$

$$x \perp y \iff G_n \models \phi_\perp (x,y).$$

To avoid continually writing "$G_n \models$", for the rest of this proof variables x,y,z etc. will range over the elements of some fixed initial segment of N containing $\{0,1,2,3,4,5\}$. (These initial segments are the intervals $[0,n]$, $n \geq 5$, and N itself.) However the formulas produced will be independent of which initial segment is used. We begin with some less ambitious definitions:

(i) $x = y \iff \forall z(x\gamma z \iff y\gamma z)$

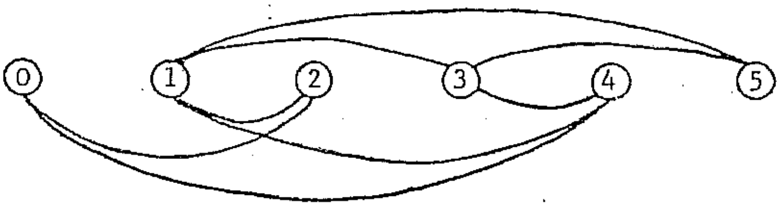
(ii) $x = 1 \iff \exists y(\neg y = x \wedge \forall z(\neg z\gamma x \iff z = y \vee z = x))$

(iii) $x = 0 \iff \neg x\gamma 1 \wedge \neg x = 1.$

To verify (i) check that for all x and y with $x < y$ (say), $(x\gamma z \iff y\gamma z)$ fails for $z = 0$ (if x,y have opposite parity), for $z = y-1$ (if x,y are both even), or for $z = x+1$ (if x,y are both odd). (i) and (ii) follow from the fact that for $x = 1$ there is a unique number $z \neq x$ (namely $z = 0$) such that $\neg z\gamma x$, but this cannot happen for $x \neq 1$, since if x is odd then $\neg z\gamma x$ holds for $z = 0,2,$

while if x is even then $\neg z\Upsilon x$ holds (with $z \neq x$) for $z = 2,4$ except for the cases $x = 0,2,4$ which can be dealt with by examining the picture:

G_5:



Note now that x *is even* $\Longleftrightarrow x = 0 \vee x\Upsilon 0$ and if x,y are both odd then $x \perp y \Longleftrightarrow x\Upsilon y$.

Also, x *is a power of an odd prime* $\Longleftrightarrow$

$$x \textit{ is odd } \wedge \forall y \forall z\,(\, y,z \textit{ are odd } \wedge y \perp z \rightarrow y \perp z \vee z \perp x),$$

and if y is odd then

$$x \sim y \Longleftrightarrow x \textit{ is odd } \wedge \forall z(\, z \textit{ is odd } \rightarrow (z \perp x \Longleftrightarrow z \perp y)).$$

Using these equivalences we can define the odd primes, since

$$x \textit{ is an odd prime} \Longleftrightarrow$$
$$x \textit{ is a power of an odd prime } \wedge x \neq 1 \wedge$$
$$\forall z(z \neq x \wedge z \sim x \rightarrow \exists w(x\Upsilon w \wedge \neg z\Upsilon w)).$$

(To see this take w to be the even number $z-1$.)

Note that the predicates $y = 2$, $x = 3$ can be defined since $x = 3$ if and only if x is odd and there exists a <u>unique</u> even number $y \neq 0$ (namely $y = 2$) such that $x\Upsilon y$. Also <u>for x even,</u>

$$3 \perp x \Longleftrightarrow x = 2 \vee 3\Upsilon x\ .$$

Now suppose x is even, $3 \perp x$, and p is an odd prime $\neq 3$. Then

$$p \leqslant x \Longleftrightarrow p\Upsilon x \vee \psi(p,x)\ ,$$

where $\psi(p,x)$ is the formula:

$$\exists z\, \exists w(\; z \textit{ is odd } \wedge w \textit{ is even } \wedge p\Upsilon w \wedge \neg z\Upsilon w \wedge z\Upsilon x$$
$$\wedge\, \forall q(\; q \textit{ is an odd prime } \wedge \neg\, q \perp z \rightarrow q\Upsilon w\;))$$

Clearly $p\Upsilon x \Rightarrow p \leqslant x$. Also if z,w are as required by $\psi(p,x)$

then $z \perp w$ but $\neg z\Upsilon w$ so $w < z$, and since $p\Upsilon w$ and $z\Upsilon x$ it follows that $p < w < z < x$.

On the other hand, suppose $p \leqslant x$. If $p \nmid x$ then $p\Upsilon x$. If $p | x$ we have two possibilities:

Case 1 $p < 3^n < x$ for some $n > 1$.

Then either $p\Upsilon(3^n - 1)$ or $p\Upsilon(3^n - 5)$ so choosing $w = 3^n - 1$ or $3^n - 5$ as appropriate, and $z = 3^n$ satisfies $\psi(p,x)$.

Case 2 $3^n < p < x < 3^{n+1}$ for some $n \geqslant 1$.

Then $\frac{x}{p} < 3$, but since $p|x$ it follows that $x = 2p$. Either $2p - 1$ or $2p - 5$ is divisible by 3 (and if $p = 5$ then it is the former) so choosing z to be that one of these, and

$$w = z - 1 \geqslant \begin{cases} 2p - 6 > p & \text{if } p \geqslant 7, \\ 2p - 2 > p & \text{if } p = 5, \end{cases}$$

satisfies $\psi(p,x)$ since $p < w < 2p$, and thus $p \perp w$.

Observe that

$$x \textit{ is a power of } 2 \Leftrightarrow x \textit{ is even } \wedge 3 \perp x \wedge$$
$$\forall q(\, q \textit{ is an odd prime} \neq 3 \wedge q \leqslant x \rightarrow q\Upsilon x\,).$$

Now if x is even and p is an odd prime, then

$$p \leqslant x \Leftrightarrow p\Upsilon x \vee$$
$$\exists z\, \exists w(\, w \textit{ is a power of } 2 \wedge z \textit{ is odd} \wedge p\Upsilon w \wedge z\Upsilon x \wedge \neg z\Upsilon w\,).$$

For clearly $p\Upsilon x \rightarrow p \leqslant x$, while if the second disjunct of this definition is satisfied then $p < 2^n < z < x$ for some n,z. Conversely if $p \leqslant x$ and $p \nmid n$ then $p\Upsilon x$, while if $p|x$ then $2p \leqslant x$ so there exists n with $p < 2^n < 2p \leqslant x$, and taking $z = x-1$, $w = 2^n$ satisfies the second disjunct above.

Also for x even and p an odd prime,

$$p \perp x \Leftrightarrow p\Upsilon x \vee \neg p \leqslant x.$$

Finally if x (say) is even and y is odd, then

$$x \perp y \iff \forall p(\ p \text{ } \mathit{is\ an\ odd\ prime} \rightarrow p \perp x \vee p \perp y),$$

and since the case where x,y are both odd has already been dealt with, the existence of an L_γ formula $\phi_\perp(x,y)$ satisfying

$$\forall x \forall y(x \perp y \iff \phi_\perp (x,y))$$

clearly follows.

We next define for each fixed natural number $k \geq 1$ a preordering $\underset{k}{\leq}$ on the even natural numbers by

$$x \underset{1}{\leqslant} y \Leftrightarrow x = y \vee \exists z(x \perp z \wedge z\Upsilon y \wedge \neg z\Upsilon x),$$

$$x \underset{k+1}{\leqslant} y \Leftrightarrow x \underset{k}{\leqslant} y \vee \exists z(z \textit{ is even} \wedge x \underset{1}{\leqslant} z \wedge z \underset{k}{\leqslant} y).$$

In the notation of lemma 4.11, for x,y both even we have $x \underset{k}{\leqslant} y$ if and only if for some $m \leqslant k$ there exists a sequence of alternatively even and odd numbers $z_0,z_1,\ldots,z_{2m}$ such that $x = z_0 \overset{*}{<} z_1 \overset{*}{<} \ldots \overset{*}{<} z_{2m} = y$. Thus if k is sufficiently large then for all even numbers x,y,

$$x \leqslant y \Leftrightarrow x \underset{k}{\leqslant} y.$$

Now for x odd and y even,

$$x \leqslant y \Leftrightarrow \exists z(z \textit{ is even} \wedge x\Upsilon z \wedge z \leqslant y)$$

(as can be seen by taking $z = x+1$ if this exists) and $y \leqslant x \Leftrightarrow \neg x \leqslant y$.

Finally, if x and y are both odd,

$$x \leqslant y \Leftrightarrow x = y \vee \exists z(z \textit{ is even} \wedge x \leqslant z \wedge z \leqslant y).$$

Since all cases have now been covered it is clear that there is an L_Υ formula $\phi_{\leqslant}(x,y)$ such that

$$\forall x \forall y(x \leqslant y \Leftrightarrow \phi_{\leqslant}(x,y)) \quad .$$

REMARK: With the exception of the trivial case $n = 0$, the restriction $n \geqslant 5$ in the above theorm is unavoidable since every graph with at least 2 and at most 5 vertices admits a nontrivial automorphism which obviously cannot preserve $\leqslant$.

To complete this section we will sketch the proof of the analogue of theorem 4.15 for a rudimentary partial ordering.

<u>THEOREM 4.16</u> *There is a rudimentary partial ordering* $H = \langle N, \textcircled{$\leqslant$}\rangle$ *such that* $\textcircled{$\leqslant$}$ *is extendible to* $\leqslant$, *and if* H_n *denotes the restriction of* H *to* $[0,n]$, *then for every rudimentary predicate* $\rho(x_1,\ldots,x_m)$ *there is an* $L_{\textcircled{$\leqslant$}}$ *formula* $\phi(x_1,\ldots,x_m)$ *with the property that for*

all $n \in N$ *and all* $x_1,\ldots,x_m \leq n$,

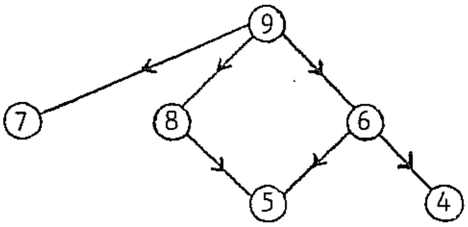
$$\rho(x_1,\ldots,x_m) \iff H_m \models \phi(x_1,\ldots,x_m).$$

Proof:

As the reader should by now be well versed in the sort of techniques used, only a sketch will be provided.

To define $\text{\textcircled{$\le$}}$ we will construct a rudimentary directed graph having N as its set of points, with the property that there is a directed edge from y to x if and only if x is an immediate $\text{\textcircled{$\le$}}$-predecessor of y. Then we will have $z \text{\textcircled{$\le$}} y$ if and only if $x = y$ or there is a directed path from y to x. (Some care is needed here as "path" is a second order concept. However the paths in the directed graph constructed will be "trivial enough" to allow the existence of a path from y to x to be described easily by an $L_{\leq,+,\cdot}$ formula, thus making $\text{\textcircled{$\le$}}$ rudimentary.)

The basic constituents of the directed graph will be of the form:



and will be called *sextets*. The points in the $4,5,\ldots,9$ positions (called i-points, $i = 4,5,\ldots,9$) will be consecutive numbers (greater than 3) congruent to $4,5,\ldots,9 \pmod 6$ respectively.

The idea is to ensure that $\leq,\perp$ can be recovered from $\text{\textcircled{$\le$}}$ on $[0,n]$. For example, for $n \equiv 9 \pmod 6$ it will be possible to define $\leq$ since it will be arranged that:

(i) The predicate "x is a 4-point" is $L_{\text{\textcircled{$\le$}}}$ definable in H_m.

(ii) $\text{\textcircled{$\le$}}$ linearly orders the 4-points.

