# BOREL SETS AND CIRCUIT COMPLEXITY

Michael Sipser*
Department of Mathematics
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

## Abstract

It is shown that for every k, polynomial-size, depth-k Boolean circuits are more powerful than polynomial-size, depth-(k-1) Boolean circuits. Connections with a problem about Borel sets and other questions are discussed.

## I. Introduction

There are strong relationships between Turing machine resource based complexity and circuit complexity. In particular, it has been shown that Turing machine time, space, and reversal correspond to (uniform) circuit size, width, and depth, these relationships even holding simultaneously [B, P, Sch, H, R]. This suggests that one way to gain insight into polynomial time would be to study the expresive power of polynomial sized circuits. Perhaps the P=?NP question will be settled by showing that some problem in NP does not have polynomial sized circuits. Unfortunately, there are currently no known techniques for establishing significant lower bounds on circuit size for NP problems. The strongest results to date give linear lower bounds [Pa], and it does not seem likely that the ideas used there can go much beyond that.

We are investigating a new class of related, though hopefully more tractable, problems. We consider and/or circuits where the nodes have arbitrary fanin and fanout. The depth of a circuit is the length of a longest directed path from input to output. Thus, a circuit of depth two is either an "and" of "ors" or an "or" of "ands." (Equivalently, we could consider nodes with fixed fanin



*Supported by NSF grant MCS-81-05555 and Air Force grant AFOSR-82-0326.

and count the number of and/or alternations on paths.) Let the parity function $p_n$ of n Boolean variables be their sum modulo 2. It is easily seen that there are linear (in n) sized circuits computing $p_n$ of depth $0(\log n)$. On the other hand, the size of circuits computing $p_n$ of depth 2 must grow exponentially rapidly. In [FSS] we proved that the functions $p_n$ cannot be computed by circuits of constant depth and polynomial size.

Our proof, which gives a negative answer to the question, introduces the method of probabilistic restriction for simplifying a circuit while preserving the difficulty of the computed function. Novel probabilistic arguments were developed to carry out this simplification.

The finite-infinite analogy [S2] played a crucial role in the discovery of this proof. This analogy states that polynomial vs. exponential growth corresponds in some sense to countability vs. uncountability. The basis for this analogy is that exponential growth and uncountability both frequently arise from the power set construction. In view of this analogy and the above question, we asked: Is there an infinite parity function which can be computed by a circuit of finite depth and countably many gates? The notions of infinite circuit and infinite parity function can be specified in a natural way. The combinatorial methods used to answer this question [S2] form the core of the solution to the original question about finite parity functions.

The following connection indicates that this line of research may lead to progress on the P=NP question. If one considers the natural notion of nondeterministic circuit, then it follows that the class of languages $L$ accepted by polynomial size, depth 2, nondeterministic circuits is exactly the class "nonuniform NP." Showing that $L$ is not closed under complement would prove that NP≠co-NP and P≠NP. The infinite analog to the class $L$ is the class of analytic or $\Sigma_1^1$ sets [K]. It is a classical theorem in descriptive set theory that the analytic sets are not closed under complement. The proof is by diagonalization and, unfortunately, it does not appear to have a finitary analog. We are currently seeking a purely combinatorial proof of this theorem about analytic sets.

A similar situation existed for the following question: Are there functions computed by polynomial size, depth k circuits which cannot be computed by polynomial size, depth k-1 circuits? The infinite analog to this question is the classical hierarchy theorem for Borel sets. The standard proof of this theorem is also a diagonalization, which seems to be of no help in the finite case. In this paper we give a new, combinatorial proof of this theorem. The methods of this proof have enabled us to settle affirmatively the finitary question.

There has been other recent work giving strong lower bounds for weak computational models. Valiant [V] shows that monotone depth-three formulae must have exponential size when computing certain NP-complete functions. Borodin, Dolev, Fich, and Paul [BDFP] consider a more powerful model, the constant-width branching program, but can only prove superpolynomial lower bounds for a special case of width-two programs. In a somewhat older paper [S1], Sipser shows that there must sometimes be an exponential increase in size when deterministic sweeping automata simulate nondeterministic ones.

