INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

U·M·I

University Microfilms International
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
313.761-4700 800.521-0600

Order Number 9130348

Circuits in bounded arithmetic

Mantzivis, Georgios Spyridon, Ph.D.

City University of New York, 1991

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

CIRCUITS IN BOUNDED ARITHMETIC

by

Georgios Spyridon Mantzivis

A dissertation submitted
to the Graduate Faculty in Mathematics
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy,
The City University of New York.

1991

This manuscript has been read and accepted for the Graduate Faculty in Mathematics in satisfaction of the requirement for the degree of Doctor of Philosophy.

May 9, 1991 — Attila Máté
Date — Chair of Examining Commitee

May 10, 1991 — Martin Moskovitz
Date — Executive Officer

Attila Máté

Kenneth McAloon

Rohit Parikh

Supervisory Committee

The City University Of New York

Acknowledgements

I wish to thank Professor Nck. Phillips for teaching me mathematical logic in 1977, my adviser, Professor Attila Máté, for reading this dissertation and pointing out errors, Professors R. Parikh and K. McAloon for being on my dissertation committee, Professors G. Takeuti and P. Clote for endorsing parts of this work for publication, and my father, Nck. Mantzivis, for his financial assistance.

Contents

Acknowledgements iv

1 Introduction, outline and preliminaries **1**

Fundamental lemma for bounded circuits 6

2 PARITY is not $\mathcal{SB}_1^{\mathrm{N}}$ definable **11**

Sharp truth lemma 17

3 The set of primes is not $\mathcal{SB}_1^{\mathrm{N}}$ definable **19**

4 PARITY is not $\mathcal{SB}_1^{\mathrm{N}}(\mathbf{BIT})$ definable **22**

Wide lemma 25

5 INDEXPARITY is not $\mathcal{SB}_1^{\mathrm{N}}(\mathbf{BIT}, \mathbf{CARD})$ definable **37**

Multicircuit-Odd-Even lemma 39

6 PARITY is not $\mathcal{SB}_1^{\mathrm{N}}(\mathbf{trunc})$ definable **44**

Cut-and-paste cluster lemma 50

7 INDEXPARITY is not $\mathcal{SB}_1^{\mathrm{N}}(\mathbf{trunc}, \mathbf{CARD})$ definable **66**

G-lemma 71

The Cluster Shift Argument 74

8 PARITY is not $\mathcal{SB}_1^{\mathrm{N}}(\mathbf{trunc}, \mathbf{flip})$ definable **86**

References **101**

Chapter 1

Introduction, outline and preliminaries

Introduction

We shall be working in the area of theoretical computer science known as complexity theory and the area of logic known as bounded arithmetic.

One of the concerns in complexity theory is finding lower bounds for the computation of certain functions. The reason for this is the successful application of lower bounds for **PARITY** to obtain separation results for the relativized Meyer-Stockmeyer polynomial-time hierarchy, [Sto77, FSS84]. Finer separation results were then obtained, first in [Yao85] and subsequently in [Hås86]. At this point interest was generated in obtaining lower bounds for other functions, [Raz87, Smo87].

In the last twenty or so years the area of bounded arithmetic – the natural offspring of Peano Arithmetic – has become a full fledged branch of logic, with R. Parikh [Par71] who for the first time introduced and studied bounded arithmetic. That system is now known as $I\Delta_0$. A variety of aspects of $I\Delta_0$ have been researched as is clear from work in [Ajt88, DG82, DP82, PW87, Woo81]. One of the earliest studies, however, of Δ_0 sets can be found in [Ben62].

A fruitfull variation on $I\Delta_0$ was introduced by S. Buss in [Bus86], with the purpose in mind to study "logical" aspects of the $P = NP$ problem. His work is mainly proof-theoretic

and he relates polynomial-time computability with provability. Since then contributions to Buss's system include [CT86, Tak90, Fer88, PWW88, KPT].

By working in the intersection of the above areas one hopes to successfully combine the uniformity of formulae, elementary combinatorial techniques, and the results about non-uniform circuits in order to tackle questions concerning polynomial-time predicates. Examples of such work can be found in [Imm87, Ruz81].

Some fundamental questions about the Meyer-Stockmeyer hierarchy, $I\Delta_0$, and Buss's system still remain unanswered:

1. Do any of the respective hierachies of sets (languages), i.e. the polynomial-time hirarchy, the Δ_0-hierarchy or Buss's-hierarchy, *collapse*?

2. Is any of $I\Delta_0$ or Buss's system finitely axiomatizable?

3. Does any of $I\Delta_0$ or Buss's system logically collapse?

Outline

We are interested in classifying sets(languages) definable by *sharply bounded* formulae introduced by Samuel R. Buss in [Bus86]. These clearly define polynomial-time predicates. We express formulae as circuits with binary input and then use a blend of college algebra and probabilistic restrictions. Restrictions and probabilistic restrictions were introduced in [FSS84]. Cylinders, as found in [Ajt83] are the same objects as restrictions.

In this chapter we give a definition of *sharply bounded closures*, i.e. $\mathcal{SB}_1^{\mathbf{N}}$, $\mathcal{SB}_1^{\mathbf{N}}(\mathbf{BIT})$, $\mathcal{SB}_1^{\mathbf{N}}(\mathbf{BIT}, \mathbf{CARD})$, $\mathcal{SB}_1^{\mathbf{N}}(\text{trunc})$, $\mathcal{SB}_1^{\mathbf{N}}(\text{trunc}, \mathbf{CARD})$, and $\mathcal{SB}_1^{\mathbf{N}}(\text{trunc}, \mathbf{flip})$. For completeness then, we state and prove a close variant of the main result about constant depth and polynomial size circuits, as found in [FSS84].

In chapter 2 we prove that **PARITY** is not in $\mathcal{SB}_1^{\mathbf{N}}$ by showing that $\mathcal{SB}_1^{\mathbf{N}}$ is contained in AC^0, the class of all relations and functions definable by bounded depth and polynomial size circuitry, and then applying [FSS84].

In chapter 3, via an easy application of the prime number theorem, we show that the set of primes is not $\mathcal{SB}_1^{\mathbf{N}}$.

All of the subsequent results are proved by showing that there are non-trivial domains

(cylinders or G-cylinders) in which the predicates at hand are oblivious, i.e. take on only one truth value.

In chapter 4 we prove that **PARITY** is not in $\mathcal{SB}_1^{\mathbf{N}}(\mathbf{BIT})$. A problem appears: unlike in the case of $\mathcal{SB}_1^{\mathbf{N}}$, here we have knowledge of the individual bits of our terms; and our terms involve multiplication, which as a function is *not* AC^0 (it is unknown whether the *relation* $x \cdot y = z$ is AC^0 !). This is overcome by the wide lemma which is used in all subsequent chapters and which provides with domains in which the *function* of multiplication becomes AC^0. In chapter 5 we prove that **INDEXPARITY** is not in $\mathcal{SB}_1^{\mathbf{N}}(\mathbf{BIT}, \mathbf{CARD})$. This follows from the fact that the predicates **BIT** and **CARD** do not interact. **INDEXPARITY** is the **PARITY** taken over the set of odd bits of the input.

In each of the following chapters we show by induction on the complexity of the terms that these terms have a certain "bit-finite" property. This property then enables us to convert the predicate at hand into a feasible circuit, in the appropriate non-trivial domain. So, in chapters 6 and 8 we prove that **PARITY** is not in $\mathcal{SB}_1^{\mathbf{N}}(\text{trunc})$ and $\mathcal{SB}_1^{\mathbf{N}}(\text{trunc}, \mathbf{flip})$, respectively.

In chapter 7 we prove that **INDEXPARITY** is not in $\mathcal{SB}_1^{\mathbf{N}}(\text{trunc}, \mathbf{CARD})$. This is done in a way similar to chapter 6. However, it was necessary to introduce restrictions (cylinders) over groups of permutations, rather than restrictions over the bits of the input.

In view of the "slow" enhancements of the basic set of terms and predicates of bounded arithmetic, it is noteworthy that we are obliged to use a *bounded* quantifier should we decide to express the predicate $\mathbf{BIT}(x, y)$, the y-th bit in the binary expansion of x, as a formula of bounded arithmetic, [Bus86, page 41, (g)]. Indeed, if we pass to the class of unary Σ_1^{b} predicates, i.e. formulae which admit *bounded* existential quantifiers, then we have formulae defining each polynomial time computable predicate, and perhaps more.

The point, however, is not that **PARITY** or **INDEXPARITY** are not definable by respective sharply bounded formulae. It is rather the *reduction to bounded depth circuits within appropriate, nontrivial domains* of the respective classes that we are labouring for. In chapter 7 one realizes that *appropriate domains* can mean quite different things, and indeed should: cardinality is *never* constant in a *non-trivial* classical restriction(cylinder).