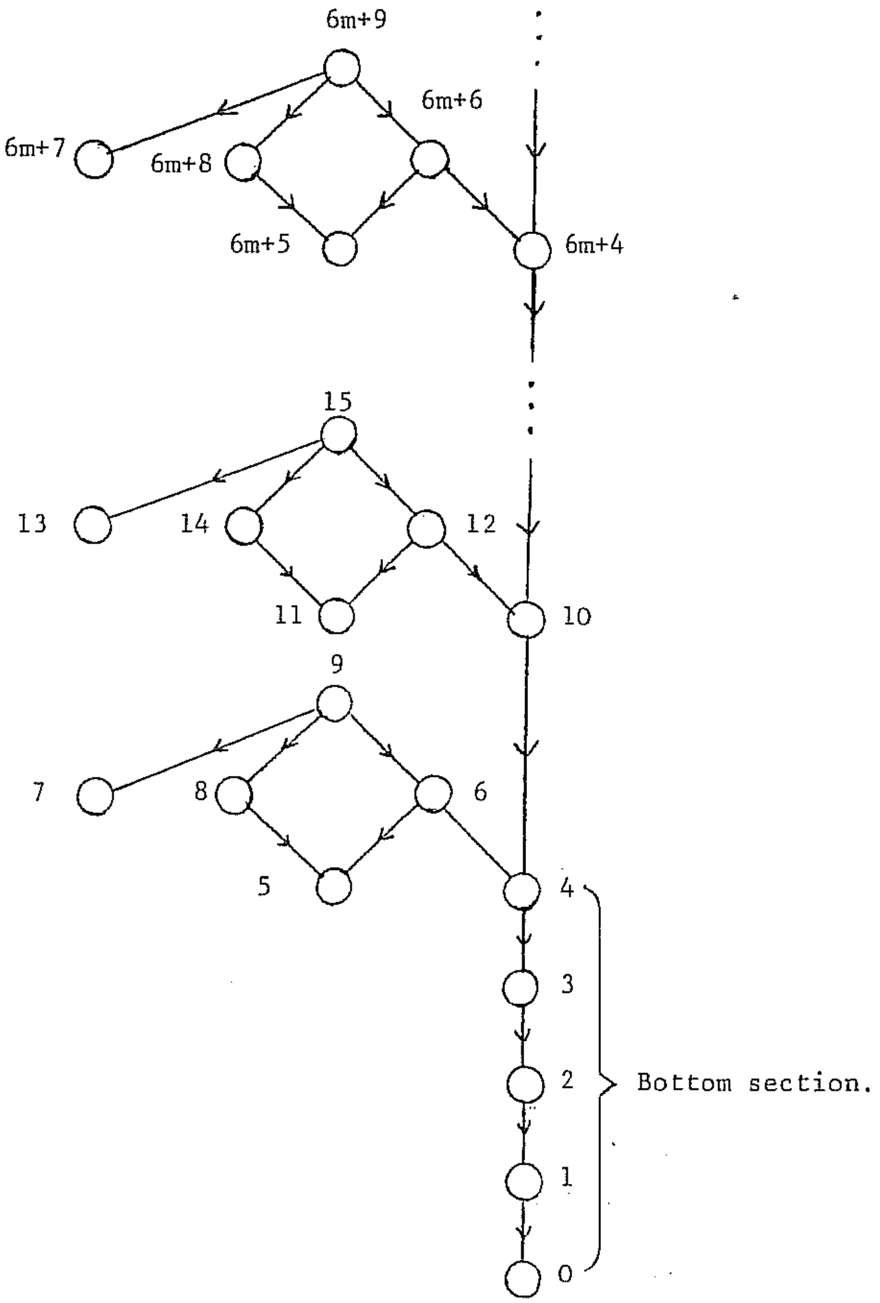
(iii) This fact can be used to linearly order the 9-points (which will be the maximal points of the $\textcircled{\leq}$ ordering).

(iv) The sextet to which a point x belongs is then determined uniquely by the smallest 9-point above x.

As the asymmetry of the sextet allows $\leq$ to be defined on the sextet, conditions (i)-(iv) will make it possible (for $n \equiv 9 \pmod 6$) to give an $L_{\textcircled{\leq}}$ definition of $\leq$ on $[0,n]$ which is independent of n.

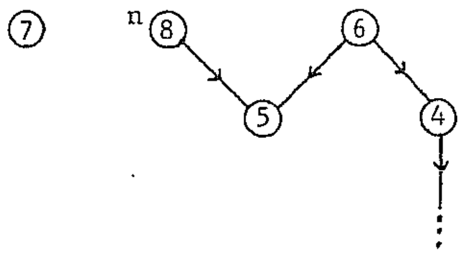
Now consider the partial ordering $H^o = \langle N, \textcircled{\leq} \rangle$ corresponding to the following directed graph:

H^o:

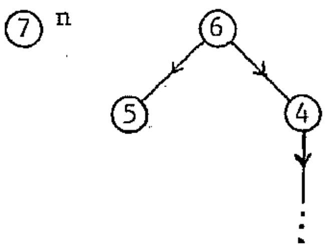


Clearly this partial ordering satisfies (i)-(iv) for $n \equiv 9 \pmod 6$. The addition of the special bottom section makes it easy to comply with (i). (Hint: first define the predicate $x = 0$.)

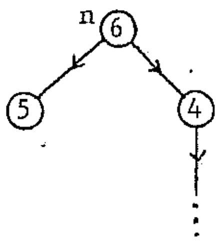
If $n \equiv 9 \pmod 6$ then only part of the uppermost sextet is included in H_m^0. These "partial sextets" are of the forms:



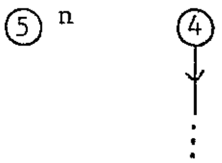
$n \equiv 8 \pmod 6$



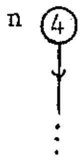
$n \equiv 7 \pmod 6$



$n \equiv 6 \pmod 6$



$n \equiv 5 \pmod 6$



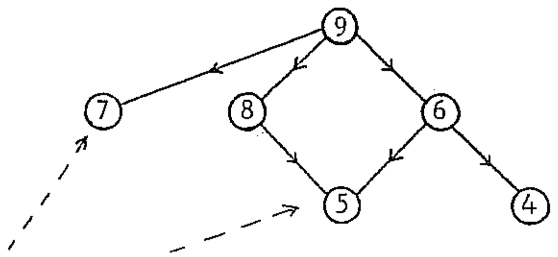
$n \equiv 4 \pmod 6$

$\leq$ can still be defined on H_n^0 in these cases (this is left for the reader to check) and hence there is a single $L_{\textcircled{\le}}$ formula which defines $\leq$ on H_n^0 in all cases.

However $\perp$ cannot be defined on all H_n^0 by a single $L_{\textcircled{\le}}$ formula (as can easily be seen by considering nonstandard models), so to produce a partial ordering $H = \langle N, \textcircled{\le} \rangle$ such that $\perp$ can be defined on all H_n by a single $L_{\textcircled{\le}}$ formula we will add new edges to the directed graph for H^0 while at the same time preserving the definition of $\leq$ (for example, in the case of $n \equiv 9 \pmod 6$ preserving (i)-(iv)).

If $p (\neq 2,3)$ is a prime less than x and $p | x$ then we will add an edge from x's sextet (or two edges from the preceeding sextet) to the point p in such a way that information about exactly which element(s) of x's sextet is (are) multiple(s) of p is coded into the corresponding partial ordering $\textcircled{\leq}$.

The only points in a sextet which can be on the "receiving" end of these additional edges will be those which are primes, and these can only be 5-points or 7-points (since if $p \geq 5$ is prime then $p \equiv 5 \pmod 6$ or $p \equiv 7 \pmod 6$). These prime points will not be on the "transmitting" end of any additional edges and will constitute exactly the minimal points of the $\textcircled{\leq}$ partial ordering (aside from 0).



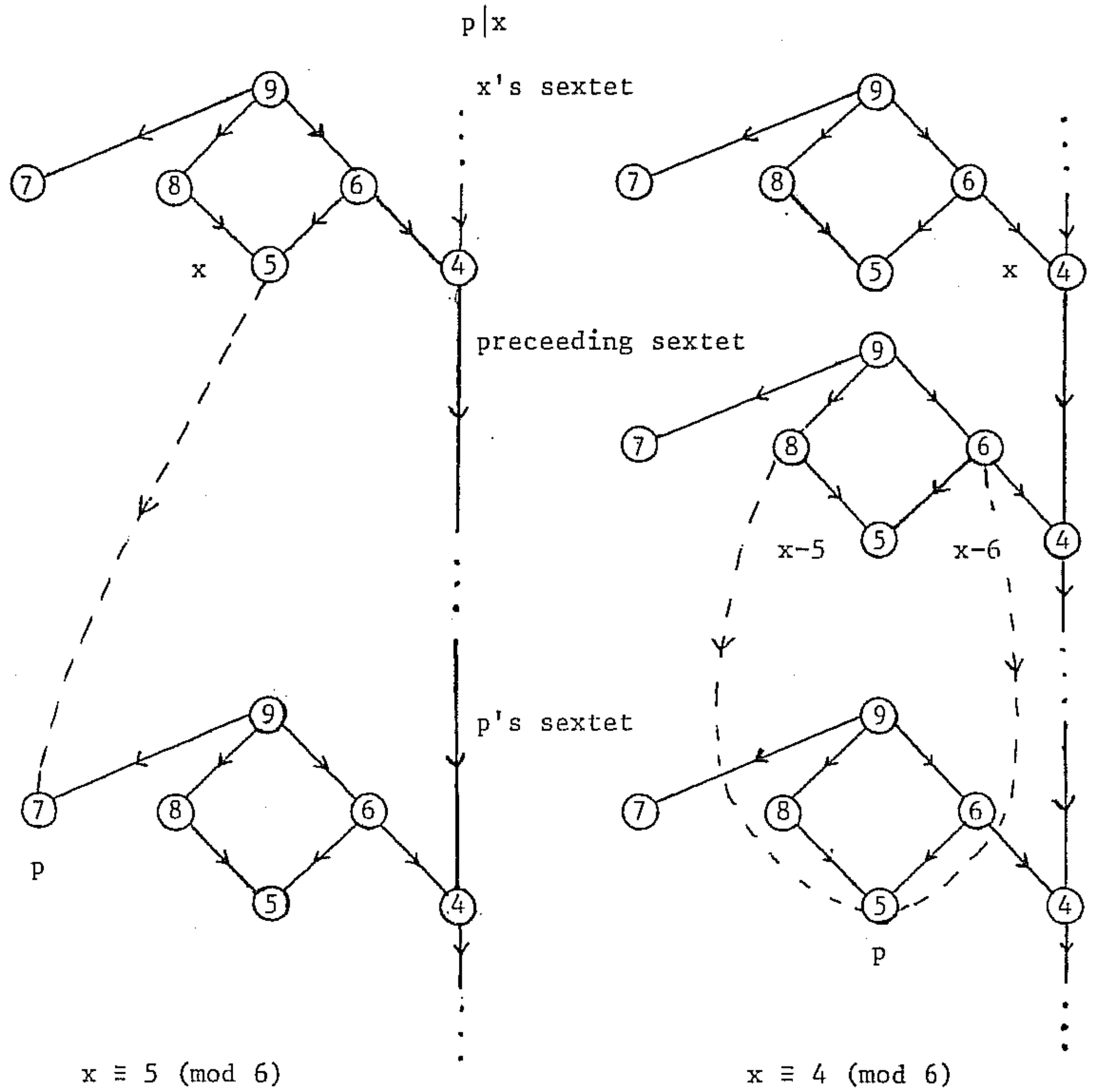
Possible receiving points
for additional edges.

Thus given x,y it will be possible to decide whether $x \perp y$ from the minimal (prime) points lying under the sextets of x and y (and/or the sextets immediately preceeding these). Note that it is trivial to decide whether 2 or 3 divides x from x's position in the sextet.