Unbounded-fanin circuits arise in several other contexts. They are a good model for PLA integrated circuits [CM, FSS] and for certain models of parallel computation [SV]. Immerman shows a relationship to expressiveness of first order sentences. Lower bounds for constant-depth circuits also have implications for establishing the alternating log-time hierarchy and the relativized polynomial-time hierarchy. Some of these connections are elaborated in sections III and IV.

## II. Main Results

There is a strong similarity in the proofs of the infinitary and finitary versions of the main theorem. Though the following are designed to be independent, the first is simpler and is a good introduction.

### A. Infinite case

Let $x=\{x_1, x_2, \ldots\}$ be the set of variables and $\bar{X}=\{x_1, \bar{x}_1, x_2, \bar{x}_2, \ldots\}$ be the set of <u>inputs</u> or <u>literals</u>. A literal is also called an $\bigwedge_0$-circuit or an $\bigvee_0$-circuit. For $i > 0$, an $\bigwedge_i$-circuit ($\bigvee_i$-circuit) is a nonempty collection of $\bigvee_{i-1}$-circuits ($\bigwedge_{i-1}$-circuits). We will sometimes refer to an $\bigwedge_i$-circuit or an $\bigvee_i$-circuit as simply an <u>i-circuit</u>. We say that one circuit <u>belongs to</u> another if it is hereditarily a member, i.e., "belong to" is the transitive closure of "member of." A $\bigvee_{i+1}$-circuit is also called a <u>$\Sigma_i$-circuit</u> if all of its 1-circuits are finite. A $\Sigma_0$-circuit is a $\Sigma_1$-circuit which is entirely finite. Dually we define $\Pi_i$-circuits. There are also two <u>constant circuits</u> called <u>0</u> and <u>1</u>. The <u>depth</u> of an i-circuit is i and its <u>size</u> is the cardinality of the set of circuits belonging to it. All circuits in this section have countably infinite size.

Letting $\Sigma=\{0,1\}$ and $\Sigma^\omega$ be the collection of infinite 0,1 sequences, we see that a circuit defines a function $f:\Sigma^\omega \to \Sigma$ in the natural way. We also say that a circuit defines the set $A \subseteq \Sigma^\omega$ if $A=f^{-1}(1)$.

We review the definition of the Borel sets in the standard product topology on $\Sigma^\omega$ [K]. An <u>interval</u> is the set of extensions of a finite sequence. An <u>open</u> or $\Sigma_1$ set is a union of intervals. A $\Sigma_i$ set is the countable union of $\Pi_j$ sets, $j < i$, where these are the complements of $\Sigma_j$ sets. A set is <u>Borel</u> if it is $\Sigma_i$ for some $i < \omega_1$.

<u>Theorem</u>: A set is $\Sigma_i$-circuit definable iff it is $\Sigma_i$-Borel.

<u>Proof</u>: Straightforward induction.

Let $\alpha \in \Sigma^\omega$ and $\alpha(i) \in \Sigma$ be the ith bit. Let $\langle i_1, \ldots, i_j\rangle$ be a d-ary pairing function from $N^d \to N$.

<u>Definition</u>: For each d, the function $F_d:\Sigma^\omega \to \Sigma$ is defined as follows: For $\alpha \in \Sigma^\omega$:

$$F_d(\alpha)=1 \text{ iff } \exists i_1 \forall i_2 \exists i_3 \ldots Q i_d$$

$$[\alpha(\langle i_1, i_2, \ldots, i_d\rangle)=1]$$

Clearly, $F_d$ is definable by a $\Sigma_d$-circuit.

<u>Theorem</u>: For every $d > 0$, the function $F_d$ is definable by a $\Sigma_d$-circuit but not by a $\Sigma_{d-1}$-circuit.

<u>Proof</u>: By induction on d.

<u>Basis</u>: (d=1) Immediate.

<u>Induction</u>: Obtain a $\Sigma_{d-2}$-circuit for $F_{d-1}$ from a $\Sigma_{d-1}$-circuit for $F_d$.