We formulate the following general question then: with Q a polynomial time computable predicate, and $\mathcal{P}$ a property on subsets of $\{0,1\}^\ell$, ℓ positive integer, we write $Q/\mathcal{P} \in AC^0$,

if there are:

1. $\mathbf{C} = \{C_\ell\}$ circuit family of constant depth and polynomial size, and
2. $\mathbf{D} = \{D_\ell\}$ domains with $D_\ell \subseteq \{0,1\}^\ell$ and $\mathcal{P}(D_\ell)$

such that for each ℓ we have that $\forall x \in D_\ell[Q_\ell(x) \Leftrightarrow C_\ell(x) = 1]$; with $\mathcal{C}$ a class of polynomial time computable predicates and $\mathcal{P}$ a property of domains, we write $\mathcal{C}/\mathcal{P} \subseteq AC^0$ if for every predicate $Q \in \mathcal{C}$ we have that $Q/\mathcal{P} \in AC^0$; given a class $\mathcal{C}$ of polynomial time computable predicates now, we ask whether there exists a non-trivial property $\mathcal{P}$ such that $\mathcal{C}/\mathcal{P} \subseteq AC^0$. The candidates we have in mind for $\mathcal{C}$ are, of course, sharply bounded closures of finitely many polynomial time computable functions or predicates.

Preliminaries

Here we introduce some notation and, for completeness, give a proof of the fundamental result in [FSS84].

Notation

The following notation and definitions are to be noted for the rest of the sequel:

1 a. The language in [Bus86] consists of the usual function and relation symbols of Peano Arithmetic plus the following function symbols:

1. $|x|$, to be interpreted as the *length of the binary expansion of* x, i.e. $|x| = \lceil \log_2(x+1) \rceil$,
2. $\lfloor x/2 \rfloor$, to be interpreted as *shift right*, and
3. $x\#y$, to be interpreted as $2^{|x|\cdot|y|}$.

Terms in this enhanced language will be called *terms of bounded arithmetic.*

1 b. $\Phi(\vec{x})$ is *sharply bounded* if Φ is a formula of bounded arithmetic and all quantifiers in Φ are of the form $\forall z < |t(\vec{x})|$ and $\exists z < |t(\vec{x})|$, where t is a term of bounded arithmetic.

It is convenient to regard sharply bounded formulae in prenex form. It is easy to see that any sharply bounded formula is *logically* equivalent to a sharply bounded formula in prenex form.

1 c. Often we shall enhance the language of bounded arithmetic by additional predicates and function symbols. The class of sharply bounded formulae in the language of bounded arithmetic is denoted by $\mathcal{SB}_1^{\mathrm{N}}$. Given new predicates $P_1, \ldots, P_n$ and new function symbols $f_1, \ldots, f_m$ we denote by

$$\mathcal{SB}_1^{\mathrm{N}}(f_1, \ldots, f_m, P_1, \ldots, P_n)$$

the class of sharply bounded formulae in the language of bounded arithmetic enhanced by $P_1, \ldots, P_n$ and $f_1, \ldots, f_m$.

The rest of this section is devoted to notation to be used in the proof of the fundamental circuit lemma.

1 d. $[n] = \{0, 1, \ldots, n-1\}$, n a positive integer.

1 e. Let $X = \{x_1, \ldots, x_\ell\}$, ℓ a positive integer. A *circuit in X over the boolean base* $\{0, 1, \bigvee, \bigwedge, \neg\}$ is a finite, directed, labeled tree satisfying the following:

1. All nodes of the tree are of at most single fanout and of arbitrary fanin.
2. Each of the leaves (inputs) of the tree are labeled by 0, 1, x or $\neg x$ for $x \in X$, while all other nodes are labeled by $\bigvee$ or $\bigwedge$.

Unless otherwise stated all circuits are over the boolean base. Let C be a circuit in X. Provided that C is non-empty, the unique node with no fanout is the *root* of C. C evaluates (in the obvious way) to 0 or 1, when all x, $x \in X$, are replaced by 0 or 1. Thus C defines a boolean function $f_C: \{0,1\}^X \longrightarrow \{0,1\}$. For $z \in \{0,1\}^X$ we write $C(z)$ for $f_C(z)$. Let $Y \subseteq \{0,1\}^X$. We then write $C|Y \equiv D|Y$ if $f_C|Y = f_D|Y$, for C, D circuits on X. If $\mathbf{C}$ and $\mathbf{D}$ are *indexed sets* of circuits on X (such objects are called *multicircuits*), we write $\mathbf{C}|Y \equiv \mathbf{D}|Y$ if $\{f_C|Y \mid C \in \mathbf{C}\} = \{f_D|Y \mid D \in \mathbf{D}\}$. If $Y = \{0,1\}^X$ we drop all reference to it in the previous notation.

1 f. Given d, s positive integers, we say that a circuit C:

i. *has depth* $d \geq 0$ if d is the number of edges in the longest path of C, and

ii. *has size* $\leq s$ if the number of nodes of C does not exceed s.

Given a circuit C the *d-children* of C are the subcircuits of C of depth d. If d is the depth of C then we call the $(d-1)$-children of C simply the children of C.

1 g. A *circuitry*, $C = \{C_\ell\}_0^\infty$, *has depth* $\leq d$ *and size* $\leq p$ if for each ℓ, C_ℓ is a circuit of depth $\leq d$ and size $\leq p(\ell)$.

1 h. Let μ be a positive integer and q a polynomial:

i. a *μ-short multicircuit of size* $\leq q$ is a sequence of *indexed sets*, one set to each ℓ, of $\leq q(\ell)$ circuits of depth 2 with all children having $\leq \mu$ inputs, and

ii. a *μ-finite multicircuit of size* $\leq q$ is a sequence of *sets*, one set to each ℓ, of $\leq q(\ell)$ circuits with $\leq \mu$ inputs.

Given $C = \{C_\ell\}_0^\infty$ a μ-short(finite) multicircuit of size q and C_ℓ a member of C, we shall again say that C_ℓ is a μ-short(finite) multicircuit of size $\leq q$.

1 i. Let λ be a positive integer and X the set of λ inputs. $\rho \in \{0,1,*\}^X$ will be called a *restriction on* X. We denote by free(ρ) the set $\rho^{-1}(*)$ and by unfree(ρ) the set $\rho^{-1}(0) \cup \rho^{-1}(1)$. We identify a restriction ρ on X with the set of $\sigma \in \{0,1\}^X$ such that $\rho(x) = \sigma(x)$, for $x \in$ unfree(ρ). Under this identification restrictions are partially ordered by $\subseteq$.

1 j. Let λ be a positive integer, X the set of λ inputs and $\epsilon \in (0,1]$. We shall denote by $\text{Restr}_\epsilon(X)$ the set of restrictions ρ on X such that $|\text{free}(\rho)| \geq \lambda^\epsilon$. We write $\text{Restr}_\epsilon(\lambda)$ or $\text{Restr}_\epsilon(\ell)$ whenever the set X is understood or whenever the inputs are identified with the set $[\lambda]$ or $[\ell]$.

1 k. We shall abuse the notation and use ρ to denote both the restriction on $X(\ell)$ of $\lambda(\ell)$ inputs and the (infinite) *sequence* of restrictions, one per $X(\ell)$.

1 l. If C is a circuit in X, the set of inputs, and ρ a restriction on X then by $C|\rho$ we understand the circuit that arises from C by relabeling in C every label $x(\neg x)$ by $\rho(x)(1-\rho(x))$ for $x \in$ unfree(ρ) $\subseteq X$, and pruning C in the obvious way. If $C = \{C_\ell\}_0^\infty$ is a circuitry and $\rho = \{\rho_\ell\}_0^\infty$ a sequence of restrictions on appropriate sets of inputs, then $C|\rho = \{C_\ell|\rho_\ell\}_0^\infty$. Whenever the context allows we shall drop the subscripts and write $C|\rho$ for $C_\ell|\rho_\ell$.

The fundamental lemma for bounded circuits

We first show how to convert a μ-short multicircuit of polynomial size to a ν-finite multicircuit of polynomial size.

Conversion Lemma 1.1. *Let μ be a positive integer and q a polynomial in one variable. Then,*
$\forall\, \epsilon \in (0,1]; \exists\, \nu$ *positive integer,* $\delta \in (0, \epsilon], \ell_0$;
$\forall$ *μ-short* $\mathbf{C}$ *multicircuit of size* q, $\ell \geq \ell_0$;
$\forall\, \rho \in \mathrm{Restr}_\epsilon(\ell)\ \exists\, \rho' \in \mathrm{Restr}_\delta(\ell), \rho' \subseteq \rho$

$$C_\ell|\rho' \equiv C'_\ell \text{ a } \nu\text{-finite multicircuit of size } q.$$

Proof. We proceed by induction on μ. To get the right restriction we use probability. Given λ we consider $\hat{\rho}$ a *probabilistic restriction*, i.e. a sequence of λ independent trials with outcomes $*, 1, 0$ and their respective probabilities of success p_*, p_1, and p_0:

i. $p_* = 1/\sqrt{\lambda}$, and

ii. $p_1 = p_0 = \frac{1}{2}(1 - \frac{1}{\sqrt{\lambda}})$.