Suppose $p|x$ where $5 \leq p < x$. Then unless x is a 4-point, add an edge from x to p. If x is a 4-point then add edges from the 6- and 8- points of the preceeding sextet (that is from the numbers $x-2$, $x-4$) to p. No confusion between these two cases can arise, because if $p > 5$ then p can divide at most one of the numbers $x, x-1, x-2, \ldots, x-6$, while if $p = 5$ and $p|x$ where x is a 4-point, then p will divide $x-5$, that is the 5-point of the preceeding sextet, and (see the example below) the resulting partial ordering is the same as if the two additional edges from $x-2, x-4$ to p were left out.

Apart from this case, none of the additional edges is redundant (that is, an omission would change $\textcircled{\leq}$) so the requirement that information be coded into the partial ordering about which element(s) of x's sextet is (are) divisible by p is satisfied. (To check this use the minimality of primes and the fact that if $p > 5$ then p can divide at most one element of a sextet.)

EXAMPLES:



x ≡ 5 (mod 6)

p ≡ 7 (mod 6)

x ≡ 4 (mod 6)

p ≡ 5 (mod 6)

(If p ≡ 5 there is also an edge from x-5 to p.)

Let $H = \langle N, \textcircled{\leq} \rangle$ be the partial ordering obtained from H^{o} by adding edges as described above. Then there is an $L_{\textcircled{\leq}}$ formula which defines $\perp$ in all H_n's. It is left to the reader to verify that the definition of $\leq$ on $[0,n]$ has been preserved.

§5. Applications to spectrum problems.

Consider any language L of first order predicate calculus (with or without identity). A model M will be said to be *normal* if for each pair $a,b \in M$ with $a \neq b$ there is some formula $\psi(u,v_1,\ldots,v_k)$ of L and elements $c_1,\ldots,c_k \in M$ such that

$$M \models \psi(a,c_1,\ldots,c_k) \text{ and } M \models \neg\psi(b,c_1,\ldots,c_k).$$

Obviously if the identity relation = can be defined on M by an L formula then M is normal. (Conversely if M is a finite normal model then = can be defined on M by some L formula, but in general this formula will of course depend on M.)

The *spectrum* of a sentence ϕ of L is the set of positive natural numbers:

$$S_\phi = \{|M| : M \text{ is a finite normal model of } \phi\}.$$

Let

$$\mathcal{S} = \{S_\phi : \phi \text{ is a sentence of some first order language}\}.$$

The *spectrum problem* (Scholz [1952]) is to characterise $\mathcal{S}$. Some characterisations are known. For example, $\mathcal{S}$ = NEXP where NEXP is the class consisting of all sets $S \subseteq N \setminus\{0\}$ with the property that there are a nondeterministic Turing machine T_S and a constant c (dependent on S) such that T_S will accept the binary representation of any element of S with n digits (and no others) in time 2^{cn}. (See Jones and Selman [1974].) However although it has long been known that the class

$$\text{RUD} = \{R \subseteq N\setminus\{0\} : x \in R \text{ is a rudimentary predicate}\}$$

is contained in $\mathcal{S}$, it is not known whether this containment is strict. A proof that $\mathcal{S}$ = RUD would settle, in the negative, the P = NP problem in computational complexity theory. In fact it would follow not only that NP ≠ co-NP, but also that the polynomial time hierarchy of Stockmeyer [1976] would collapse at some level higher than this. (For details see chapter 4. A general discussion of the $P = NP$ problem may be found in Hopcroft and Ullman [1979].)

Analogues of the spectrum problem can be obtained by starting with some theory T in a first order language L, and defining the T-*spectrum* of a sentence ϕ to be:

$$S_\phi^T = \{|M| : M \text{ is a finite normal model of } T \cup \{\phi\}\}.$$

The problem is to characterise

$$\mathcal{S}_T = \{S_\phi^T : \phi \text{ is an } L \text{ sentence}\}.$$

For example, suppose L is the language $L_{=,\gamma}$ with two binary predicate symbols and that T is axiomatized by the axioms of equality (for =) plus the sentence

$$\forall x\, \forall y(x\gamma y \Leftarrow\!\Rightarrow y\gamma x) \wedge \forall x(\neg x\gamma x) \qquad (\dagger)$$

Then $\langle V,=,\gamma\rangle$ is a normal model of T if and only if $\langle V,\gamma\rangle$ is a graph. (Normal here simply implies = is interpreted by identity.) Thus $n \in S_\phi^T$ if and only if some graph with n vertices has the first order "property" ϕ, and the spectrum problem is to characterise the sets S_ϕ^T which can be obtained in this way.

If instead we use the language L_γ and take T to be the theory with $(\dagger)$ as its only axiom, then $\langle V,\gamma\rangle$ is a normal model of T if and only if $\langle V,\gamma\rangle$ is a graph in which no two vertices are joined by edges to exactly the same set of vertices. (We will call $\langle V,\gamma\rangle$ a *normal graph* in these circumstances.)

Similarly other examples yielding nontrivial spectrum problems are obtained by taking L to be L and T to be the theory of partial orderings, or the theory of quasiorderings. (a *quasiordering* is a preordering which can be extended to just one linear ordering, that is, one whose transitive closure is a linear ordering.)

As a final example consider the language $L_{\leqslant,A,M}$ (where

$A(x,y,z) \iff x + y = z$ and $M(x,y,z) \iff x.y = z)$ and let T be the theory FA of *finite arithmetic*:

$$FA = \{\psi : \psi \text{ is an } L_{\leqslant,A,M} \text{ sentence } \wedge \forall n \in N\,(<[0,n], \leqslant, A, M> \models \psi)\}.$$

In this case $S_T = RUD$ by virtue of the second part of the following lemma which states that FA is categorical in every finite cardinality.

LEMMA 5.1 *There is a finite set* Σ *of axioms in the language* $L_{\leqslant,A,M}$ *such that*

(i) *For each* $n \in N$, $<[0,n], \leqslant, A, M> \models \Sigma$.

(ii) *Every finite normal model* $<X, \leqslant^*, A^*, M^*>$ of Σ *with cardinality* n *is isomorphic to* $<[0, n-1], \leqslant, A, M>$.

Proof:

The axioms in Σ describe the basic "algebraic" properties of $\leqslant$,A,M, and assert that A and M satisfy the inductive definitions of addition and multiplication as far as is possible in a finite initial segment of N. (ii) is proved by induction on n.

Now if $R \in RUD$ then (by proposition 0.2) there is a bounded $L_{\leqslant,A,M}$ formula $\psi_R(x)$ such that

$$\forall n(n \in R \iff \psi_R(n-1)),$$

and it follows by the lemma that $R = S_\phi^{FA}$, where ϕ is the sentence

$$\exists x(\forall y(y \leqslant x) \wedge \psi_R(x)).$$

On the other hand, if ϕ is an $L_{\leqslant,A,M}$ sentence and $\phi^*(x)$ is the bounded formula obtained from ϕ by bounding all quantifiers in ϕ by x (where x is a variable not occurring in ϕ) then

$$n \in S_\phi^{FA} \iff \exists x \leqslant n\,(n = x + 1 \wedge \phi^*(x)).$$

Thus $S_{FA} = RUD$.

(Note also that, as mentioned earlier, RUD $\subseteq \mathcal{S}$, since for all ϕ, $S_{\phi}^{FA} = S_{\phi\wedge\sigma} \in \mathcal{S}$ where σ is the conjunction of the axioms in $\sum$.)

An obvious question now is to ask how RUD is related to the other theories T cited above as examples.

THEOREM 5.2 *If* T *is the theory of graphs, the theory of normal graphs, the theory of quasiorderings, or the theory of partial orderings, then* RUD $\subseteq \mathcal{S}_T$.

Proof:

Since RUD = $\mathcal{S}_{FA}$ it will suffice to show that in each case $\mathcal{S}_{FA} \subseteq \mathcal{S}_T$. Therefore suppose ψ is an $L_{\leqslant,A,M}$ sentence and take ψ^o to be $\psi \wedge \sigma$ where σ is the conjunction of the axioms Σ for FA given by lemma 5.1. Fix three formulas $\phi_{\leqslant}(x,y)$, $\phi_A(x,y,z)$, $\phi_M(x,y,z)$ in the language L corresponding to T, and let ψ^* be the L sentence obtained by replacing all occurrences of $\leqslant$,A,M in ψ^o by $\phi_{\leqslant},\phi_A,\phi_m$ in the obvious way. By lemma 5.1, $S_{\psi*}^T \subseteq S_{\psi}^{FA}$, since if $<X,\ldots> \models \psi^*$ then taking $\leqslant^*,A^*,M^*$ to be the predicates defined on X by $\phi_{\leqslant},\phi_A,\phi_M$ yields a model $<X,\leqslant^*,A^*,M^*>$ of FA (with the same cardinality) in which ψ is satisfied.

Thus we will have $S_{\psi*}^T = S_{\psi}^{FA}$ if we can choose $\phi_{\leqslant},\phi_A,\phi_M$ having the property that for each $n \in N \setminus \{0\}$ there is some model $<X,\ldots> \models T$ with $|X| = n$ such that $<X,\leqslant^*,A^*,M^*> \cong <[0,n-1], \leqslant,A,M>$. But for T the theory of graphs (or normal graphs), partial orderings or quasiorderings, the existence of formulas $\phi_{\leqslant},\phi_A,\phi_M$ possessing this property is guaranteed by theorems 4.15, 4.16 and 4.14 respectively. (Note that the preordering $\overset{*}{\leqslant}$ in theorem 4.14 is a quasiordering by lemma 4.13. Also strictly speaking in the case of (normal) graph theory the construction fails because it is impossible to define $\phi_{\leqslant},\phi_A,\phi_M$ which will work for n = 2,3,4,5. However theorem 4.15 does allow the construction of a sentence ψ^* such that $S_{\psi*}^T$ and S_{ψ}^{FA} differ only on [1,5] and it is easy to do a "finite modification" on this sentence so that for $n \leqslant 5$ the normal models of ψ^* with exactly n distinct elements will be all the (normal) graphs with n vertices if $n \in S_{\psi}^{FA}$, and nonexistent otherwise. Since there are normal graphs with 1,2,3,4 and 5 vertices, we will then have $S_{\psi*}^{T} = S_{\psi}^{FA}$.)

COROLLARY 5.3 *Every* $R \in RUD$ *can be represented in the form* $R = S_\phi$ *where* ϕ *is a sentence of first order predicate calculus with only a single binary predicate symbol.*

This corollary shows that the RUD = $\mathcal{S}$ problem would be settled in the negative if it could be shown that there is some $S \in \mathcal{S}$ which cannot be represented in the form $S = S_\phi$ for any sentence ϕ involving only a single binary predicate symbol. Similarly it follows from theorem 5.2 that RUD $\neq \mathcal{S}$ if $\mathcal{S}_T \neq \mathcal{S}$ where T is the theory of graphs, normal graphs, partial orderings or quasiorderings, since in each case $\mathcal{S}_T \subseteq \mathcal{S}$. In fact to prove the existence of a nonrudimentary spectrum it would suffice to show that one of these theories T has the property that for each $n \in N$ there is a sentence ϕ such that $S_\phi^T \neq S_\psi^T$ for all sentences ψ of the form $Q_1Q_2\ldots Q_n X$ where X is quantifier free and each Q_i denotes a block of similar quantifiers of the form $\exists\exists\ldots\exists$ or $\forall\forall\ldots\forall$. This is because (as shown in Harrow [1978], theorem 3) the assumption that $\mathcal{S}$ = RUD implies the existence of an absolute bound on the number of changes of quantifier (but not the total number of quantifiers(†), see Wilkie [1979]) required in giving a bounded $L_{\leqslant,+,\cdot}$ (or $L_{\leqslant,A,M}$) definition $\psi_R(x)$ for any $R \in$ RUD. This bound implies the existence of a similar (but larger) bound on the number of blocks of quantifiers which need appear in the prenex normal forms of the sentences ψ^* constructed in the proof of theorem 5.2.