<u>Definition</u>: [$S_2$, FSS] A (<u>partial</u>) <u>restriction</u> is a (partial) function $\rho:X \to \{0,1,*\}$. We extend $\rho$ in the natural way to $\bar{X}$ by assigning

$$\rho(\bar{x}_i) = \begin{cases} * \text{ if } \rho(x_i) = * \\ \neg\rho(x_i) \text{ if } \rho(x_i) \neq * \end{cases}$$

Restrictions are used to fix some of the inputs to a function while leaving others (those assigned *) variable. For any restriction $\rho$ and function $f:\Sigma^n \to \Sigma$,

$\rho$ induces a function $f^{\rho}:\Sigma^{k}\to\Sigma$ where k is the (possibly infinite) number of literals assigned *. If $\alpha\in\Sigma^{k}$, $f^{\rho}(\alpha)=f(\alpha^{\rho})$ where $\alpha^{\rho}$ is $\rho(\alpha_1)\rho(\alpha_2)$ . . . with the k *'s replaced by the k elements of $\alpha$.

In the same spirit that restrictions induce new functions, they also induce new circuits. Say a partial restriction $\rho$ <u>forces an $\wedge$-circuit($\vee$-circuit) to be 0</u> if $\rho$ forces any (all) of its members to be 0. Dually, $\rho$ <u>forces an $\wedge$-circuit ($\vee$-circuit) to be 1</u> if $\rho$ forces all (any) of its members to be 1. Also, $\rho$ forces a 0-circuit to be a 0 or 1 if it assigns it a 0 or 1. Given a circuit C, a restriction $\rho$ <u>induces</u> circuit $C^{\rho}$. If $\rho$ forces C, then $C^{\rho}$ is the constant circuit equal to the forced value. If C is unforced, then $C^{\rho}$ is the collection of the unforced members of C.

<u>Lemma</u>: For any circuit C, function f, and restriction $\rho$, if C computes f, then $C^{\rho}$ computes $f^{\rho}$. □

<u>Definition</u>: A <u>row</u> is a set $\{<i_1, \ldots, i_{d-1}, j>: j\in N\}$, and a <u>box</u> is a set $\{<i_1, \ldots, i_{d-2}, j, k>: j,k\in N\}$ for some $i_1, \ldots, i_{d-1}\in N$. Assume d is even. A restriction <u>kills</u> a row r if for some $i\in r$, $\rho(x_i)=0$, and it <u>preserves</u> r if for a unique $i\in r$, $\rho(x_i)=*$ and for every other $j\in r$, $\rho(x_j)=1$. A restriction is <u>good</u> if:

1) every row is killed or preserved

2) every box contains infinitely many preserved rows.

If d is odd, exchange the roles of 0 and 1.

<u>Lemma</u>: If $\rho$ is a good restriction, then $F_d^{\rho}$ is $F_{d-1}$, on a suitably chosen (d-1)-ary pairing function.

<u>Proof</u>: Use the following (d-1)-ary pairing function. Number all of the starred positions. Let $<i_1, \ldots, i_{d-1}>$ be the number of $<i_1,\ldots,i_{d-2},j,k>$ where j is the index of the $i_{d-1}$th preserved row in box $i_1, \ldots, i_{d-2}$ and k is the position of the * in that row. □

<u>Lemma</u>: For every $\Sigma_{d-1}$-circuit C, there is a good restriction inducing a circuit equivalent to one all of whose 2-circuits are finite.

<u>Proof</u>: We construct $\rho$ in stages, alternating between extending the definition to ensure that the next infinite 2-circuit is induced to a finite one, and ensuring that the conditions of goodness are being met. Assume d is even, the odd case is dual. Initially, the partial restriction $\rho$ is everywhere undefined. An essential feature of the construction is that at every stage, $\rho$ assigns * to only finitely many new variables. Let $C_1$, $C_2$, ... be an enumeration of the infinite 2-circuits of C. Note that each is a $\Pi_i$-circuit. Begin at stage 1.

<u>Stage i</u>. First, we extend $\rho$ so that the circuit induced from $C_i$ is equivalent to a finite one. Let s be the number of *'s currently assigned. Let $\rho_1, \rho_2, \ldots, \rho_{2^s}$ be the $2^s$ partial restrictions obtained from $\rho$ by setting these *'s to 0's and 1's in all possible ways. When we now extend $\rho$, this implicitly extends the $\rho_i$ in the obvious way. Perform substages 1 through $2^s$.