Using the multinomial distribution with parameters λ, p_0, p_1, and p_*, we show that desired properties on a probabilistic restriction, $\hat{\rho}$, have non-zero probability. This renders the subset of the sample space with those desired properties, nonempty. So, pick such a restriction. $\mu = 1$. Given is a 1-short multicircuit of size q. With ϵ and $\rho \in \mathrm{Restr}_\epsilon(\ell)$ reset $C = C|\rho$, $X = \mathrm{free}(\rho)$ and $\lambda = |\mathrm{free}(\rho)|$. We say that $C \in C_\ell$ is *wide* if C has more than $c \cdot \log \lambda$ children, otherwise C is called *narrow*. Note that here the children are literals because $\mu = 1$. We now see that:

1. For wide C and $\lambda > 4$,

$$\begin{aligned} \mathbf{Prob}(C|\hat{\rho} \text{ is not constant}) &\leq [1 - \tfrac{1}{2}(1 - \tfrac{1}{\sqrt{\lambda}})]^{c \cdot \log \lambda} \\ &< (3/4)^{c \cdot \log \lambda} = \lambda^{-c \cdot \log(4/3)}. \end{aligned}$$

2. For narrow C, sufficiently large λ and γ = number of distinct inputs of C,

$$\mathbf{Prob}(C|\hat{\rho} \text{ has at least } \nu \text{ *-ed inputs}) \leq \sum_{j=\nu}^{\gamma} \binom{\gamma}{j} \cdot (\tfrac{1}{\sqrt{\lambda}})^j \cdot (1 - \tfrac{1}{\sqrt{\lambda}})^{\gamma - j}$$

$$< \binom{\gamma}{\nu} \cdot (\tfrac{1}{\sqrt{\lambda}})^{\nu} < \lambda^{-(\frac{\nu}{2}-1)}$$

because $\gamma < c \cdot \log \lambda$ and because

$$\binom{\gamma}{j+\nu} \leq \binom{\gamma}{\nu} \cdot \binom{\gamma-\nu}{j}.$$

3. By Chebychev's inequality applied to $\sigma^2 = \sqrt{\lambda} - 1$ and $\mu = \sqrt{\lambda}$, *where μ is the mean for this calculation and should not be confused with the μ of the induction,*

$$\mathbf{Prob}(|\text{free}(\hat{\rho})| < \frac{1}{2} \cdot \sqrt{\lambda}) < \frac{4}{\sqrt{\lambda}}.$$

The above are the probabilities that something goes wrong. So the probability of everything being right is

$$\alpha_\ell = 1 - \frac{4}{\sqrt{\lambda}} - q(\ell) \cdot (\lambda^{-c \cdot \log(4/3)} + \lambda^{-(\frac{\nu}{2}-1)}).$$

Choose c and ν so that $\alpha_\ell > 0$ for $\ell \geq \ell_0$ (possible, since q has finite degree and $\lambda \geq \ell^\epsilon$). Clearly now, we can choose $\rho' \subseteq \rho$ such that $\rho' \in \text{Restr}_{\epsilon/4}(\ell)$ and such that $C|\rho'$ is ν-finite. The size is at most q. Note that ν only depends on $\text{degr}(q)$ and ϵ.

$\mu > 1$. Again with q, ϵ and $\rho \in \text{Restr}_\epsilon(\ell)$ fixed we reset $C = C|\rho$, $X = \text{free}(\rho)$ and $\lambda = |\text{free}(\rho)|$ $(\geq \ell^\epsilon)$. We say $C \in C_\ell$ is *wide* if there are $\geq c \cdot \log \lambda$ children of C with mutually disjoint sets of inputs, otherwise we say C is *narrow*. For wide C and $\lambda > 4$ we have that

$$\begin{aligned} \mathbf{Prob}(C|\hat{\rho} \text{ is not constant}) &\leq \left[1 - [\tfrac{1}{2} \cdot (1 - \tfrac{1}{\sqrt{\lambda}})]^{\mu}\right]^{c \cdot \log \lambda} \\ &< [1 - (\tfrac{1}{4})^{\mu}]^{c \cdot \log \lambda} = \lambda^{-c \cdot \beta} \end{aligned}$$

where $\beta = \log[1 - (\frac{1}{4})^{\mu}]^{-1} > 0$. For narrow C we let $H(C)$ be a maximally disjoint family of sets of inputs chosen from the sets of inputs to the children of C. $|\bigcup H(C)| < \mu \cdot c \cdot \log \lambda$. We have that

$$\mathbf{Prob}(|\bigcup H(C) \cap \text{free}(\hat{\rho})| \geq \nu') \leq \binom{\lfloor \mu \cdot c \cdot \log \lambda \rfloor}{\nu'} \cdot (\tfrac{1}{\sqrt{\lambda}})^{\nu'} < \lambda^{-(\frac{\nu'}{2}-1)}$$

by an analogous calculation to that concerning narrow circuits in the base case of this induction. So again we get $\rho' \subseteq \rho$, $\rho' \in \text{Restr}_{\epsilon/4}(\ell)$ such that wide circuits are constants, thus with no inputs. To see what happened to the narrow circuits of C_ℓ, let C be narrow and let $h(C)$ be the set $*$-ed inputs in $\bigcup H(C)$. By the choice of ρ', $|h(C)| < \nu'$. For $\sigma \in \{0,1\}^{h(C)}$ let:

1. for $x \in h(C)$, $\tilde{\sigma}(x) = \begin{cases} x, & \text{if } \sigma(x) = 1; \\ \neg x, & \text{if } \sigma(x) = 0. \end{cases}$;
2. $\tilde{\sigma}(h(C)) = \bigwedge_{x \in h(C)} \tilde{\sigma}(x)$, and
3. $\rho' \cup \sigma$ be the finer restriction resulting from uniting the functions $\rho' | X \setminus h(C)$ and σ, which is possible since $\mathrm{dom}(\sigma) = h(C) \subseteq \mathrm{free}(\rho')$.

Clearly then,

$$C|\rho' \equiv \bigvee_{\sigma \in \{0,1\}^{h(C)}} [\tilde{\sigma}(h(C)) \wedge (C|\rho' \cup \sigma)].$$

We now collect the circuits $D = C|\rho' \cup \sigma$ for narrow $C \in C_\ell$ and $\sigma \in \{0,1\}^{h(C)}$, into a new multicircuit. This new multicircuit is $(\mu - 1)$-short because the circuits have depth ≤ 2, because $h(C)$ intersects the inputs of every child in C, and because all of $h(C)$ are fixed by $\rho' \cup \sigma$. It has size $\leq q' = 2^{\nu'} \cdot q$. Apply the induction hypothesis to $\mu - 1$ with $\epsilon/4$ and ρ' and get ν'', δ and $\rho'' \subseteq \rho'$, $\rho'' \in \mathrm{Restr}_\delta(\ell)$ so that $D|\rho'' \equiv D''$ a ν''-finite multicircuit of size $\leq q'$. Set $\nu = \nu' + \nu''$. $C|\rho'' \equiv C''$ a ν-finite multicircuit of size $\leq q$. □

Next, the fundamental lemma. The proof of this lemma uses the conversion lemma. Of course, if we want wider cylinders we would have to use Håstad's switching lemma [Hås86].

Fundamental Lemma 1.2. (Furst-Saxe-Sipser) *Let d be a positive integer, q a polynomial in one variable. Then,*
$\forall\, \epsilon \in (0,1]; \exists\, \nu$ *positive integer,* $\delta \in (0, \epsilon], \ell_0$;
$\forall$ **C** *circuitry of depth d and size q,* $\ell \geq \ell_0$;
$\forall\, \rho \in \mathrm{Restr}_\epsilon(\ell)\ \exists\, \rho' \in \mathrm{Restr}_\delta(\ell), \rho' \subseteq \rho$

$$C_\ell|\rho' \equiv C'_\ell \text{ a circuit of depth } d-1 \text{ and size} \leq 2^\nu \cdot q.$$

Proof. Collect all 1-children of C into a multicircuit, D. $D \equiv D'$ a 1-short multicircuit of size $\leq q$. Apply, therefore, the conversion lemma with ϵ fixed to get μ such that given $\rho \in \mathrm{Restr}_\epsilon(\ell)$ we can find $\rho' \subseteq \rho$, $\rho' \in \mathrm{Restr}_{\epsilon/4}(\ell)$ which renders $D'|\rho' \equiv D''$ a μ-finite multicircuit of size $\leq q$. Next, collect the 2-children of $C|\rho'$ into a multicircuit, E. $E \equiv E'$ a μ-short multicircuit of size $\leq q$ because the circuits of D'' are circuits of depth 1 and have at most μ inputs. Apply, therefore, the conversion lemma with μ q, $\epsilon/4$ and $\rho' \in \mathrm{Restr}_{\epsilon/4}(\ell)$ to get ν and ℓ_0 such that for any $\ell \geq \ell_0$ there is $\rho'' \subseteq \rho'$, $\rho'' \in \mathrm{Restr}_\delta(\ell)$, which renders

$E|\rho'' \equiv E''$ a ν-finite multicircuit of size $\leq q$. Clearly now, $C|\rho'' \equiv C''$ a circuit the 2-children of which are circuits of no more than ν inputs. Express each of them in disjunctive or conjunctive normal form getting C''', with $C''' \equiv C''$. Evidently C''' has depth $d-1$, size $\leq 2^{\nu} \cdot q$ and $C|\rho'' \equiv C'''$. $\square$

It is a trivial corrolary of the fundamental lemma, that **PARITY** cannot be calculated by bounded circuitry of polynomial size.