It should also be mentioned that Fagin [1975] has conjectured

(†) at least if the matrix is an $L_{=,+,\cdot}$ formula.

that for each k there is some element of $\mathcal{S}$ which is not the spectrum of any sentence having only k-ary predicate symbols. He highlighted the nature of the problem by proving:

PROPOSITION 5.4 (Fagin [1975]) *For every* $S \in \mathcal{S}$ *there is some* $k \in N \setminus \{0\}$ *such that* $\{n^k : n \in S\} \in \mathcal{S}_T$ *where* T *is the theory of graphs.*

An easy modification of Fagin's argument shows that the same is true if T is the theory of normal graphs. Furthermore, the theories of quasiorderings and partial orderings have similar properties since:

LEMMA 5.5 *If* T_1, T_2, T_3 *are the theories of graphs, quasiorderings, and partial orderings respectively, then for all* S,

(i) $S \in \mathcal{S}_{T_1} \to \{2n : n \in S\} \in \mathcal{S}_{T_2}$

(ii) $S \in \mathcal{S}_{T_1} \to \{n^2 : n \in S\} \in \mathcal{S}_{T_3}$.

Proof:

(i) It will suffice to show that there are L formulas $\phi_G(x), \phi_\gamma(x,y)$ such that:

(I) For any graph $\langle G, \gamma \rangle$ it is possible to construct a quasiordering $\langle Q, \preccurlyeq \rangle$ with $|Q| = 2|G|$ on which ϕ_G, ϕ_γ define a subset $G^* \subseteq Q$ and a binary relation γ^* on G^* with $\langle G^*, \gamma^* \rangle \cong \langle G, \gamma \rangle$.

(II) There is an L sentence which if satisfied by a quasiordering $\langle Q^*, \preccurlyeq^* \rangle$ will ensure that the structure $\langle G^*, \gamma^* \rangle$ defined by ϕ_G, ϕ_γ is a graph with $2|G^*| = |Q^*|$.

Given $\langle G, \gamma \rangle$ to construct $\langle Q, \preccurlyeq \rangle$ let $Q = G \cup H$ where $G \cap H = \phi$ and $|H| = |G| = n$ (say). Linearly order Q as $h_1 \leqslant g_1 \leqslant h_2 \leqslant g_2 \leqslant \ldots \leqslant h_n \leqslant g_n$ where $h_1, \ldots, h_n$ and $g_1, \ldots, g_n$ are the elements of H and G respectively (in some order). Now define a quasiordering $\preccurlyeq$ on Q by:

$$h_i \preccurlyeq h_j \iff i \leqslant j$$

$$g_i \preccurlyeq g_j \iff (i \leqslant j \wedge g_i \gamma g_j) \vee i = j$$

$$h_i \precsim g_j \iff i = j$$

$$g_i \precsim h_j \iff j = i + 1 .$$

Clearly for all $x,y \in Q$,

$$x \leqslant y \iff \exists z \, \exists w (x \precsim z \wedge z \precsim w \wedge w \precsim y) \qquad (*)$$

so the $\leqslant$ - least and $\leqslant$ - least-but-one elements h_1, g_1 are L definable. But

$$x \in G \iff x = g_1 \vee \neg h_1 \precsim x \qquad (\dagger)$$

and for $x,y \in G$,

$$x \gamma y \iff (x \preceq y \vee y \preceq x) \wedge x = y, \qquad (\dagger\dagger)$$

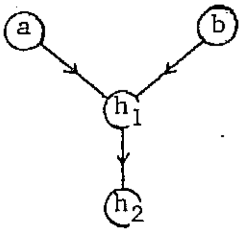
which proves (I).

Also we can define the predicate $y = x'$, that is, y is the immediate $\leqslant$ - successor of x, by an L formula, so using the definitions for G and g_1 given above, there is an L sentence which says:

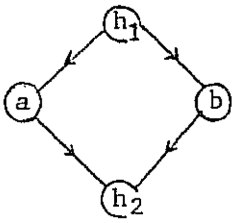
$$g_1 \in G \wedge \forall x (\neg x \in G \iff \exists y \in G (y = x')).$$

Clearly if a quasiordering $\langle Q^*, \precsim^* \rangle$ satisfies the conjunction of this sentence with the sentence saying that the formula on the right of (*) defines a linear ordering, then the formula $\phi_G(x)$ given by ($\dagger$) will define a subset $G^* \subseteq Q^*$ with $2|G^*| = |Q|$, and on G^* the formula on the right of ($\dagger\dagger$) will define a relation γ^* such that $\langle G^*, \gamma^* \rangle$ is a graph.

(ii) The second part of the lemma can be proved in a similar way by considering the following method for coding an arbitrary graph $\langle G, \gamma \rangle$ with $|G| = n$ into a partial ordering $\langle Q, \precsim \rangle$ with $Q = G \cup H$, where $G \cap H = \phi$ and $|H| = n^2 - n$. The partial ordering $\precsim$ is defined so as to associate the elements of H in a two-to-one fashion with the (unordered) pairs of elements of G. Each pair $\{a,b\} \subseteq G$ is associated with two distinct elements $h_1, h_2 \in H$ in one of the following two ways:



Case 1. $a \Upsilon b$



Case 2. $\neg a \Upsilon b$

$\precsim$ is taken to be the (reflexive) partial ordering generated by the directed graph containing just the edges required by these two cases. Since there are $n(n-1)/2$ unordered pairs from G, this is possible if and only if $|H| = n(n-1)$.

The rest of the details are left to the reader.

Now suppose T is one of the theories considered in theorem 5.2. Then

$$\mathcal{S}_T \subseteq \mathrm{RUD} \iff \mathcal{S}_T = \mathrm{RUD}.$$

For if $S \in \mathcal{S}$, then by proposition 5.4 and lemma 5.5 there are non-zero natural numbers j,k such that $\{jn^k: n \in S\} \in \mathcal{S}_T$. Thus if $\mathcal{S}_T \subseteq \mathrm{RUD}$ then $\{jn^k: n \in S\} \in \mathrm{RUD}$. But it is easy to see that if $\{jn^k: n \in S\}$ can be defined by a bounded $L_{\leq,+,\cdot}$ formula, then S can also be defined by a bounded $L_{\leq,+,\cdot}$ formula, so $S \in \mathrm{RUD}$.

Combining this observation with theorem 5.2 yields:

THEOREM 5.6 $\mathcal{S} = \mathrm{RUD} \iff \mathcal{S}_T = \mathrm{RUD}$, *for* T *the theory of graphs, normal graphs, quasiorderings, or partial orderings.*

Thus (using the equivalence of RUD and $\mathcal{S}_{FA}$) the problem of whether $\mathcal{S} \neq \mathrm{RUD}$ can be viewed as a question asking whether graph theory, the theory of partial orderings, etc., are each more complex than the theory of finite arithmetic, in the sense that they generate a broader class of spectra.

Obviously there is much more to be learnt about the classes of spectra generated by the various theories of finite structures, and the relationships between these classes. It might be interesting to know, for example, what can be said about the spectra generated by field theory, group theory, the theory of trees, etc..

CHAPTER 3 Some logical consequences of the linear case of Schinzel's hypothesis H.

Let $L_{=,+,P}$ be the first order language for N (or more generally for some model of Peano arithmetic, PA) with primitive symbols =,+, and the predicate $P(x)$ defined by:

$P(x) \Leftrightarrow x$ *is a prime number.*

Obviously several famous open problems in number theory can be expressed by very simple formulas in this language. For example noting that the predicates $x \leq y$, $x = 0$, $x = 1$, $x = 2$, ..., $y = 2x$, $y = 3x$, ... all have existential (or quantifier free) $L_{=,+,P}$ definitions we have:

$$\forall x \exists p(p \geq x \wedge P(p) \wedge P(p+2))$$

(There exist infinitely many twin primes.)

$$\exists x \forall y \exists p(p \geq y \wedge P(p) \wedge P(p+2x))$$

($\liminf(p_{i+1} - p_i) < \infty$, *where* p_i *denotes the* i *th prime number.)*

$$\forall x(x \geq 2 \rightarrow \exists p \exists q(P(p) \wedge P(q) \wedge 2x = p+q))$$

(Goldbach's conjecture.)

This suggests the following question:

Open problem: *Is there some* $L_{=,+,P}$ *sentence* ϕ *such that* $PA \vdash \phi$?

(Recall that by a well known theorem of Presburger [1930] the $L_{=,+}$ theory of N is identical with the $L_{=,+}$ consequences of PA and therefore decidable.)

In this chapter a partial answer to this question is deduced from the following conjecture (in which $Z[x]$ denotes the ring of all polynomials over the integers Z):

HYPOTHESIS H (Schinzel) *Let* $f_1(x), f_2(x), \ldots, f_n(x)$ *be irreducible polynomials in* $Z[x]$ *with positive leading coefficients and suppose that*

$$\forall y \neq 1 \; \exists x(\; y \nmid \prod_{1 \leq i \leq n} f_i(x) \;),$$

then there exist infinitely many numbers x *for which* $f_1(x), f_2(x), \ldots, f_n(x)$ *are all primes.*

Actually we will require only the linear case where $f_1(x), f_2(x), \ldots, f_n(x)$ are all polynomials of degree 1:

CONJECTURE H_L (Dickson [1904] *If* $a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_n \in Z$ *with all* $a_i > 0$ *and*

$$\forall y \neq 1 \; \exists x(\; y \nmid \prod_{1 \leq i \leq n} (a_i x + b_i) \;)$$

then there exist infinitely many numbers x *such that* $a_i x + b_i$ *is prime for all* i.

REMARK: Note that in each case the conjecture states that an obviously necessary condition for the existence of primes is also sufficient. Also the words "infinitely many" are actually redundant - see Schinzel and Sierpiński [1958]. Special cases of H_L include Dirichlet's theorem on the existence of infinitely many primes $p \equiv a \pmod b$ provided a and b are coprime, the twin primes conjecture, and the conjecture that there exist arbitrarily long arithmetic progressions consisting entirely of prime numbers.

Assuming H_L has the following consequence for N:

THEOREM 1 *If* H_L *is true then the predicate* $z = x.y$ *can be defined by an* $L_{=,+,P}$ *formula.*

However the definition cannot be an existential formula since:

THEOREM 2 *If* H_L *is true then the existential* $L_{=,+,P}$ *theory of* N *is decidable.*

In fact it will be easy (and left to the reader) to check that the proof of theorem 1 shows that there is an $L_{=,+,P}$ formula which defines $z = x.y$ on any model of the $L_{=,+,P}$ consequences of $PA + H_L$ (or indeed of $I\Delta_0 + \forall x\,(2^x$ *exists*$) + H_L$) where H_L is a single sentence expressing the conjecture H_L. (Note that here Z will be interpreted as "integers in the sense of the model".) Therefore turning any undecidable $L_{=,+,\cdot}$ sentence into an $L_{=,+,P}$ sentence in the obvious way gives:

COROLLARY 3 *If* $PA + H_L$ *is consistent then there is some* $L_{=,+,P}$ *sentence* ϕ *such that* $PA \nvdash \phi$.

Similarly the proof of theorem 2 will show that the existential $L_{=,+,P}$ theory of N is the set of existential sentences ϕ such that

$$PA + H_L^1 + H_L^2 + H_L^3 + \ldots \vdash \phi$$

(assuming this axiom system is sound, where H_L^i, $i \in N$, is an enumeration of the instances of H_L with $n, a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_n$ standard integers. (Actually PA can be replaced here by an axiom system considerably weaker than $I\Delta_0$.)