<u>Substage j</u>: Our goal is to extend $\rho$ in such a way that $\rho_j$ forces $C_i$. If that is already true, go to substange j+1. Otherwise, there must be a member D of $C_i$ which is not forced by $\rho_j$. Since D is an $\vee$-circuit, all of its members are either forced to be 0 or unforced. D is a 1-circuit, and thus all of its members are literals. Simply extend $\rho$ to assign 0 to all unassigned literals of D.

This extended $\rho_j$ forces $C_i$. Go to substage j+1.

Upon completion of substage $2^s$, it is not hard to show that $\rho$ has the property that for any restriction $\sigma$ extending $\rho$, the function computed by $C_i\sigma$ depends only upon these s *'d literals. This finite function can be computed by a finite 2-circuit.

Second, we further extend $\rho$ so that the conditions of goodness are being met. For each row containing variables assigned above, we extend $\rho$ so it assigns 0 to all other variables in that row, thereby killing it. In each box containing a row so killed we choose a new row and preserve it by assigning * to its first variable and 1 to all others. This adds only finitely many new *'s. Go to stage i+1.

Upon completion of all stages, extend $\rho$ to preserve all remaining untouched rows. Replacing all 2-circuits in $C^\rho$ by their finite equivalents we obtain the desired circuit having only finite 2-circuits. □

By the preceeding two lemmas we can convert a $\Sigma_{d-1}$-circuit defining $F_d$ into a $\Sigma_{d-1}$-circuit all of whose 2-circuits are finite and which defines $F_{d-1}$. Inverting the finite 2-circuits and merging non-alternating levels we obtain a $\Sigma_{d-2}$ circuit defining $F_{d-1}$ satisfying the induction step. This completes the proof.

### B. Finite Case

For each n, let $X_n=\{x_1, x_2, \ldots , x_n\}$ be the set of _inputs_ or _literals_. We define $\bigwedge_i$- and $\bigvee_i$-, and i-circuits exactly as in the infinite case and use the same notions of circuit size and depth.

To exhibit complexity growth rates we consider families of circuits $C_1$, $C_2$, ... , where $C_n$ defines a function of n variables. The sizes of the circuits in these families will be limited to growth rates at most a polynomial in n. A $\Sigma_i$-family is one containing only $\bigvee_{i+1}$-circuits all of whose 1-circuits are bounded in size uniformly in n. $\Pi_i$-families are dually defined. For simplicity we will sometimes speak of representative members of families as $\Sigma_i$- or $\Pi_i$-circuits.

For $d,m > 0$ and $n=m^d$, let $N=\{1, \ldots , n\}$ and $M=\{1, \ldots , m\}$ and $\langle i_1, \ldots , i_d\rangle$ be a d-ary pairing function from $M^d \to N$. Let $\alpha \in \Sigma^m$, $\Sigma=\{0,1\}$ and $\alpha(i) \in \Sigma$ be the ith bit.

_Definition_: For $d,m > 0$ and $n=m^d$, we define $F_d^n$: $\Sigma^n \to \Sigma$ as follows: For $\alpha \in \Sigma^n$,

$$F_d^n(\alpha)=1 \text{ iff } \exists i_1 < m \ \forall i_2 < m \ \ldots \ Qi_d < m$$

$$[\alpha(\langle i_1, \ldots , i_d\rangle)]$$

If n is not a power of d, the $F_d^n$ is everywhere 0. Clearly, the functions $F_d^n$ are definable by a $\Sigma_d$-circuit family.

_Theorem_: For every $d > 0$, the functions $F_d^n$, $n > 0$ are definable by a $\Sigma_d$-circuit family but not by a $\Sigma_{d-1}$-circuit family.

_Proof_: By induction on d.

_Definition_: [FSS] A _restriction_ is a function $\rho$: $X_n \to \{0,1,*\}$. We extend $\rho$ to $\overline{X}_n$ as in the infinite case. Restrictions induce functions and circuits as in the infinite case.