Chapter 2

PARITY is not $\mathcal{SB}_1^N$ definable

In [Bus86] it was quickly mentioned that sharply bounded formulas in one variable (i.e. sets) using atomic relations *not* involving *multiplication*, could not define **PARITY**. However, here it is shown that ,

Theorem. **PARITY** *is not sharply bounded definable.*

Of course, by sharply bounded we mean in the *full* language of bounded arithmetic.

The proof of the restricted result in [Bus86] uses the weak lower bounds for **PARITY** found in [Ajt83] and [FSS84]. These bounds as well as the optimal bound found in [Hås86] are calculated for *nonuniform* circuits. Note that by a circuit we mean a *sequence* of boolean circuits, one for each binary length.(For more see [FSS84])

In [Bus86] the resulting circuits are not only of constant depth and polynomial size but also *uniform*, in the sense that there is *one* formula which prescribes the circuit at each length, a formula which involves particularly simple algebraic operations.

Here, the circuits result from formulas involving multiplication, which is not of constant depth. Their uniformity, however, forces them to be of constant depth and more than that, these circuits are of a particularly simple nature, as can be seen from the sharp truth lemma.

Notation and Conventions

The following are to be noted:

2a. $\mathbf{PARITY}(x) = \begin{cases} 1, & \text{if the binary expansion of } x \text{ has odd \# of 1's;} \\ 0, & \text{otherwise.} \end{cases}$

2b. Let $A(x,\ldots)$ be a property of x, possibly with parameters. Given a set P, we say A is *oblivious* in P if,

$$\forall x, y \in P \ [A(x,\ldots) \leftrightarrow A(y,\ldots)] \ .$$

2c. Throughout, by a sharply bounded formula in x, say $\Theta(x)$, we shall mean a formula of the form,

$$Q_1 \, \vec{z}^{(1)} < |x|^m \ldots Q_n \, \vec{z}^{(n)} < |x|^m \, \Psi(x, \vec{z}) \ ,$$

where, $Q_i = \forall$ or $\exists$, for $i = 1,\ldots,n$ in an alternating fashion, Ψ is a boolean combination of atomic statements in the language of bounded arithmetic and $\vec{z} = \vec{z}^{(1)} \cup \cdots \cup \vec{z}^{(n)}$. An arbitrary sharply bounded formula is logically equivalent to a formula of the above form.

2d. Let $t(x, \vec{z})$ be an arbitrary but fixed term of bounded arithmetic, where $\vec{z} = z_1, \ldots, z_k$ are k distinct variables different from x. Recall that the language of bounded arithmetic is the same as the language of Peano arithmetic enhanced by *three* new function symbols:

1. $\lfloor x/2 \rfloor$, the *integer part* of $x/2$,
2. $|\, x \,|$, the *length of the binary expansion* of x, and
3. $x \# y$, interpreted as $2^{|x| \cdot |y|}$.

2e. Let f be a function $f\colon A \longrightarrow B$ and $C \subseteq A$. We then set,

$$f(C) = \{f(x) \mid x \in C\} \ ,$$

e.g. $f(x) = \lfloor x/2 \rfloor$, and then $\lfloor C/2 \rfloor = \{\lfloor x/2 \rfloor \mid x \in C\}$.

2f. $[2^\ell, 2^{\ell+1}) \bigcap \mathbf{N} = \left\{x \in \mathbf{N} | 2^\ell \leq x < 2^{\ell+1}\right\}$.

2g. $\nu(t)$ = maximum depth of nested occurrences of $\lfloor \_/2 \rfloor$ in $t(x, \vec{z})$ which contain occurrences of x.

2h. We say $|s|$ *maximally occurs* in t, if $|s|$ is a subterm of t or $s\#r$ is a subterm of t, for some other term r, *and* both $|s|$ and $s\#r$ can only occur in t within the scope of $\times$, $+$ and $\lfloor\, \_/2 \,\rfloor$. Now, set $\mathcal{M}(t) = \{s : |s| \text{ occurs maximally in } t\}$.

2i. With $\sigma < 2^\nu$, ν a positive integer we let $[\sigma]_\nu = \{x \in \mathbf{N} |\ x \equiv \sigma \pmod{2^\nu}\}$.

2j. By $\vec{z} < x$ we mean that all members of the sequence $\vec{z}$ are less than x.

2k. Let P and Q be two partitions of a set A; by *the superimposition of* or *superimposing* the two partitions P and Q, we mean the creation of a new partition of A, namely $P \bigwedge Q = \{X \bigcap Y | X \in P \text{ and } Y \in Q\}$. If P and Q happen to be collections of intervals then an easy argument shows that: $|P \bigwedge Q| \leq |P| + |Q|$.

The lemmata

This lemma states that the length of a term is constant in each of *finitely* many intervals,

Length Lemma 2.1. *Let m be a fixed positive, nonzero integer and $t(x, \vec{z})$ a term as above. There are positive integers ℓ_t and κ_t such that,*
$\forall \ell \geq \ell_t$, $\forall \vec{c} < \ell^m$,
$\exists \mathcal{P}$, *a partition of* $[2^\ell, 2^{\ell+1}) \bigcap \mathbf{N}$, *into at most κ_t many intervals,*
$\forall P \in \mathcal{P}$, $|t(x, \vec{c})|$ *is constant in P.*

Proof. First observe that for terms r and s:

(1) $a \in [2^\ell, 2^{\ell+1}) \bigcap \mathbf{N} \Rightarrow |a| = \ell + 1$,

(2) $|r + s| = \epsilon_{\max}(r, s) \cdot |r| + (1 - \epsilon_{\max}(r, s)) \cdot |s| + \epsilon_+(r, s)$,

(3) $|r \cdot s| = |r| + |s| - \epsilon_\times(r, s)$,

(4) $|r\#s| = |r| \cdot |s| + 1$, and

(5) $|\lfloor r/2 \rfloor| = |r| - 1$,

where $\epsilon_{\max}$, $\epsilon_\times$ and ϵ_+ are 0 or 1, depending on their arguments.
Fix $\vec{c}$, $\vec{c} < \ell^m$. Apply (2), (3), (4) and (5) above to $|t(x, \vec{c})|$. We get a polynomial expression in $|x|$ with coefficients in the lengths of the binary expansions of members of $\vec{c}$, and in $\vec{\epsilon}$.

Let $\hat{t}(|x|, \vec{c}, \vec{\epsilon})$ be that expression. Note that the length λ of the vector $\vec{\epsilon}$ only depends on the *syntactic structure* of the term t.
We have:

$(1)'$ $\forall a \in [2^\ell, 2^{\ell+1}) \bigcap \mathbf{N}$, $\exists \vec{\epsilon} < 2^\lambda$ $|t(a, \vec{c})| = \hat{t}(\ell + 1, \vec{c}, \vec{\epsilon})$, by (1), and

$(2)'$ $|t(x, \vec{c})|$ is increasing in x, because $t(x, \vec{c})$ is increasing in x.

We conclude by $(1)'$ that there are 2^λ possible values for $|t(a, \vec{c})|$, $a \in [2^\ell, 2^{\ell+1}) \bigcap \mathbf{N}$. By $(2)'$ now, the inverse of each point in the range of $|t(x, \vec{c})|$ in $[2^\ell, 2^{\ell+1}) \bigcap \mathbf{N}$ must be an interval. Set $\kappa_t = 2^\lambda$.□

This is the first of the "oblivious" lemmata and states that atomic statements are oblivious in each of polynomially many intervals,

Atomic Truth Lemma 2.2. *Let $t(x, \vec{z})$ and $s(x, \vec{z})$ be two arbitrary but fixed terms of bounded arithmetic. Let k be the length of the vector $\vec{z}$. Let n be a fixed non-zero positive integer and $m = n \cdot k$.*
With $\nu \geq \max\{\nu(t), \nu(s)\}$ there are positive integers $\kappa_{t,s}$ and $\ell_{t,s}$ such that,
$\forall \sigma < 2^\nu$, $\forall \ell \geq \ell_{t,s}$,
$\exists \mathcal{P}_\sigma$, *a partition of* $[2^\ell, 2^{\ell+1}) \bigcap [\sigma]_\nu$ *into at most* $\kappa_{t,s} \cdot \ell^m$ *many intervals,*
$\forall P \in \mathcal{P}_\sigma, \forall \vec{c} < \ell^m$, $s(x, \vec{c}) < t(x, \vec{c})$, $s(x, \vec{c}) = t(x, \vec{c})$ *and* $s(x, \vec{c}) > t(x, \vec{c})$ *are oblivious in* P.