The proofs of theorems 1 and 2 depend on the following lemma (in which $\equiv$ denotes the identity relation between polynomials).

PROPOSITION 4 (Schinzel [1961]) *Suppose* $a_1x + b_1, a_2x + b_2, \ldots, a_nx + b_n$ *satisfy the conditions* of H_L *and let* $c_1x + d_1, c_2x + d_2, \ldots, c_mx + d_m$ *be in* $Z[x]$ *with* $c_j > 0$ *and* $c_jx + d_j \not\equiv a_ix + b_i$ *for all* i,j. *Then* H_L *implies the existence of infinitely many numbers* x *such that:*

(i) $a_ix + b_i$ *is a prime for each* $i \in [1,n]$.

(ii) $c_jx + d_j$ *is composite for each* $j \in [1,m]$.

Using this we now prove:

LEMMA 5 H_L *implies that for all* $a > 1$, $n \in N$, *there exist* $n + 1$ consecutive *prime numbers* $p_0, p_1, \ldots, p_n$ *with* $p_i - p_{i-1} = 2ia$ *for all* $i \in [1,n]$.

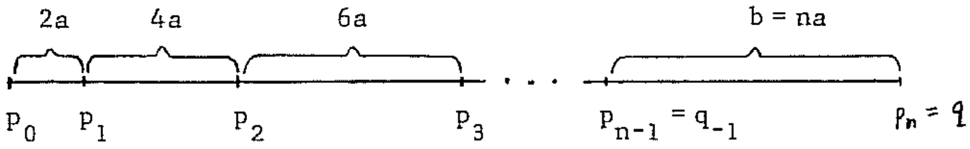
Proof:

Take the polynomials $x, x+2a, \ldots, x+i(i+1)a, \ldots, x+n(n+1)a$ as the $a_i x + b_i$'s and the polynomials $x+m$ with $0 < m < n(n+1)a$, $m \neq i(i+1)a$ for all i, as the $c_j x + d_j$'s. We need only check that for every prime p there is some x such that $\forall i \leq n \; (p \nmid (x + i(i+1)a))$, in other words that there is some residue class $x \pmod{p}$ such that $x \not\equiv -i(i+1)a \pmod{p}$ for all i. But clearly the residue class of $-i(i+1)a \pmod{p}$ is determined by that of $i \pmod{p}$, so since $i \equiv 0 \pmod{p}$ and $i \equiv -1 \pmod{p}$ both produce $-i(i+1)a \equiv 0 \pmod{p}$ we see that $-i(i+1)a$ takes fewer than p values (mod p). Therefore such an x exists.

THEOREM 1 H_L *implies that* $z = x.y$ *can be defined by an* $L_{=,+,P}$ *formula.*

Proof:

Since $z = x.y$ can be defined by an $L_{+,|}$ formula (see Robinson [1949] or chapter 2 of this thesis) it suffices to show that the predicate $a|b$ is $L_{=,+,P}$ definable. Suppose $a|b$ with $n = \frac{b}{a}$ (and $b \neq 0$). Then assuming H_L, it follows by lemma 5 that a pattern of the following sort exists:



where $p_0, p_1, \ldots, p_n$ are consecutive prime numbers with $p_i - p_{i-1} = 2ia$. Conversely if such a pattern exists then obviously $a|b$. But we can construct an $L_{=,+,P}$ formula $\phi(a,b)$ asserting that there exist primes $p_0 < p_1 < q_{-1} < q$ such that

(i) p_0, p_1 *are consecutive primes* $\wedge \; p_1 = p_0 + 2a$,

(ii) q_{-1}, q *are consecutive primes* $\wedge \; q = q_{-1} + 2b$,

(iii) $\forall p \; \forall r \; \forall s \; (p_0 < p < r < s < q \; \wedge \; p, r, s$ *are consecutive primes* $\rightarrow s - r = r - p + 2a)$.

Clearly, $a|b \iff \phi(a,b)$.

THEOREM 2 H_L *implies that the existential* $L_{=,+,P}$ *theory of* N *is decidable.*

Proof:

Suppose ϕ is an existential $L_{=,+,P}$ sentence. Using the equivalences:

$$(\psi \rightarrow \chi) \iff \chi \vee \neg\psi, \qquad \neg(\psi \vee \chi) \iff \neg\psi \wedge \neg\chi,$$

$$\neg(\psi \wedge \chi) \iff \neg\psi \vee \neg\chi, \qquad \chi\wedge(\psi \vee \rho) \iff (\chi\wedge\psi) \vee (\chi \wedge \rho),$$

we can effectively find an equivalent sentence in *disjunctive normal form*:

$$\exists\vec{x}(\psi_1(\vec{x}) \vee \psi_2(\vec{x})\vee\ldots\vee\psi_m(\vec{x})),$$

where each $\psi_j(\vec{x})$ is a conjunction of atomic and negated atomic formulas of the forms:

$$P(\, a_0 + a_1x_1 + a_2x_1 + \ldots + a_nx \,)$$

$$\neg P(\, b_0 + b_1x_1 + b_2x_2 + \ldots + b_nx_n \,)$$

$$c_0 + c_1x_1 + c_2x_2 + \ldots + c_nx_n = 0$$

$$\neg d_0 + d_1x_1 + d_2x_2 + \ldots + d_nx_n = 0$$

where we have taken the liberty of replacing terms by linear polynomials with coefficients in Z, so $a_i, b_i \in$ N and $c_i, d_i \in$ Z.

But since

$$\exists\vec{x}(\psi_1(\vec{x})\vee\ldots\vee\psi_m(\vec{x})) \iff \exists\vec{x}\psi_1(\vec{x})\vee\ldots\vee\exists\vec{x}\psi_m(\vec{x}),$$

the individual disjuncts $\exists\vec{x}\psi_j(\vec{x})$ can be considered separately, so it suffices to show that there is an algorithm for deciding whether or not a system of relations of the form:

$$\begin{aligned}
&P(\, a_0 + \sum_{1\leq i\leq n} a_i x_i) \;, && \vec{a} \in A,\\
&\neg P(\, b_0 + \sum_{1\leq i\leq n} b_i x_i \,), && \vec{b} \in B,\\
&c_0 + \sum_{1\leq i\leq n} c_i x_i = 0 \;, && \vec{c} \in C,\\
&\neg\, d_0 + \sum_{1\leq i\leq n} d_i x_i = 0, && \vec{d} \in D,
\end{aligned}$$

with $A, B \subset N^{n+1}$, $C,D \subset Z^{n+1}$ finite sets, has a simultaneous solution $\vec{x} \in N^n$.

At the expense of some more "separating out" of disjuncts we may replace the inequalities

$$d_0 + \sum_{1\leq i\leq n} d_i x_i = 0$$

using the equivalence:

$$d_0 + \sum_{1\leq i\leq n} d_i x_i = 0 \iff$$

$$(d_0 + \sum_{1\leq i\leq n} d_i x_i > 0) \vee ((-d_0) + \sum_{1\leq i\leq n} (-d_i) x_i > 0),$$

and thus obtain systems of the form:

$$\begin{aligned}
&P(\, a_0 + \sum_{1\leq i\leq n} a_i x_i \,), && \vec{a} \in A,\\
&\neg P(\, b_0 + \sum_{1\leq i\leq n} b_i x_i \,), && \vec{b} \in B,\\
&c_0 + \sum_{1\leq i\leq n} c_i x_i = 0 \;, && \vec{c} \in C,\\
&d_0 + \sum_{1\leq i\leq n} d_i x_i > 0 \;, && \vec{d} \in D.
\end{aligned} \tag{1}$$

We will actually consider systems of this sort where A,B (as well as C,D) are allowed to be finite subsets of Z^{n+1}, but with the following added requirement:

RESTRICTION R: *For every* $\vec{a} \in A$, $\vec{b} \in B$, *the system* (1) *must contain:*

$$\text{(i)}\quad (a_0 - k_{\vec{a}}) + \sum_{1 \le i \le n} a_i x_i > 0 \quad \textit{for some} \quad k_{\vec{a}} \in N,$$

$$\text{(ii)}\quad (b_0 - m_{\vec{b}}) + \sum_{1 \le i \le n} b_i x_i > 0 \quad \textit{for some} \quad m_{\vec{b}} \in N.$$

It will be left to the reader to verify that this requirement can be preserved through all the reductions made below.

We will also assume that if any $\vec{a}$, $\vec{b}$, $\vec{c}$, or $\vec{d}$ is zero apart (possibly) from the first term a_0, b_0, c_0 or d_0 (respectively) then the truth or falsity of the relevant formula is determined immediately. If it is false then the system has no solution. If it is true then the formula is deleted.

Now consider two cases:

Case 1 $C \neq \emptyset$.

Multiplying equations by -1 and renumbering variables (if necessary) we may suppose that the system contains an equation

$$c_0^* + \sum_{1 \le i \le n} c_i^* x_i = 0$$

with $c_n^* > 0$, for some $\vec{c}^* \in C$. Thus if $\vec{x}$ is a solution of (1), then

$$c_0^* + \sum_{1 \le i \le n-1} c_i^* x_i \equiv 0 \pmod{c_n^*}. \tag{2}$$

But we can test effectively whether this congruence has a solution, and if so, find a finite set E of solutions $\vec{X} = \langle X_1, X_2, \ldots, X_{n-1} \rangle$ such that E includes each solution $\vec{X} (\text{mod } c_n^*)$ with $0 \le X_i < c_n^*$ for all i. Now if $\vec{x}$ is a solution of (1) then there is some $\vec{t} \in Z^{n-1}$, $\vec{X} \in E$, such that:

$$x_i = X_i + c_n^* t_i \quad \text{for} \quad i \in [1, n-1],$$

$$x_n = \frac{-1}{c_n^*} (c_0^* + \sum_{1 \le i \le n-1} c_i^* X_i) - \sum_{1 \le i \le n-1} c_i^* t_i . \qquad (3)_{\vec{X}}$$

Thus system (1) has a solution if and only if for some $\vec{X} \varepsilon E$ there exists a solution of the system (called $(1)_{\vec{X}}$) obtained from (1) by substituting the expressions on the right of $(3)_{\vec{X}}$ for the variables $x_1,\ldots,x_n$ in (1), and adjoining the inequality:

$$1 - \frac{1}{c_n^*}(c_0^* + \sum_{1\leq i\leq n-1} c_i^* X_i) - \sum_{1\leq i\leq n-1} c_i^* t_i > 0.$$

Clearly the systems $(1)_{\vec{X}}$ in the variables $t_1, t_2, \ldots, t_{n-1}$ may now be considered individually, and each is of the form (1) but with one less variable.

Continuing in this way we either decide that the system has or does not have a solution, or eventually arrive at:

Case II $C = \phi$.

Here the system is of the form:

$$\begin{array}{lll} P(\ a_0 + \sum_{1\leq i\leq n} a_i x_i\ )\ , & \vec{a}\ \varepsilon\ A, & \\ \neg P(\ b_0 + \sum_{1\leq i\leq n} b_i x_i\ )\ , & \vec{b}\ \varepsilon\ B, & (4) \\ d_0 + \sum_{1\leq i\leq n} d_i x_i > 0\ , & \vec{d}\ \varepsilon\ D. & \end{array}$$

We may assume that there is no $a_0 = 0$. For since $\vec{0} \notin A$, the system of linear inequalities

$$a_0 + \sum_{1\leq i\leq n} a_i x_i \neq 0, \qquad \vec{a}\ \varepsilon\ A,$$

certainly has an (effectively found) solution $\vec{X} \varepsilon N^n$ (since some $\vec{X} \varepsilon N^n$ lies outside the union of the n-1 dimensional "subspaces" comprised of the rational solutions of the equations $a_0 + \sum_{1\leq i\leq n} a_i x_i = 0$), and clearly all $\vec{x} \varepsilon N$ satisfy

$$(\bigwedge_i x_i = X_i + t_i) \vee \bigvee_i (x_i = 0 \vee x_i = 1 \vee \ldots \vee x_i = X_i - 1).$$

Thus the problem reduces to considering:

(i) the systems obtained from (4) by replacing some x_i with a number $< X_i$,

(ii) the system with variables $\vec{t}$ obtained from (4) by putting $x_i = X_i + t_i$ for $i \in [1,x]$.