_Definition_: As in the infinite case, a _row_ is a set $\{\langle i_1, \ldots , i_{d-1}, j\rangle: j \in M\}$ and a _box_ is a set $\{\langle i_1, \ldots , i_{d-1},j,k\rangle: j,k \in M$ for some $i_1, \ldots , i_{d-1} \in M$. Assume d is even. A restriction $\rho$ _kills_ a row r if for some $i \in r$, $\rho(x_i)=0$, it _satisfies_ r if for every $i \in r$, $\rho(x_i)=1$, and it _preserves_ r if for a unique $i \in r$, $\rho(x_i)=*$ and for every other $j \in r$, $\rho(x_j)=1$. A restriction is _good_ if

1) Every row is killed or preserved;

2) every box contains $> \sqrt{m}/2$ preserved rows.

Define dual notions for odd d by exchanging 0's and 1's.

Lemma B1: A good restriction $\rho$ can be strengthened, in the sense of changing some *'s to 0's and 1's, to a restriction that induces from the function $F_d^n$ the function $F_{d-1}^{n'}$ where $n'=(\sqrt{m}/2)^{d-1}$, for a suitably chosen (d-1)-ary pairing function.

Proof: Let $m'=(\sqrt{m}/2)$. We first strengthen $\rho$ so that all boxes which are "out of bounds," i.e., have coordinates $i_1, \ldots, i_{d-2}$ where some $i_j > m'$, do not affect the value of the function, by either killing or satisfying all remaining preserved rows in these boxes depending on whether j is odd or even.

We then number the preserved rows in the remaining boxes and either kill or satisfy those numbered larger than m' depending on whether d is even or odd. Number all the starred positions. The new pairing function is given by assigning $\langle i_1, \ldots, i_{d-1}\rangle$ the number of $\langle i_1, \ldots, i_{d-2}, j, k\rangle$ where j is the $i_{d-1}$th row in the box given by $i_1, \ldots, i_{d-2}$ and k is the position of the * in that row. ◻

Now, we show that there are good restrictions converting $\Sigma_d$-circuit families into $\Sigma_{d-1}$-circuit families. A deterministic method for obtaining restrictions with this property is not known. Instead, we give a certain probability distribution on restrictions and show that a random one has a nonzero chance of satisfying all conditions.

Probability distribution: First, independent Bernoulli trials are run for each row to determine which are to be preserved, where Pr[row is preserved] = $1/\sqrt{m}$. For each preserved row, the restriction assigns * to the first variable and 1 (0) to the remaining variables if d is even (odd). For each variable in the remaining rows, an independent trial is run to determine whether it gets assigned a 0 or a 1, where Pr[variable is assigned 0] = Pr[variable is assigned 1] = 1/2.

The next lemmas show that a random restriction is likely to have the desired properties.

Lemma B2: The probability that a randomly chosen restriction is not good is $< 2^{-O(\sqrt{m})}$.

Proof: The expected number of preserved rows per box is $\sqrt{m}$ and the expected number of 1's and 0's in the other rows is m/2. A straightforward calculation shows that significant deviation from this in any row or box is very unlikely. ◻

Lemma B3: For every c and k, there is a constant $b_c$, such that for any 2-circuit A of size at most $n^k$ all of whose 1-circuits are of size at most c, if a random $\rho$ is chosen:

$$\Pr[A \text{ depends upon } > b_c \text{ inputs}] = o(n^{-k})$$

Proof: By induction on c. To simplify the basis of the induction, we permit $\bigwedge_1$- and $\bigvee_1$-circuits to be empty, representing the constant functions 1 and 0. They are thus forced by any restriction.

Basis: (c=0, show $b_o$ exists) $b_o=0$.

Induction step: (show $b_c$ exists given $b_{c-1}$) Assume A is an $\bigvee_2$-circuit, the $\bigwedge_2$ case is dual. Call two inputs independent if they appear on different rows, and say two 1-circuits are independent if the inputs of one are all independent of the inputs of the other.

Let $b = 2k \cdot 4^c$. Consider two cases, one in which A is wide, i.e., has more than $b \ln n$ pairwise independent rows, and the other in which A is narrow.