Proof. Fix σ, $\sigma < 2^\nu$. Until further notice, fix $\vec{c}$, $\vec{c} < \ell^m$. Apply the length lemma to each $r \in \mathcal{K}(t(x, \vec{c})) \bigcup \mathcal{K}(s(x, \vec{c}))$ to find a partition of $[2^\ell, 2^{\ell+1}) \bigcap \mathbf{N}$ into not more than κ_r intervals, in each of which r is constant. Superimpose these partitions and let $\mathcal{Q}_{\vec{c}}$ be the resulting partition of $[2^\ell, 2^{\ell+1}) \bigcap \mathbf{N}$. Let $\mathcal{Q}_{\vec{c},\sigma} = \{Q' \bigcap [\sigma]_\nu \mid Q' \in \mathcal{Q}_{\vec{c}}\}$ and until further notice, fix Q, $Q \in \mathcal{Q}_{\vec{c},\sigma}$. Notice now, that with,

$$\kappa = \sum \left\{ \kappa_r \mid r \in \mathcal{K}(t(x, \vec{c})) \bigcup \mathcal{K}(s(x, \vec{c})) \right\} ,$$

we have that $|\mathcal{Q}_{\vec{c}}| \leq \kappa$ and hence also $|\mathcal{Q}_{\vec{c},\sigma}| \leq \kappa$. Again, κ *only* depends on the *syntactic structure of* the terms s and t. Since $Q = Q' \bigcap [\sigma]_\nu$, for some $Q' \in \mathcal{Q}_{\vec{c}}$, we have that $|r(x, \vec{c})|$ is constant in Q, for any $r \in \mathcal{K}(t(x, \vec{c})) \bigcup \mathcal{K}(s(x, \vec{c}))$. Next, replace in $t(x, \vec{c})$ and $s(x, \vec{c})$ all subterms by their (*unique*) value achieved in Q' (and hence in Q) and call the

resulting expression in x, $t_{Q,\vec{c}}(x)$ and $s_{Q,\vec{c}}(x)$ respectively. Obviously, $t(a,\vec{c}) = t_{Q,\vec{c}}(a)$ and $s(a,\vec{c}) = s_{Q,\vec{c}}(a)$, for all $a \in Q$. We get the polynomials, $t_{Q,\vec{c},\sigma}(x)$ and $s_{Q,\vec{c},\sigma}(x)$ as follows: Observe first that $t_{Q,\vec{c}}(x)$ only involves multiplication, addition and $\lfloor\, \_/2 \,\rfloor$, unless it has become a constant. Define the following operation on terms of the above type,

$$t^{(\epsilon)}(x) = t(2x + \epsilon) \quad \text{where } \epsilon = 1 \text{ or } 0,$$

where it is understood that $\lfloor (2 \cdot s + r)/2 \rfloor$ is replaced by $s + \lfloor r/2 \rfloor$ and if σ is a binary sequence of length $\nu > 0$ then,

$$t^{(\sigma)}(x) = t^{(\sigma(0))\cdots(\sigma(\nu-1))}(x).$$

Claim 1: *For $\epsilon = 0$ or 1 we have that $\nu(t^{(\epsilon)}(x)) < \nu(t(x))$ and if $\nu(t(x)) > 0$, and $\nu(t^{(\epsilon)}(x)) = 0$ if $\nu(t(x)) = 0$.*

Proof of claim 1. The second statement of the claim is obvious, and the first is proved by induction on $\nu(t(x))$. It suffices to show that if $\nu(t(x)) = 1$ then $\nu(t^{(\epsilon)}(x)) = 0$; but this is obvious too.□

We let $t_{Q,\vec{c},\sigma}(x) = t^{(\sigma)}_{Q,\vec{c}}(x)$. The length of σ is ν, so that applying claim 1, ν times we see that $\nu(t_{Q,\vec{c},\sigma}(x)) = 0$, i.e. $t_{Q,\vec{c},\sigma}(x)$ is a polynomial expression in x with non-negative integer coefficients.

Claim 2: *$t(a,\vec{c}) = t_{Q,\vec{c},\sigma}(\lfloor a/2^\nu \rfloor)$, for $a \in Q$.*

Proof of claim 2. First observe that:

(1) $t(x) = t(2\lfloor x/2 \rfloor + x(0)) = t^{(x(0))}(\lfloor x/2 \rfloor)$,

(2) $\lfloor \lfloor x/2^\nu \rfloor /2 \rfloor = \lfloor x/2^{\nu+1} \rfloor$, and

(3) $a \in Q \Rightarrow a \equiv \sigma \pmod{2^\nu}$ and $t(a,\vec{c}) = t_{Q,\vec{c}}(a)$.

So for $a \in Q$,

$$\begin{aligned} t_{Q,\vec{c},\sigma}(\lfloor a/2^\nu \rfloor) &= t^{(\sigma)}_{Q,\vec{c}}(\lfloor a/2^\nu \rfloor), \text{ by definition;} \\ &= t_{Q,\vec{c}}(a), \text{by (1) and (2);} \\ &= t(a,\vec{c}), \text{ by (3).} \end{aligned}$$

□

Similarly for $s_{Q,\vec{c},\sigma}(x)$. Consider now, $h(x) = t_{Q,\vec{c},\sigma}(x) - s_{Q,\vec{c},\sigma}(x)$. Now, h is a polynomial in x with integer coefficients; let d be the degree of h. Therefore, there are at most $2d+1$ subintervals of $[2^{\ell-\nu}, 2^{\ell-\nu+1}) \bigcap \mathbf{N}$ in which h is either all zero or all negative or all positive. As a set, $Q = 2^{\nu} \cdot \lfloor Q/2^{\nu} \rfloor + \sigma$, and $\lfloor Q/2^{\nu} \rfloor$ is a subinterval of $[2^{\ell-\nu}, 2^{\ell-\nu+1}) \bigcap \mathbf{N}$; so $\lfloor Q/2^{\nu} \rfloor$ breaks up into at most $2d+1$ intervals, in each of which h is either all zero or all positive or all negative.

Recalling that for $a \in Q$, $h(\lfloor a/2^{\nu} \rfloor) = t(a,\vec{c}) - s(a,\vec{c})$, we conclude that Q splits in at most $2d+1$ intervals in each of which $t(a,\vec{c}) - s(a,\vec{c})$ is either all zero or all positive or all negative. For each Q, $Q \in \mathcal{Q}_{\vec{c},\sigma}$, collect all of their new pieces and form a new partition, $\mathcal{P}_{\vec{c},\sigma}$.

Note that $d \leq d_{\max}$ = maximum degree of x in t and s. and so for each $\vec{c}$, $\vec{c} < \ell^m$, we have that $|\mathcal{P}_{\vec{c},\sigma}| \leq (2d_{\max}+1) \cdot \kappa$.

Let $\mathcal{P}_{\sigma}$ be the superimposition of all the $\mathcal{P}_{\vec{c},\sigma}$, $\vec{c} < \ell^m$.

Clearly now:

(1) $|\mathcal{P}_{\sigma}| \leq \kappa_{s,t} \cdot \ell^m$, where $\kappa_{s,t} = (2d_{\max}+1) \cdot \kappa$, and

(2) for any $\vec{c}$, $\vec{c} < \ell^m$, and any P, $P \in \mathcal{P}_{\sigma}$, $A(x,\vec{c})$ is oblivious in P,
where $A(x,\vec{c})$ is either $t(x,\vec{c}) < s(x,\vec{c})$ or $t(x,\vec{c}) = s(x,\vec{c})$
or $t(x,\vec{c}) > s(x,\vec{c})$.

□

The next lemma tells us that a boolean combination of atomic wff's is oblivious in each of polynomially many intervals,

Boolean Truth Lemma 2.3. *Let $\Psi(x,\vec{z})$ be a boolean combination of atomic statements in bounded arithmetic in x and $\vec{z}$, n a positive integer and k the length of the vector $\vec{z}$. With $\nu \geq \max\{\nu(t) | t$ a term in $\Psi\}$ and $m = n \cdot k$, there are positive integers κ_{Ψ} and ℓ_{Ψ} such that,*

$\forall \sigma < 2^{\nu}$, $\forall \ell \geq \ell_{\Psi}$,

$\exists \mathcal{P}_{\sigma}$ *a partition of* $[2^{\ell}, 2^{\ell+1}) \bigcap [\sigma]_{\nu}$ *into at most* $\kappa_{\Psi} \cdot \ell^m$ *many intervals,*

$\forall P \in \mathcal{P}_{\sigma}$, $\forall \vec{c} < \ell^m$, $\Psi(x,\vec{c})$ *is oblivious in* P.