But the systems in (i) have fewer than n variables, while the system in (ii) satisfies the requirement that all constant terms be nonzero, since

$$a_0 + \sum_{1\leq i\leq n} a_i x_i = (a_0 + \sum_{1\leq i\leq n} a_i X_i) + \sum_{1\leq i\leq n} a_i t_i$$

where $a_0 + \sum_{1\leq i\leq n} a_i X_i \neq 0$.

We can also assume that every $d_0 \leq 0$, since if $d_0 > 0$ then

$d_0 + \sum_{1\leq i\leq n} d_i x_i > 0 \iff$

$(\sum_i d_i x_i > 0) \vee (\sum_i d_i x_i = 0) \vee (1 + \sum_i d_i x_i = 0) \vee \ldots \vee ((d_0 - 1) + \sum_i d_i x_i = 0)$.

so we may reduce the problem to considering the system with $d_0 + \sum_{1\leq i\leq x} d_i x_i > 0$ replaced by $\sum_{1\leq i\leq n} d_i x_i > 0$, and d_0 systems of the form to which case I applies. (Note that each time we return to Case I the number of variables is <u>reduced</u> by the procedure used there, so the "complexity" of the systems which need be considered is decreased.)

We now have only systems of the form:

$$\begin{aligned} &P(\, a_0 + \sum_{1\leq i\leq n} a_i x_i\,)\ , && \vec{a} \in A,\ (a_0 \neq 0), \\ &\neg P(\, b_0 + \sum_{1\leq i\leq n} b_i x_i\,)\ , && \vec{b} \in B, \\ &d_0 + \sum_{1\leq i\leq n} d_i x_i > 0\ , && \vec{d} \in D,\ (d_0 \leq 0), \end{aligned} \tag{5}$$

left to consider, where by the restriction R introduced earlier there are inequalities present which imply $a_0 + \sum_{1\leq i\leq n} a_ix_i > 0$ and $b_0 + \sum_{1\leq i\leq n} b_ix_i \geqslant 0$ for all $\vec{a} \in A$, $\vec{b} \in B$. Now let $S = \{p\colon P(p) \wedge \exists\vec{a} \in A\ (\ p|a_0\ )\}$. Consider the system of relations:

$$\begin{aligned}
&a_0 + \sum_{1\leq i\leq n} a_ix_i \not\equiv 0 \pmod{p}, && p \in S,\ \vec{a} \in A,\\
&a_0 + \sum_{1\leq i\leq n} a_ix_i \neq b_0 + \sum_{1\leq i\leq n} b_ix_i, && \vec{a} \in A,\ \vec{b} \in B, \qquad (6)\\
&d_0 + \sum_{1\leq i\leq n} d_ix_i > 0, && d \in D.
\end{aligned}$$

Since $$x \equiv 0 \pmod{p} \iff \exists z(x = \underbrace{z + z + \ldots + z}_{p\ \text{times}}),$$

there is clearly an $L_{=,+}$ sentence which expresses the statement that system (6) has a solution. But using Presburger's algorithm we can decide effectively whether a given $L_{=,+}$ sentence is true and thus whether (6) has a solution. If it does not have a solution, then the only way (5) can have a solution is if some $a_0 + \sum_{1\leq i\leq n} a_ix_i = p$ for some prime $p \in S$. Therefore the problem can be reduced to considering the $|S|.|A|$ systems obtained by adding one of the possible equations of the form

$$(a_0 - p) + \sum_{1\leq i\leq n} a_ix_i = 0$$

to (5), and of course case I applies to all of these.

If (6) does have a solution, then some such solution $\vec{X}$ can obviously be found effectively. It is claimed that in this case (5) will also have a solution (provided H_L is true). To prove this we will apply proposition 4 to the polynomials (in x):

$$\begin{aligned}
&a_0 + (\sum_{1\leq i\leq n} a_iX_i)x, && \vec{a} \in A,\\
& && \qquad\qquad (7)\\
&b_0 + (\sum_{1\leq i\leq n} b_iX_i)x, && b \in B.
\end{aligned}$$

The restriction R introduced earlier and the requirement that $d_0 \leq 0$ ensure that $\sum_{1\leq i\leq n} a_iX_i > 0$ and $\sum_{1\leq i\leq n} b_iX_i > 0$. To check that the product of the polynomials in the first line of (7) has no fixed divisor (other than 1), suppose that some prime $p | \prod_{\vec{a}\in A} (a_0 + (\sum_{1\leq i\leq n} a_iX_i)x)$ for every $x \in N$. Taking $x = 0$ we see that $p \in S$, while taking $x = 1$ shows that $p | \prod_{\vec{a}\in A} (a_0 + \sum_{1\leq i\leq n} a_iX_i)$, that is, that $a_0 + \sum_{1\leq i\leq n} a_iX_i \equiv 0 \pmod p$ for <u>some</u> $\vec{a} \in A$, which is impossible since X is a solution of (6).

Thus assuming H_L it follows from proposition 4 that $x > 0$ can be chosen satisfying:

$$P(\ a_0 + \sum_{1\leq i\leq n} a_i(X_ix))\ , \qquad \vec{a} \in A,$$

$$\neg P(\ b_0 + \sum_{1\leq i\leq n} b_i(X_ix))\ , \qquad \vec{b} \in B.$$

Taking $x_i = X_ix$ gives a solution of (5) since

$$\begin{aligned} d_0 + \sum_{1\leq i\leq n} d_ix_i &= d_0 + (\sum_{1\leq i\leq n} d_iX_i)x \\ &\geq d_0 + \sum_{1\leq i\leq n} d_iX_i > 0, \end{aligned}$$

(because $\vec{X}$ is a solution of (6) and thus $\sum_{1\leq i\leq n} d_iX_i > - d_0 \geq 0$).

Having proved theorem 2 it now follows immediately (by proposition 0.1 of chapter 2) that:

<u>COROLLARY 6</u> *If* H_L *is true then* $z = x.y$ *cannot be defined by an existential* $L_{=,+,P}$ *formula.*

<u>REMARK</u>: The argument above is easily extended to give a decision procedure for all $L_{\leq,+,P}$ sentences of the form $\tilde{\exists}\vec{x}\,\exists\vec{y}\,\phi(\vec{x},\vec{y})$ where $\phi(\vec{x},\vec{y})$ is quantifier free and $\tilde{\exists}x_i$ denotes $\forall z_i \exists x_i \geq z_i$ (where z_i does not occur in $\phi(\vec{x},\vec{y})$), that is, $\tilde{\exists}$ is the quantifier:

there exist infinitely many x_i.

Corollary 6 can also be proved for $L_{\leqslant,+,P}$ formulas with only $\exists$ and $\tilde{\exists}$ quantifiers, by analysing the properties of predicates defined by such formulas.

The theorems proved above suggest asking what the situation regarding decidability is for the languages $L_{=,\prime,P}$, $L_{\leqslant,P}$ (in the notation of chapter 2). In the first case we have

THEOREM 7 H_L *implies that the* $L_{=,\prime,P}$ *theory of* N *is decidable.*

This can be proved directly by quantifier elimination, but in view of theorem 2 it also follows from the general result:

PROPOSITION 8 (Hanf [1965], Thomas [1978]) *Let* $P(x)$ *be any unary predicate on* N. *Then the* $L_{\prime,P}$ *theory of* N *is decidable if and only if the existential* $L_{\prime,P}$ *theory of* N *is decidable.*

Open Question: *Is the* $L_{\leqslant,P}$ *theory of* N *undecidable?*

As observed by Thomas [1978], this question has a negative answer if H_L fails so badly that $\lim\inf(p_{i+1} - p_i) = \infty$ (where p_i denotes the i th prime).

It also seems worthwhile to remark that although it is obviously hopeless to seek unconditional proofs of the decidability results above using only currently known number theory, it does not seem quite so implausible that a proof of the undecidability of the $L_{=,+,P}$ theory of N might be obtained without going very far beyond what is currently known about primes. Certainly there are other ways of deducing the undecidability result from H_L. For example, it is a consequence of the conjecture:

If r *is divisible by all prime numbers* $\leqslant$ n *then there exists a sequence of* n *consecutive prime numbers in arithmetic progression with common difference* r. (This states that arithmetic progressions of consecutive primes of all conceivable sorts exist, and is a consequence of H_L - see Schinzel and Sierpiński [1958].)

Finally there is the question of what natural axiom systems it is reasonable to expect H_L (or H for that matter) to be provable in (assuming it is true). Let $I\Sigma_1$ be the axiom system similar to the system $I\Delta_0$ defined in chapter 1, but with induction allowed for induction hypotheses expressed by Σ_1 formulas, that is, formulas of the form $\exists\vec{x}\,\theta(\vec{x},y)$, where $\theta(\vec{x},y)$ is a Δ_0 formula. (Intuitively $I\Sigma_1$ corresponds to doing inductive arguments with recursively enumerable predicates as the induction hypotheses - a large proportion of known number theory can be developed this way.)

CONJECTURE: $I\Sigma_1 \not\vdash H_L$

(where H_L denotes the single sentence expressing the conjecture, as described above).

APPENDIX I An alternative proof that addition and multiplication can be defined by bounded $L_{\leqslant,\perp}$ formulas.

Since the proof of theorem 1.3 shows that $z = x.y$ can be defined by a bounded $L_{\leqslant,A,\perp}$ formula where $A(x,y,z) \iff x+y = z$, it suffices to prove that $z = x+y$ can be defined by a bounded $L_{\leqslant,\perp}$ formula.

Recall that for all real numbers a,b the integer part sign $[\ ]$ has the property:

$$[a] + [b] \leqslant [a+b] \leqslant [a] + [b] + 1.$$

LEMMA $z = x+y$ *if and only if* z *is the least number such that*:

(I) $x \leqslant z \wedge y \leqslant z$

(II) $z \equiv x+y \pmod 6$

(III) *For every prime* $p \leqslant z$, $p \neq 7$, *either*

$$[\tfrac{z}{p}] \equiv [\tfrac{x}{p}] + [\tfrac{y}{p}] \pmod 7$$

or $$[\tfrac{z}{p}] \equiv [\tfrac{x}{p}] + [\tfrac{y}{p}] + 1 \pmod 7).$$

Proof:

Obviously $z = x+y$ satisfies (I), (II), (III), so it is only necessary to show that if $z < x+y$ at least one of these must fail. Suppose $z < x+y$ satisfies (I) and (II). Then by (II), $x+y-z = 6m$ for some $m \geqslant 1$. But for any $n \geqslant 1$ there is a prime p with $n < p \leqslant 2n$ by Bertrand's postulate (a theorem of Chebychev - see, for example, Hardy and Wright [1979] or chapter 1 of this thesis). Hence for any $m \geqslant 1$ there is a prime $p \neq 7$ with $\frac{3}{2}m \leqslant p \leqslant 3m$, and thus some prime $p \neq 7$ satisfies $2p \leqslant x+y-z \leqslant 4p$.

But then $2 \leqslant [\frac{x}{p} + \frac{y}{p} - \frac{z}{p}] \leqslant 4$

so $2 + [\frac{z}{p}] \leqslant [\frac{x}{p} + \frac{y}{p}] \leqslant 5 + [\frac{z}{p}]$

and $1 + [\frac{z}{p}] \leqslant [\frac{x}{p}] + [\frac{y}{p}] \leqslant 5 + [\frac{z}{p}]$.