Case 1: A is wide. Since independent variables are assigned by the random restriction independently:

$$
\begin{aligned}
&\Pr[\text{A is not forced}] \\
&\leq \Pr[\text{ no member of A is forced to 1}] \\
&\leq \Pr[\text{a 1-circuit of size} \leq c \text{ is not forced to 1}]^{b \ln n} \\
&\leq (1-4^{-c})^{b \ln n} \\
&= n^{-b\ln(1-4^{-c})} \\
&\leq n^{-b4^{-c}} \\
&= o(n^{-k}) \text{ for } b=2k\cdot 4^c
\end{aligned}
$$

Case 2: A is narrow. Choose a maximal collection of pairwise independent 1-circuits in A, and let H be the set of rows that contain an input appearing in this collection. Since A is narrow, $|H| \leq bc \ln n$ and H hits (contains a row containing an input appearing in) each of A's 1-circuits. Let h be the number of preserved rows appearing in H, and let $\ell=2^h$. Each preserved row contains exactly one *. Let $\rho_1, \rho_2, \ldots, \rho_\ell$ be the $2^h$ restrictions obtained by setting the *'s appearing in the preserved rows in H to 0 and 1 in all possible ways. The value of $A^\rho$ is determined by the values of $A^{\rho_1}, \ldots, A^{\rho_\ell}$ and the h *'ed inputs. Furthermore, since H hits every 1-circuit, the 1-circuits in each $A^{\rho_i}$ are of size at most c-1. Thus, by the induction hypothesis, $\Pr[A^{\rho_i}$ depends upon $>b_{c-1}$ inputs] is $o(n^{-k})$.

Since H contains at most $bc \ln n$ rows and each row has a $1/\sqrt{m}$ chance of being preserved, an easy calculation verifies that $\Pr[h > 4k]$ is $o(n^{-k})$. If $A^\rho$ depends upon more than $4k+\ell\cdot b_{c-1}$ inputs, then either $h>4k$ or one of the $A^{\rho_i}$ depends upon more than $b_{c-1}$ inputs. Since $\ell < 2^{4k}$ if $h < 4k$, letting $b_c = 4k+2^{4k}\cdot b_{c-1}$, we have:

$$
\begin{aligned}
&\Pr[A^\rho \text{ depends upon } >b_c \text{ inputs}] \\
&\leq o(n^{-k})+2^{4k}\cdot o(n^{-k}) \\
&= o(n^{-k}) \qquad \square
\end{aligned}
$$

A consequence of this lemma is that given any $\Sigma_i$-circuit of size at most $n^k$ whose 1-circuits are of size at most c, the probability that a random restriction fails to induce a circuit all of whose 2-circuits depend upon $b_c$ inputs is o(1). Such a circuit can be converted into an equivalent $\Sigma_{i-1}$-circuit using the distributive law to exchange the order of the lowest two levels and merging nonalternating levels. This increases the size of the circuit by only a constant factor depending upon $b_c$, and its 1-circuits are of size at most $2^{b_c}$.

Hence, given a $\Sigma_{d-1}$-circuit for $F_d$, with nonzero probability, a random restriction is good and induces a circuit equivalent to a $\Sigma_{d-2}$-circuit. By lemma B1, this may be converted into a $\Sigma_{d-2}$-circuit computing $F_{d-1}$. This computes the proof of the theorem.

## III. The alternating log-time hierarchy

A consequence of the preceeding theorem is that the levels of the alternating log-time hierarchy are all distinct. Alternating Turing machines that run in sublinear time were defined by Chandra, Kozen, and Stockmeyer [CKS] and investigated by [Rv, DC, C2]. They have a random access mechanism to read their input tape. Alternating machines that run in logarithmic time form a hierarchy within LOGSPACE analogous to the polynomial time hierarchy.

For the following theorem, assume that the random access input query tape is erased after each use. The more general model can be converted into this one at the cost of a few alternations.

Theorem: If $L \in \Sigma_d$-LOGTIME, then there is a $\Sigma_d$-circuit family accepting the finite segments of L.

Theorem: $L_d = \{\alpha: |\alpha| = 2^{nd} \text{ and } F_d(\alpha) = 1\} \in \Sigma_d$-LOGTIME.

Corollary: For each $d > 0$, $\Sigma_{d-1}$-LOGTIME $\subsetneq \Sigma_d$-LOGTIME.