Proof. Fix σ, $\sigma < 2^\nu$. We apply the atomic truth lemma to every positive atomic subformula of $\Psi(x, \vec{z})$. For each such subformula, say $A(x, \vec{z})$, we obtain partition $\mathcal{P}_{\sigma,A}$ such that:

(1) $|\mathcal{P}_{\sigma,A}| \leq \kappa_A \cdot \ell^m$, and

(2) for any $\vec{c}$, $\vec{c} < \ell^m$, $A(x, \vec{c})$ is oblivious in any P, $P \in \mathcal{P}_{\sigma,A}$.

Let $\mathcal{P}_\sigma$ be the superimposition of all the $\mathcal{P}_{\sigma,A}$, for A a positive atomic in Ψ, and,

$$\kappa_\Psi = \sum \{\kappa_A | A \text{ a positive atomic in } \Psi\} .$$

Then clearly:

(1) $|\mathcal{P}_\sigma| \leq \kappa_\Psi \cdot \ell^m$, and

(2) for any $\vec{c}$, $\vec{c} < \ell^m$, $\Psi(x, \vec{c})$ is oblivious in any P, $P \in \mathcal{P}_\sigma$.

□

This last lemma tells us that sharply bounded predicates are oblivious in each of polynomially many intervals,

Sharp Truth Lemma 2.4. *Let $\Theta(x)$ be a sharply bounded formula in x. With $\nu \geq \max\{\nu(t) | t \text{ is a term in } \Theta\}$, there are positive integers m_Θ, κ_Θ and ℓ_Θ such that,*
$\forall \sigma < 2^\nu$, $\forall \ell \geq \ell_\Theta$,
$\exists \mathcal{P}_\sigma$ a partition of $[2^\ell, 2^{\ell+1}) \bigcap [\sigma]_\nu$ into at most $\kappa_\Theta \cdot \ell^{m_\Theta}$ many intervals,
$\forall P \in \mathcal{P}_\sigma$, $\Theta(x)$ is oblivious in P.

Proof. Fix σ, $\sigma < 2^\nu$. Let n_Θ be the power in the sharp bounds of the quantifiers of Θ and k the maximum length of the vectors appearing in the sharp quatifiers of Θ. Let $\Psi(x, \vec{z})$ be the quantifier-free matrix of Θ. Apply the boolean truth lemma to Ψ with $n = n_\Theta$. Put $\kappa_\Theta = \kappa_\Psi$ and consider $\mathcal{P}_\sigma$, the partition obtained for Ψ. Then:

(1) $|\mathcal{P}_\sigma| \leq \kappa_\Theta \cdot \ell^{m_\Theta}$ where $m_\Theta = k \cdot n_\Theta$, and

(2) $\Theta(x)$ is oblivious in any P, $P \in \mathcal{P}_\sigma$.

□

Proof of the theorem

Note that for any integer x,

$$\mathbf{PARITY}(2 \cdot x) = 1 - \mathbf{PARITY}(2 \cdot x + 1)\,. \qquad (\star)$$

Let $\Theta(x)$ be a sharply bounded formula in x. Fix $\sigma < 2^{\nu}$, apply the sharp truth lemma to $\Theta(x)$ and get the partition $\mathcal{P}_\sigma$. Since $|\mathcal{P}_\sigma| \leq \kappa_\Theta \cdot \ell^{m_\Theta}$, we can find $P \in \mathcal{P}_\sigma$ with $|P| \geq 2^{\ell-\nu}/(\kappa_\Theta \cdot \ell^{m_\Theta})$. Consider $Q = \lfloor P/2^{\nu} \rfloor$, which is a subinterval of $[2^{\ell-\nu}, 2^{\ell-\nu+1}) \bigcap \mathbf{N}$. Let $Q' = [a, b]$, with a =least even number in Q, and b =largest odd number in Q. Then $|Q| \geq |Q'| \geq (|Q| - 2)$, and $Q' \subseteq Q$. It is now an easy calculation using $(\star)$ to show that **PARITY** is satisfied by exactly half of Q''s elements and hence **PARITY** is satisfied by exactly half of P''s elements, where $P' = 2^{\nu} \cdot Q' + \sigma$. $|P'| \geq 2^{\ell-\nu-1}/\kappa_\Theta \cdot \ell^{m_\Theta}$. But Θ is oblivious in P, and $P' \subseteq P$. It follows that Θ is oblivious in P', and hence cannot be **PARITY**. ■

Chapter 3

The set of primes is not $\mathcal{SB}_1^{\mathbf{N}}$ definable

Gaisi Takeuti asked whether the set of prime numbers is sharply bounded definable. Here it is shown that

Theorem. *The set of primes is not sharply bounded definable.*

The proof uses the prime number theorem, namely that: $\pi(x) \approx x/\ln x$. The rest depends on the fact that the number of "graded" intervals in which a sharply bounded formula of one free variable is oblivious, is no more than ℓ^m, where ℓ is the length of the current arguments to the formula, and m depends entirely on the syntax of the formula.

Proof of the theorem

Let $\Theta(x)$ be a sharply bounded formula in x. Recall the statement of the sharp truth lemma, and apply it to $\Theta(x)$ to get that with

$$\nu \geq \max\{\nu(t) | t \text{ a term in } \Theta\}$$

there are positive integers m_Θ, κ_Θ, and ℓ_Θ such that,
$\forall \sigma < 2^\nu$, $\forall \ell \geq \ell_\Theta$;
$\exists \mathcal{P}_\sigma$ a partition of $[2^\ell, 2^{\ell+1}) \bigcap [\sigma]_\nu$ into at most $\kappa_\Theta \cdot \ell^{m_\Theta}$ many intervals;
$\forall P \in \mathcal{P}_\sigma$,

$\Theta(x)$ is oblivious in P.

For all large ℓ the interval $[2^\ell, 2^{\ell+1}) \bigcap \mathbf{N}$ contains at least $2^\ell/(2 \cdot \ell)$ primes: we prove this with the help of the prime number theorem which states that,

$\forall \epsilon > 0 \exists \ell_\epsilon \forall \ell \geq \ell_\epsilon$,

$$-\epsilon < \frac{\pi(2^\ell) \cdot \ell}{2^\ell} - \frac{1}{\ln 2} < \epsilon$$

or equivalently,

$$\left(\frac{1}{\ln 2} - \epsilon\right) \cdot \frac{2^\ell}{\ell} < \pi(2^\ell) < \left(\frac{1}{\ln 2} + \epsilon\right) \cdot \frac{2^\ell}{\ell},$$

with $\pi(x) =$ number of primes $< x$. In general if A is a set of numbers we put $\pi(A) =$ number of primes $\in A$.

We work with fixed ϵ, to be determined later. Consider $\ell \geq \ell_\epsilon$. Then,

$$\left(\frac{1}{\ln 2} - \epsilon\right) \cdot \frac{2^\ell}{\ell} < \pi(2^\ell) < \left(\frac{1}{\ln 2} + \epsilon\right) \cdot \frac{2^\ell}{\ell}$$

and

$$\left(\frac{1}{\ln 2} - \epsilon\right) \cdot \frac{2^{\ell+1}}{\ell+1} < \pi(2^{\ell+1}) < \left(\frac{1}{\ln 2} + \epsilon\right) \cdot \frac{2^{\ell+1}}{\ell+1}.$$

Now,

$$\begin{aligned} \pi([2^\ell, 2^{\ell+1})) &= \pi(2^{\ell+1}) - \pi(2^\ell) \\ &> \left(\frac{1}{\ln 2} - \epsilon\right) \cdot \frac{2^{\ell+1}}{\ell+1} - \left(\frac{1}{\ln 2} + \epsilon\right) \cdot \frac{2^\ell}{\ell} \\ &= \frac{2^\ell}{\ell} \cdot \left[2 \cdot \frac{\ell}{\ell+1} \cdot \left(\frac{1}{\ln 2} - \epsilon\right) - \left(\frac{1}{\ln 2} + \epsilon\right)\right]. \end{aligned}$$

$1.4 < \frac{1}{\ln 2} < 1.5$. For $\epsilon < 0.1$ and $\ell > 5$ we have that,

$$\begin{aligned} 2 \cdot \frac{\ell}{\ell+1} \cdot \left(\frac{1}{\ln 2} - \epsilon\right) - \left(\frac{1}{\ln 2} + \epsilon\right) &> 2 \cdot \frac{\ell}{\ell+1} \cdot 1.3 - 1.6 \\ &= 2.6 \cdot \left(1 - \frac{1}{\ell+1}\right) - 1.6 \\ &= 1 - \frac{2.6}{\ell+1} > \frac{1}{2}. \end{aligned}$$

This finishes the first stage of the proof.