In other words,

$$[\tfrac{z}{p}] - [\tfrac{x}{p}] - [\tfrac{y}{p}] = -1,-2,-3,-4, \text{ or } -5.$$

Hence

$$[\tfrac{z}{p}] - [\tfrac{x}{p}] - [\tfrac{y}{p}] \not\equiv 0,1 \pmod 7,$$

and since $p \leqslant x+y-z \leqslant z+z-z = z$ (by (I)), condition (III) fails.

<u>THEOREM</u> $z = x+y$ *can be defined by a bounded* $L_{\leqslant,\perp}$ *formula.*

<u>Proof:</u>

We show that conditions (I), (II), (III) in the lemma can be defined by bounded $L_{\leqslant,\perp}$ formulas.

As the predicates $y = x'$, $x = 0$ are definable by bounded $L_{\leqslant}$ formulas, $x \equiv i \pmod 6$ can be defined for $i = 0,1,\ldots,5$ using:

$x \equiv i \pmod 6 \iff x = i \vee (x \geqslant 6 \wedge \neg(2 \perp x-i) \wedge \neg(3 \perp x-i))$.

Thus $x+y \equiv z \pmod 6$ can be defined by a bounded $L_{\leqslant,\perp}$ formula, since

$$x+y \equiv z (\text{mod } 6) \iff \bigvee_{\substack{i,j,k\leqslant 5\\ i+j\equiv k(\text{mod } 6)}} (x \equiv i(\text{mod } 6) \wedge y \equiv j(\text{mod } 6) \wedge z \equiv k(\text{mod } 6)).$$

Similarly condition (III) will be expressible by a bounded $L_{\leqslant,\perp}$ formula if it can be shown that for $i = 0,1,\ldots,6$,

$$p \textit{ is prime} \wedge p \neq 7 \wedge [\tfrac{x}{p}] \equiv i \pmod 7$$

can be defined by a bounded $L_{\leqslant,\perp}$ formula. But this is the case since the predicate p *is prime* can obviously be handled (as in §1), and $[\frac{x}{p}] \equiv i \pmod 7$ for a prime $p \neq 7$ if and only if there exists $u \leqslant x$ with $\neg p \perp u \wedge \neg 7 \perp u$ such that there are <u>exactly</u> i distinct numbers v satisfying $u < v \leqslant x \wedge \neg p \perp v$. (To see this consider the numbers:

$u = 7pw < p(7w+1) < p(7w+2) < \ldots < p(7w+i) \leqslant x < p(7w+i+1)$.)

REMARK: This method can be refined to show that the full stength of $\perp$ (with $\leqslant$) is not required to define the addition and multiplication predicates on $[0,n]$. Write $p|_S\ x$ for $p \in S \wedge p|x$, where S is a set of primes. Then there are $L_{\leqslant,|_S}$ formulas $\phi_A(x,y,z)$, $\phi_m(x,y,z)$ such that for all n and all $x,y,z \in [o,n]$,

$$x+y = z \Leftrightarrow \langle [0,n], \leqslant, |_S \rangle \models \phi_A(x,y,z)$$

$$x\,.\,y = z \Leftrightarrow \langle [0,n], \leqslant, |_S \rangle = \phi_M(x,y,z)$$

provided $\{p: p \leqslant C \log n\} \subseteq S$ (where C is a suitably large constant, for example $C > 1$).

This is because for $y \leqslant 2C \log n$ the definition of $x+y = z$ given above requires only a knowledge of which numbers are divisible by primes $p \leqslant C \log n$. Thus there is an $L_{\leqslant,|_S}$ formula $\theta(z,y,p)$ such that for all primes $p \leqslant C \log n$ and all z,y,

$$\theta(z,y,p) \Leftrightarrow y < p \wedge z \equiv y \pmod{p}.$$

Also since the predicates $x+y = z$ and thus $x\,.\,y = z$ can be defined on $[0,2C \log n]$, it is possible to define $x+y \equiv z \pmod{p}$ and $x\,.\,y \equiv z \pmod{p}$ (the latter in two steps) for all $x,y \in [0,n], p \leqslant C \log n$, and hence obtain definitions for $x+y = z$ and $x\,.\,y = z$ on $[0,n]$ from the equivalences:

$$x+y = z \Leftrightarrow \forall p \leqslant C \log n\ (x+y \equiv z \pmod{p})$$

$$x\,.\,y = z \Leftrightarrow \forall p \leqslant C \log n\ (x\,.\,y \equiv z \pmod{p}).$$

Using these ideas, the method can also be extended to show that there are bounded $L_{\leqslant,\perp}$ formulas $\psi_A(x,y,z)$, $\psi_M(x,y,z)$ with the property that if a suitable axiom ϕ defining $\perp$, for example:

$$\forall x\, \forall y\, (x \perp y \Leftrightarrow \forall z(\exists u(x = z.u) \wedge \exists v(y = z.v) \rightarrow z = 1)),$$

is added to the axiom system $I\Delta_0$ considered in chapter 1, then

$$I\Delta_0 + \phi \vdash \forall x\, \forall y\, \forall z(x+y = z \Leftrightarrow \psi_A(x,y,z)),$$

$$I\Delta_0 + \phi \vdash \forall x\, \forall y\, \forall z(x\,.\,y = z \Leftrightarrow \psi_M(x,y,z)).$$

APPENDIX II Some problems in computational complexity

As mentioned in chapter 2 a positive solution to the $RUD = \mathcal{S}$ problem would solve several open problems in computational complexity theory. For a start the class of all sets S of positive integers such that S can be accepted or rejected by some deterministic Turing machine using linear working space (or equivalently such that S has an $\mathcal{E}^2_*$ definition) is easily seen to include RUD and to be contained in $\mathcal{S}$. Thus we have:

$$RUD(= \Delta_0 \ \textit{sets}) \subseteq \textit{linear space}(= \mathcal{E}^2_* \ \textit{sets}) \subseteq \mathcal{S}(= \mathrm{NEXP}),$$

where it is not known whether *linear space* = NEXP or whether $RUD =$ *linear space*.

A proof that $RUD = \mathcal{S}$ would also settle the $P = NP$ problem which asks whether every $S \subseteq N$ which can be accepted in polynomial time by some nondeterministic Turing machine (that is, every $S \in NP$) is a member of the class P of all sets which can be accepted (or rejected) in polynomial time by a deterministic Turing machine. Let $co - NP = \{N \setminus S : S \in NP\}$. At present it is not known whether $NP = co - NP$, although $P = NP \Rightarrow NP = co - NP$ since if $S \in P$ then obviously $N \setminus S \in P$. These problems were generalised by Stockmeyer [1976]. The *Stockmeyer polynomial time hierarchy* is defined by taking

$\Sigma^P_0 = \Pi^P_0 = P,$

$S \in \Sigma^P_{n+1} \iff$

S *is accepted in polynomial time by some nondeterministic Turing machine using an oracle* $A \in \Pi^P_n$,

$\Pi^P_{n+1} = \{N \ \ S : S \in \Sigma^P_{n+1}\}.$

The problem of whether the classes $\Sigma_0 \subseteq \Sigma^P_1 \subseteq \Sigma_2 \subseteq \ldots$ form a strict hierarchy is open. Let $\Delta^P = \bigcup_n \Sigma^P_n = \bigcup_n \Pi^P_n$. Since $\Sigma^P_1 = NP$ it is obvious that $NP = \Delta^P \iff NP = co - NP$. Also as observed by Paris and

Dimitracopoulos [????], the sets $S \varepsilon \Delta^P$ are precisely those subsets of N with the property that $S = \{x: \phi(x)\}$ for some $\phi(x)$ of the form:

$$Q_1 y_1 \leqslant x^{[\log_2 x]^m} Q_2 y_2 \leqslant x^{[\log_2 x]^m} \ldots Q_s y_s \leqslant x^{[\log_2 x]^m} \lambda(x,\vec{y}) \qquad (1)$$

where $Q_1, Q_2, \ldots, Q_s$ *denote quantifiers* ($\forall$ *or* $\exists$), $m \varepsilon N$, *and* λ *is a quantifier free* $L_{\leqslant,+,\cdot}$ *formula.*

Furthermore, the polynomial time hierarchy collapses at some point (that is, $\Delta^P = \Sigma^P_n = \Pi^P_n$ for some n) if and only if a fixed bound independent of S can be placed of the number of quantifiers required in (1).

An analogous open problem for Δ_0 formulas is whether there is some fixed k such that every R ε R U D can be defined by a formula of the form:

$$Q_1\vec{v}^{(1)} \leqslant x\; Q_2\vec{v}^{(2)} \leqslant x \ldots Q_k\vec{v}^{(k)} \leqslant x\; \rho(x,\vec{v}) \qquad (2)$$

where $Q_i\vec{v}^{(1)} \leqslant x$ *denotes a* *block* *of similar* *bounded* *quantifiers of the form* $\exists v_1^{(i)} \leqslant x\, \exists v^{(i)} \leqslant x \ldots \exists v_r^{(i)} \leqslant x$ *or* $\forall v_1^{(i)} \leqslant x\, \forall v_2^{(i)} \leqslant x \ldots \forall v_r^{(i)} \leqslant x$ *, and* ρ *is a quantifier free* $L_{=,+,\cdot}$ *formula.* A bound k on the number of changes of quantifier does exist if R U D = *linear space*, as is shown in Harrow [1978]. (Wilkie [1979] has proved it is not possible to put a bound on the number of individual quantifiers.)

On the other hand Paris and Dimitracopoulos [????] show that for every k there is some formula $\phi_k(x)$ of type (1) which cannot be defined by any bounded $L_{\leqslant,+,\cdot}$ formula with only k changes of quantifier. This is because it is possible to use numbers $u \leqslant x^{[\log_2 x]}$ as codes for sequences $v_1, \ldots, v_r \leqslant x$ and thus replace blocks of quantifiers of the form $Qv_i \leqslant x$ by single quantifiers $Qu \leqslant x^{[\log_2 x]}$, to obtain a sort of *universal formula* $\tau_k(x,n)$ with

$$\tau_k(x, \ulcorner\theta\urcorner) \iff \theta(x)$$

for all $\theta(x)$ of form (2) having Gödel numbers $\ulcorner\theta\urcorner$ sufficiently small compared with x.

THEOREM 1 RUD = $\mathcal{S} \Rightarrow$ NP $\neq$ co-NP.

Proof:

Suppose RUD = $\mathcal{S}$. Then RUD = *linear space* so there exists some fixed k such that every $R \in$ RUD can be defined by some formula of form (2). Thus there is some formula $\phi_k(x)$ of form (1) such that the set

$$S = \{x\colon \phi_k(x)\} \notin \text{RUD},$$

that is, there is some $S \in \Delta^P$ such that $S \notin$ RUD. But since NP $\subseteq$ NEXP = $\mathcal{S}$ = RUD, we see that $S \notin$ NP, so NP $\neq \Delta^P$, that is, NP $\neq$ co-NP.

Not only does assuming RUD = $\mathcal{S}$ ensure the existence of an upper bound on k in (2), but it also implies a lower bound, namely that some bounded $L_{\leqslant,+,\cdot}$ formula $\phi(x)$ is not equivalent to any *bounded diophantine formula*:

$$\exists v_1 \leqslant x\, \exists v_2 \leqslant x \ldots \exists v_r \leqslant x\, \lambda(x,\vec{v})$$

where λ is quantifier free and can be assumed without loss of generality to be an equation between two polynomials.

For all such formulas obviously define sets $S \in$ NP, and NP $\neq$ NEXP by a well known nondeterministic time hierarchy theorem.

Finally it should be remarked that *assuming the existence of an upper bound on* k *allows us to deduce* $\exists n(\Sigma_n^P = \Delta^P)$.