## IV. Other connections

A. A more careful analysis shows that the lower bound established in section II is barely superpolynomial: To compute $F_d$ with a (d-1)-circuit, at least $n^{\log^{(3d)} n}$ size is necessary, where $\log^{(i)} n$ is the i'th iterate of the log function. An improved lower bound would establish the existence of oracles separating the levels of the polynomial time hierarchy.

Theorem: If for all d, $F_d$ requires more than $n^{\text{poly} \log n}$ size with depth d-1, then there is an oracle A such that for all d, $\Sigma^A_{d-1}$-PTIME $\subsetneq \Sigma^A_d$-PTIME.

The proof involves the methods of [FSS] and [A].

B. Stockmeyer and Vishkin [SV] have shown that there is a strong relationship between certain models for parallel computation and the unbounded fan-in circuit model. Specifically, the concurrent-read, concurrent-write parallel RAM (CRCW PRAM) with addition but not multiplication can simulate and be simulated by circuits, with a constant factor increase in time and polynomial increase in number of processors. Thus, viewing the depth of a circuit as its time and size as number of processors and considering PRAM's to be uniform, an immediate consequence of this result and section II is:

Theorem: For $t > 0$, CRCW PRAM's operating in constant time t with a polynomial number of processors have strictly more power than do those operating in time t-1 with a polynomial number of processors.

C. Immerman [I] establishes connections between the expressive power of constant-depth circuits and that of sentences in first order predicate calculus.

## References

[A] D. Angluin, "On counting problems and the polynomial time hierarchy," TCS, to appear.

[BS] T. Baker and A. Selman, "A second step toward the polynomial time hierarchy," 17th FOCS, 1976, 71-75

[B] A. Borodin, "On relating time and space to size and width," SIAM J. Comput. 6, 1977, 733-744.

[BDFP] A. Borodin, D. Dolev, F. Fich, and W. Paul, "Bounds for width two branching programs," 15th STOCL.

[CSV] A. Chandra, L. Stockmeyer, V. Vishkin, "A complexity theory for unbounded fan-in parallelism." 23rd FOCS, 1982, 2-13.

[CKS] A. Chandra, D. Kozen, and L. Stockmeyer, "Alternation," JACM, 28, 1981

[C2] S. Cook, "Towards a complexity theory of synchronous parallel computation," Univ. of Toronto Computer Science Dept. Tech. Report 141/80, 1980.

[DC] P. Dymond and S. Cook, "Hardware complexity and parallel computation, " 21st FOCS, 360-372, 1980.

[FSS] M. Furst, J. B. Saxe, and M. Sipser, "Parity, circuits, and the polynomial time hierarchy," 22nd FOCS, 1981, 260-270.

[H] J. Hong, "On simularity and duality of computation," 21st FOCS, 1980, 348-359.

[I] N. Immerman, "Languages which capture complexity classes," 15th STOC, 1983.

[K] K. Kuratowski, Topology, Academic Press, 1966.

[MC] C. Mead and C. Conway, Introduction to VLSI Systems, Addison-Wesley, Reading, Mass., 1980.

[Pa] W. Paul, "A 2.5n lower bound for the combinatorial complexity of Boolean functions," 7th SIGACT, 1975, 27-36.

[P] N. Pippenger, "On simultaneous resources bounds," 20th FOCS, 1979, 307-311.

[R] W. L. Russo, "On uniform circuit complexity," 20th FOCS, 1979, 312-318.

[Ru] W. L. Russo, "Tree size bounded alternation," JCSS, 21, 218-235, 1980.

[Sch] C. Schnorr, "The network complexity and Turing Machine complexity of finite functions," Acts Informa., 7, 1976, 95-107.

[SV] L. J. Stockmeyer and V. Vishkin, "Simulation of parallel random access machines by circuits," RC 9362, IBM research report, 1982.

[S1] M. Sipser, "Lower bounds on the size of sweeping automata," JCSS, 21

[S2] M. Sipser, "On polynomial vs. exponential growth," MIT Tech. Report, to appear.

[V] L. Valiant, "Exponential lower bounds for restricted monotone formulae," 15th STOC.