Pick an $\ell \geq \ell_\Theta$ with the previous property. Then there are $\lambda \geq 2^{\ell-\nu-1}/(\kappa_\Theta \cdot \ell^{m_\Theta+1})$ primes in some interval P, $P \in \mathcal{P}_\sigma$ some $\sigma < 2^\nu$. We claim that if three numbers in P are

consecutive then one of them is divisible by 3: if $x \in P$ then $x \equiv \sigma \bmod 2^\nu$; take n_i, n_{i+1}, and n_{i+2} consecutive in P. We then have that,

$$\begin{aligned} n_i &= f \cdot 2^\nu + \sigma\,; \\ n_{i+1} &= (f+1) \cdot 2^\nu + \sigma = n_i + 2^\nu\,; \\ n_{i+2} &= (f+2) \cdot 2^\nu + \sigma = n_i + 2^{\nu+1}\,; \end{aligned}$$

let $\epsilon_\nu = (2^\nu \bmod 3)$; observe that,

$$\epsilon_\nu = \begin{cases} 1, & \text{if } \nu \text{ is even;} \\ 2, & \text{if } \nu \text{ is odd.} \end{cases}$$

let $\delta = (n_i \bmod 3)$; it follows now that,

$$\begin{aligned} n_i &\equiv \delta \bmod 3\,; \\ n_{i+1} &\equiv \delta + \epsilon_\nu \bmod 3\,; \\ n_{i+2} &\equiv \delta + \epsilon_{\nu+1} \bmod 3\,; \end{aligned}$$

δ, $\delta+\epsilon_\nu$, and $\delta+\epsilon_{\nu+1}$ are *distinct* mod 3, because ϵ_ν and $\epsilon_{\nu+1}$ are distinct, < 3, and nonzero; but then one of δ, $\delta + \epsilon_\nu$, $\delta + \epsilon_{\nu+1}$ is 0 mod 3, which in turn implies that one of n_i, n_{i+1}, n_{i+2} is divisible by 3.

We conclude, thus, that P contains composites as well as the λ primes. This then implies that since Θ is oblivious in P, Θ cannot possibly define the set of primes. ■

Chapter 4

PARITY is not $\mathcal{SB}_1^{\mathbf{N}}(\mathbf{BIT})$ definable

The main result in this chapter is ,

Michael's Theorem. **PARITY** *is not bit sharply bounded definable.*

Now, for some comments on the issue: in [FSS84] it is quickly mentioned that **PARITY** is reducible to bit-extraction with multiplication. (Although, in circuit theory bit-extraction is never mentioned because it is "build-in".) However, in [FSS84] the reduction is to $\ell \cdot \log_2 \ell$, where ℓ is the length of the current binary expansion. In [Weg87] it is mentioned that the reduction to a *linear* multiple of the length is still not proved. Here it is established that under *uniform* conditions the location at which the extra $\ell \cdot \log_2 \ell$ bits are needed to define **PARITY** is *within* the performance of multiplication and *not after*: we *first* need to "expand" x, *then* apply multiplication and finally bit-extraction. It is the *limited iteration* involved in the "expansion" that bit sharply bounded predicates cannot handle:

$$expansion(x) = \sum_{j=0}^{\ell} x(j) \cdot 2^{j \cdot \lceil \log_2 \ell \rceil}.$$

The key lemma here is the *wide lemma*, which gives us a cylinder in which multiplication as a function and multiplication as a relation reduces to a polynomial size constant depth circuit.

It it worth mentioning, however, that at this point in time [1] it is not known whether the

[1] Conference and workshop on Proof Theory, Arithmetic, and Complexity in U.C.S.D, 28 June 1990

graph of multiplication is a polynomial size, constant depth circuit in general, i.e. without any restriction on the domains of input.

Notation and Conventions

In addition to the notation and conventions of the previous part, the following are to be noted:

4a. When it is clear from the context that X is a set, $|X|$ will denote the cardinality of the *set* X.

4b. We enhance the language of bounded arithmetic by *one* binary predicate symbol, $\mathbf{BIT}(x, y)$. The interpretation of $\mathbf{BIT}(x, y)$ is:

$$\begin{aligned} \mathbf{BIT}(x,y) \text{ is true} &\iff x \bmod 2^y > 2^{y-1} \\ &\iff \text{"the } y\text{-th digit in the binary} \\ &\qquad \text{expansion of } x \text{ is 1"} . \end{aligned}$$

The new class of sharply bounded well formed formulae thus arising, will be called *bit sharply bounded formulae.*

4c. Let $t(x, \vec{z})$ and $s(x, \vec{z})$ be terms of bounded arithmetic. We say that, $\mathbf{BIT}(t(x, \vec{z}), s(x, \vec{z}))$ is a *bit-extractor* or just an *extractor* with *argument* $t(x, \vec{z})$ and *bit* $s(x, \vec{z})$.

4d. For integers x and i we let $x(i)$ denote the i-th digit in the binary expansion of x. Now let x and y be two positive integers. We say x *does not interfere with* y or *there is no interference between the blocks* x *and* y just in case,

$$(x + y)(i) = x(i) \vee y(i),$$

for all i.

4e. Let $A \subseteq [\lambda]$, λ a positive integer, and and $x < 2^\lambda$. We say, $x \subseteq A$ if $x = \sum_{a \in A} x(a) \cdot 2^a$. The set of x such that $x \subseteq A$ is denoted by 2^A and is called *a cylinder (with base A).*

4f. With $k \geq 1$, let $j_1, \ldots, j_k$, $j_1 \geq \cdots \geq j_k$, be nonzero positive integers, and $x_1, \ldots, x_k$ be variables. The *multiplicity group* of $j_1 \cdot x_1 + \cdots + j_k \cdot x_k$ is the subgroup $G_{j_1,\ldots,j_k}$

of the symmetric group S_k which fixes the expression $j_1 \cdot x_1 + \cdots + j_k \cdot x_k$, under the action:

$$\pi(j_1 \cdot x_1 + \cdots + j_k \cdot x_k) = j_1 \cdot x_{\pi(1)} + \cdots + j_k \cdot x_{\pi(k)} \text{ for } \pi \in G_{j_1,\ldots,j_k}.$$

4g. The *multiplicity number*, $\mu(j_1, \ldots, j_k)$, of $j_1 \cdot x_1 + \cdots + j_k \cdot x_k$ is $|G_{j_1,\ldots,j_k}|$.

4h. Given $a_1, \ldots, a_k$ a vector of pairwise distinct positive integers (or anything pairwise distinct for that matter), by the *multiplicity class* of $a_1, \ldots, a_k$ *with respect to* $j_1 \cdot x_1 + \cdots + j_k \cdot x_k$, we mean the orbit of $a_1, \ldots, a_k$ by $G_{j_1,\ldots,j_k}$ under the action:

$$\pi(a_1, \ldots, a_k) = a_{\pi(1)}, \ldots, a_{\pi(k)} .$$

Clearly the cardinality of the multiplicity class is $\mu(j_1, \ldots, j_k)$. Also note that if $a_1, \ldots, a_k$ and $b_1, \ldots, b_k$ are two vectors in each of which the entries are pairwise distinct and they belong to the same multiplicity class with respect to $j_1 \cdot x_1 + \cdots + j_k \cdot x_k$, then $j_1 \cdot a_1 + \cdots + j_k \cdot a_k = j_1 \cdot b_1 + \cdots + j_k \cdot b_k$.

4i. Let X be a set of pairwise distinct variables, $\tau = \sum_{x \in X} \tau(x) \cdot 2^x$ and j a positive nonzero integer. It is not hard to see that,

$$\tau^j = \sum_{k=1}^{j} \overset{\bullet}{\sum_{\substack{j_1 \geq \cdots \geq j_k \\ j_1 + \cdots + j_k = j \\ x_1, \ldots, x_k \in X \\ \text{pairwise distinct}}}} \mu(j_1, \ldots, j_k) \cdot \tau(x_1) \cdots \tau(x_k) \cdot \frac{j!}{j_1! \cdots j_k!} \cdot 2^{j_1 \cdot x_1 + \cdots + j_k \cdot x_k}$$

where the $\bullet$ over the second summation symbol signifies that we are summing over *one* representative of the multiplicity classes with respect to $j_1 \cdot x_1 + \cdots + j_k \cdot x_k$. The factor $\mu(j_1, \ldots, j_k)$ compensates for the missing summands.

4j. Let j be a nonzero positive integer. If $0 < k \leq j$ we let,

$$H(j, k : x_1, \ldots, x_k) = \{j_1 \cdot x_1 + \cdots + j_{k'} \cdot x_{k'} \mid j_1 \geq \cdots \geq j_{k'} > 0, \\ j_1 + \cdots + j_{k'} = j \text{ and } 0 < k' \leq k\};$$

otherwise we let,

$$H(j, k : x_1, \ldots, x_k) = H(j, j : x_1, \ldots, x_j).$$

Finally let $H(j : x_1, \ldots, x_j) = H(j, j : x_1, \ldots, x_j)$.