For suppose $S \in \Delta^P$. Then S is defined by some formula $\phi(x)$ of the form:

$$Q_1 y_1 \leq x^{[\log_2 x]^m} Q_2 y_2 \leq x^{[\log_2 x]^m} \ldots Q_s y_s \leq x^{[\log_2 x]^m} \lambda(x,\vec{y}).$$

We need only show that an upper bound can be placed on the s required here. Let $\theta(x,z)$ be the bounded $L_{\leq,+,\cdot}$ formula:

$$Q_1 y_1 \leq z\ Q_2 y_2 \leq z \ldots Q_s y_s \leq z\ \lambda(x,\vec{y}).$$

Choose a suitable <u>polynomial</u> $p(x,z)$ coding ordered pairs $<x,z>$ by the single numbers $p(x,z)$, with $p(x,z) \leq Cz^2$ (C constant) whenever $x \lesssim z$. Obviously we can find a bounded formula $\theta^*(w)$ such that:

$$\theta^*(p(x,z)) <=> \theta(x,z).$$

But <u>if</u> every bounded $L_{\leq,+,\cdot}$ formula is equivalent to one with only k changes of quantifier then $\theta^*(w)$ can be assumed to be of the form:

$$Q_1 \vec{v}^{(1)} \leq w\ Q_2 \vec{v}^{(2)} \leq w \ldots Q_k \vec{v}^{(k)} \leq w\ \rho(w,\vec{v})$$

where $\rho(w,\vec{v})$ is quantifier free. Using coding, each block $Q_i \vec{v}^{(i)} \leq w$ of quantifiers of the form $Q_i v_1^{(i)} \leq w\ Q_i v_2^{(i)} \leq w \ldots Q_i v_r^{(i)} \leq w$ can be replaced by a single quantifier of the form $Q_i y_i \leq w^r$. Thus $\theta^*(w)$ is equivalent to some formula of the form:

$$Q_1 y_1 \leq w^r\ Q_2 y_2 \leq w^r \ldots Q_k y_k \leq w^r\ \rho_1(w,\vec{y})$$

where $\rho_1(w,\vec{y})$ is a formula with all quantifiers bounded by $w^{[\log_2 w]^n}$ for some n, and <u>the number of these quantifiers is less than some bound independent of S</u>. (This is possible since there is a polynomial bound on the time required, given the y_i's, to compute the sequences $v_1^{(i)}, v_2^{(i)}, \ldots, v_r^{(i)}$ coded by the y_i's and then determine whether $\rho(w,\vec{v})$ is satisfied.)

Clearly it follows that

$$\phi(x) <=> \theta(x, x^{[\log_2 x]^m}) <=> \theta^*(p(x, x^{[\log_2 x]^m})) <=>$$

$$Q_1 y_1 \leq x^{[\log_2 x]^{m+1}} Q_2 y_2 \leq x^{[\log_2 x]^{m+1}} \ldots Q_k y_k \leq x^{[\log_2 x]^{m+1}} \rho_2(x,\vec{y})$$

where all the quantifiers in ρ_2 are bounded by $x^{[\log_2 x]^n}$ for some n, and their number does not depend on S.

In particular it follows from this that:

THEOREM 2. $RUD = \mathcal{S} \Rightarrow \exists n(\Sigma_n^P = \Delta^P)$,

(that is, the Stockmeyer polynomial time hierarchy collapses if $RUD = \mathcal{S}$).

Of course by theorem 1 we must have $n > 1$, but this could conceivably be consistent with the conjecture of Baker and Selman [1979] that $\Sigma_3^P = \Delta^P \neq \Sigma_2^P$.

BIBLIOGRAPHY

BAKER, T.P.; SELMAN, A.L. [1979] *A second step toward the polynomial hierarchy*. Theor. Comput. Sci. 8 (1979), 177-187.

BEL'TYUKOV, A.P. [1976] Zap. Naučn. Sem. Leningrad. Otdel. Mat. Inst. Steklov (LOMI) 60 (1976), 15-28.

BENNETT, J.H. [1962] *On Spectra*. Doctoral dissertation, Princeton University, Princeton, N.J..

BIRKHOFF, G.D.; VANDIVER, H.S. [1904] *On the integral divisors of* $a^n - b^n$. Ann. of Math. 5 (1904), 173-180.

BRUN, V. [1920] *Le crible d'Eratosthène et le théorème de Goldbach*. Skr. Norske Vid.-Akad. Kristiana I (1920) no. 3, 36 pp.

COOK, S.A.; RECKHOW, R.A. [1979] *The relative efficiency of propositional proof systems*. J. Symbolic Logic 44 (1979), 36-50.

DAVIS, M. [1973] *Hilbert's tenth problem is unsolvable*. Amer. Math. Monthly 80 (1973), 233-269.

DICKSON, L.E. [1904] *A new extension of Dirichlet's theorem on prime numbers*. Messanger of Mathematics 33 (1904), 155-161.

DIMITRACOPOULOS, C. [1980] *Matijasevič's Theorem and Fragments of Arithmetic*. Ph.D. thesis, University of Manchester.

ERDÖS, P. [1960] *Über die kleinste quadratfreie Zahl einer arithmetischen Reihe*. Monatsh. Math. 64 (1960), 314-316.

ERDÖS, P. [1980] *How many pairs of products of consecutive integers have the same prime factors*? Amer. Math. Monthly 87 (1980), 391-394.

ERDÖS, P.; SHOREY, T.N. [1976] *On the greatest prime factor of* $2^p - 1$ *for a prime* p *and other expressions*. Acta Arith. 30 (1976), 257-265.

FAGIN, R.G. [1975] *A spectrum hierarchy*. Z. Math. Logik Grundlagen Math. 21 (1975), 123-134.

GEL'FOND, A.O.; LINNIK, Yu. V. [1962] *Elementary Methods in the Analytic Theory of Numbers.* International Series of Monographs in Pure and Applied Mathematics, vol. 92. Pergamon Press 1962.

GRIMM, C.A. *A conjecture on consecutive composite numbers.* Amer. Math. Monthly 76 (1969), 1126-1128.

GRZEGORCZYK, A. [1953] *Some classes of recursive functions.* Rozprawy Matematyczne, 4 (1953).

HANF, W. [1965] *Model-theoretic methods in the study of elementary logic.* Symposium on the Theory of Models, North-Holland, Amsterdam (1965), 132-145.

HARDY, G.H.; WRIGHT, E.M. [1979] *An Introduction to the Theory of Numbers.* Fifth edition. Oxford University Press, Oxford 1979.

HARROW, K. [1978] *The bounded arithmetic hierarchy.* Inform. and Control 36 (1978), 102-117.

HOPCROFT, J.E.; ULLMAN, J.D. [1979] *Introduction to Automata Theory, Languages, and Computation.* Addison-Wesley (1979).

JONES, N.D.; SELMAN, A.L. [1974] *Turing machines and the spectra of first order formulas.* J. Symbolic Logic 39 (1974), 139-150.

LANGEVIN, M. [1975a] *Plus grand facteur premier d'entiers consécutifs.* C.R. Acad. Sci. Paris Sér. A,280 (1975), 1567-1570.

LANGEVIN, M. [1975b] *Plus grand facteur premier d'entiers voisins.* C.R. Acad. Sci. Paris Sér. A,281 (1975), 491-493.

LANGEVIN, M. [1979] *Facteurs premiers des coefficients binomiaux.* Séminaire Delange-Pisot-Poitou (Théorie des nombres) 20e année (1978/79), no. 27, 15 pp.

LEHMER, D.H. [1964] *On a problem of Störmer.* Illinois J. Math. 8 (1964), 57-79.

LESSAN, H. [1978] *Models of Arithmetic.* Ph.D. thesis, University of Manchester.

LIPSHITZ, L. [1978] *The diophantine problem for addition and divisibility.*

Trans. Amer. Math. Soc. 235 (1978), 271-283.

PARIS, J.B.; DIMITRACOPOULOS, C. [????] *Truth definitions for* Δ_0 *formulae*. Preprint.

PARIS, J.; HARRINGTON, L. [1977] *A mathematical incompleteness in Peano Arithmetic*. Handbook of Mathematical Logic (Ed. J. Barwise) North Holland, 1977.

PRACHAR, K. [1957] *Primzahlverteilung*. Springer-Verlag, Berlin 1957.

PRACHAR, K. [1958] *Über die kleinste quadratfreie zahl einer arithmetischen Reihe*. Monatsh. Math. 62 (1958), 173-176.

PRESBURGER, M. [1930] *Über die Vollständigkeit eines gewissen Systems der Arithmetik ganzer Zahlen, in welchem die Addition als einzige Operation hervortritt*. Sprawozdanie z I Kongresu Matematyków Krajów Słowiańskich (Comptes-rendus du I Congrès des Mathematiciens des Pays Slaves), Warsaw (1930), 92-101, 395.

RITCHIE, R.W. [1963] *Classes of predictably computable functions*. Trans. Amer. Math. Soc. 106 (1963), 139-173.

ROBINSON, J. [1949] *Definability and decision problems in arithmetic*. J. Symbolic Logic 14 (1949), 98-114.

SCHINZEL, A. [1961] *Remarks on the paper "Sur certaines hypothèses concernant les nombres premiers"*. Acta Arith. 7 (1961), 1-8.

SCHINZEL, A; SIERPÍNSKI, W. [1958] *Sur certaines hypothèses concernant les nombres premiers*. Acta Arith 4 (1958), 185-208.

SCHNIRELMANN, L. [1933] *Über additive Eigenschäften von Zahlen*. Math. Ann. 107 (1933), 649-690.

SCHOLZ, H. [1952] Problems, J. Symbolic Logic, 17 (1952), 160.

SHAPIRO, H.N. [1950] *On primes in arithmetic progression (II)*. Ann. of Math. 52 (1950), 231-243.

SHEPHERDSON, J.C. [1965] *Non-standard Models for fragments of number theory*, the Theory of Models, (Ed. Addison, J.W.; Henkin, L.; Tarski, A.,) North-Holland, Amsterdam (1965), 342-358.

STARK, H.M. [1973] *Effective estimates of solutions of some diophantine equations.* Acta Arith. 24 (1973), 251-259.

STOCKMEYER, L.J. [1976] *The polynomial time hierarchy.* Theor. Comput. Sci. 3 (1976), 1-22.

STØRMER, C. [1897] *Quelques theoremes sur l'équation de Pell* $x^2-Dy^2 = \pm 1$ *et leurs applications.* Skrifter Videnskabs. selskabet (Christiana) I, Mat-Naturv. Kl. (1897) no. 2. 48 pp.

SYLVESTER [1891] *On arithmetical series (Part I).* Messenger of Mathematics 21 (1891), 1-19. Alternatively this paper can also be found in *The Collected Mathematical Papers of James Joseph Sylvester*, Chelsea, New York (1973), Vol. IV, 687-703.

TARSKI, A. [1949] *Undecidability of group theory.* (Abstract) J. Symbolic Logic 14 (1949), 76-77.

THOMAS, W. [1978] *The theory of successor with an extra predicate.* Math. Ann. 237 (1978), 121-132.

WILKIE, A.J. [1977] *Some results and problems on weak systems of arithmetic.* Logic Colloguium '77 - Proceedings of the colloquium held in Wroclaw, August 1977. (Ed. Macintyre, A., Pacholski, L. Paris, J.) North-Holland, Amsterdam (1978).

WILKIE, A.J. [1979] *Applications of complexity theory to* Σ_0*-definability problems in arithmetic.* Model Theory of Algebra and Arithmetic. (Proceedings, Karpacz, Poland 1979.) Lecture Notes in Mathematics 834, Springer-Verlag, Berlin (1980).

WILKIE, A.J. [????] *Axioms for division.* To appear.

WOODS, A.R. [1977] *Algebraic Properties of Non-Standard Numbers.* M.Sc. thesis, Monash University, Clayton, Victoria.