4k. For arbitrary positive nonzero integers d and k let

$$\mathcal{H}(d,k:x_1,\ldots,x_k)=\bigcup_{j=1}^{d}H(j,k:x_1,\ldots,x_k)$$

and

$$\mathcal{H}(d:x_1,\ldots,x_d)=\bigcup_{k=1}^{d}\mathcal{H}(d,k:x_1,\ldots,x_k).$$

4l. Let A be a set of positive integers with $|A|\geq d$. We say A is *wide* for d, if whenever $a_1,\ldots,a_\mu$ and $b_1,\ldots,b_\nu$ are vectors each of which has pairwise distinct entries from A and $j_1\cdot x_1+\cdots+j_\mu\cdot x_\mu$, $i_1\cdot x_1+\cdots+i_\nu\cdot x_\nu\in\mathcal{H}(d:x_1,\ldots,x_d)$,

$$j_1\cdot a_1+\cdots+j_\mu\cdot a_\mu\neq i_1\cdot b_1+\cdots+i_\nu\cdot b_\nu,$$

unless $\mu=\nu$, $j_l=i_l$ for $l=1,\ldots,\mu$, and $a_1,\ldots,a_\mu$ and $b_1,\ldots,b_\nu$ belong to the same multiplicity class with respect to $j_1\cdot x_1+\cdots+j_\mu\cdot x_\mu$. If $|A|=k<d$ then the same definition holds except that the length of the vectors are adjusted to k and $\mathcal{H}(d:\ldots)$ is replaced by $\mathcal{H}(d,k:x_1,\ldots,x_k)$.

The lemmata

This is the central lemma which renders parts of the graph of multiplication constant depth and polynomial size.

Wide Lemma 4.1. *Let d be a fixed positive nonzero integer. There is positive nonzero integer $\rho(d)$ such that, for all integers $\lambda>0$ there is a set A of positive integers such that:*

(λ1) *A is wide for d;*

(λ2) *$|A|=\lambda$, and*

(λ3) *$\forall a\in A\ a\leq\lambda^{2\cdot d}\cdot\rho(d)$.*

Proof. Set $\rho(d)=|\mathcal{H}(d:x_1,\ldots,x_d)|^2\cdot(2\cdot d+d^2)$. We shall proceed by induction on λ.
When $\lambda=1$: set $A_1=\{1\}$. Certainly all of (λ1), (λ2) and (λ3) hold.
Induction step at λ: here, the existence of A_λ, (λ1), (λ2) and (λ3) are assumed to hold. Let $x_1,\ldots,x_d$ and $y_1,\ldots,y_d$ be pairwise distinct variables and define the sets of expressions,

$$\mathcal{H}_x=\mathcal{H}(d,\min\{\lambda+1,d\}:x_1,\ldots,x_{\min\{\lambda+1,d\}})$$

and

$$\mathcal{H}_y = \mathcal{H}(d, \min\{\lambda+1, d\} : y_1, \ldots, y_{\min\{\lambda+1,d\}}).$$

Now form the set, $\mathcal{E}_{x,y}$, of all possible linear equations from $\mathcal{H}_x$ and $\mathcal{H}_y$, namely all equations of the form,

$$j_1 \cdot x_1 + \cdots + j_\mu \cdot x_\mu = i_1 \cdot y_1 + \cdots + i_\nu \cdot y_\nu,$$

where $\mu, \nu \leq \min\{\lambda+1, d\}$ and $j_1 + \cdots + j_\mu, i_1 + \cdots + i_\nu \leq d$. Let z be a new variable and form a new set of equations, $\mathcal{E}^z_{x,y}$, which we get from $\mathcal{E}_{x,y}$ by:

- In each member of $\mathcal{E}_{x,y}$, substitute z for at least *one* and at most *two* variables.
- If we substitute *two* variables then one of them is an x-variable and the other is a y-variable.
- If $\lambda < d$ and if an equation involves $\lambda+1$ many x or y-variables, then we are *obliged* to decrease the number of x or y-variables to λ by the appropriate substitution(s).

The members of $\mathcal{E}^z_{x,y}$ have at most $\min\{\lambda, d\}$ many x or y-variables. We now substitute pairwise distinct numbers from A_λ for the x-variables in the equations of $\mathcal{E}^z_{x,y}$; do the same for the y-variables and call the set of these new linear equations in z, $\mathcal{E}^z_{A_\lambda}$.
Linear equations in one variable can have a unique solution, infinitely many solutions or no solution at all. Next we calculate an upper bound for the number of solutions resulting from members of $\mathcal{E}^z_{A_\lambda}$ which have at most one solution. Note that,

$$|\mathcal{E}^z_{x,y}| \leq \left(2 \cdot d + d^2\right) \cdot |\mathcal{H}(d : x_1, \ldots, x_d)|^2 = \rho(d)$$

and that an upper bound for the number of substitutions is $\lambda^{2 \cdot d}$. We conclude that the number of *unique* solutions is $\leq \rho(d) \cdot \lambda^{2 \cdot d}$ and that there is $a \leq \rho(d) \cdot (\lambda+1)^{2 \cdot d}$ which is not a unique solution. Pick the least such a and let $A_{\lambda+1} = A_\lambda \bigcup \{a\}$. That $((\lambda+1)2)$ and $((\lambda+1)3)$ hold is obvious. We now show that $((\lambda+1)1)$ holds as well.
Assume that,

$$\begin{aligned} j_1 \cdot a_1 + \cdots + j_k \cdot a_k + j \cdot a + j_{k+1} \cdot a_{k+1} + \cdots + j_\mu \cdot a_\mu = \\ = i_1 \cdot b_1 + \cdots + i_{k'} \cdot b_{k'} + i \cdot a + i_{k'+1} \cdot b_{k'+1} + \cdots + i_\nu \cdot b_\nu \end{aligned} \qquad (\mathbf{E}_a),$$

with $j_1 \cdot a_1 + \cdots + j_k \cdot a_k + j \cdot z + j_{k+1} \cdot a_{k+1} + \cdots + j_\mu \cdot a_\mu = i_1 \cdot b_1 + \cdots + i_{k'} \cdot b_{k'} + i \cdot z + i_{k'+1} \cdot b_{k'+1} + \cdots + i_\nu \cdot b_\nu \in \mathcal{E}^z_{A_\lambda}$. We need to show that:

(i) $\mu = \nu$;

(ii) the two sequences of numbers, $j_1, \ldots, j_k, j, j_{k+1}, \ldots, j_\mu$ and $i_1, \ldots, i_{k'}, i, i_{k'+1}, \ldots, i_\nu$ are identical, and

(iii) the two vectors, $a_1, \ldots, a_k, a, a_{k+1}, \ldots, a_\mu$ and $b_1, \ldots, b_{k'}, a, b_{k'+1}, \ldots, b_\nu$ are in the same multiplicity class with respect to the expression $j_1 \cdot x_1 + \cdots + j_k \cdot x_k + j \cdot x_{k+1} + j_{k+1} \cdot x_{k+2} + \cdots + j_\mu \cdot x_{\mu+1}$.

A number of claims will prove the above. But first observe that if $j = i = 0$ then by ($\lambda 1$) we get the desired conclusions (i), (ii), and (iii). So we assume that either $j \neq 0$ or that $i \neq 0$, but actually,

Claim 1: *In* $\mathbf{E}_a$, *both* $j \neq 0$ *and* $i \neq 0$.

Proof of claim 1. If exactly one of j or i is 0, then a would be the unique solution to an equation in $\mathcal{E}^z_{A_\lambda}$, contradicting our choice of a.□

Claim 2: *The equation* $\mathbf{E}_z$, *which results from* $\mathbf{E}_a$ *by replacing in* $\mathbf{E}_a$ *the number* a *by the variable* z, *is true for any* z.

Proof of claim 2. Again, if not then a would be the unique solution to an equation in $\mathcal{E}^z_{A_\lambda}$, contradicting the choice of a.□

Claim 3: $j = i$, *and*

$$j_1 \cdot a_1 + \cdots + j_k \cdot a_k + j_{k+1} \cdot a_{k+1} + \cdots + j_\mu \cdot a_\mu = \\ = i_1 \cdot b_1 + \cdots + i_{k'} \cdot b_{k'} + i_{k'+1} \cdot b_{k'+1} + \cdots + i_\nu \cdot b_\nu \, .$$

Proof of claim 3. Obvious from claim 2.□

By the induction hypothesis now, we have that $\mu = \nu$, $j_l = i_l$ for $l = 1, \ldots, \mu$ and the vectors $a_1, \ldots, a_\mu$ and $b_1, \ldots, b_\mu$ belong to the same multipliciy class with respect to the expression $j_1 \cdot x_1 + \cdots + j_\mu \cdot x_\mu$.

Claim 4: *The two sequences of numbers,*

$$j_1, \ldots, j_k, j, j_{k+1}, \ldots, j_\mu \text{ and } i_1, \ldots, i_{k'}, i, i_{k'+1}, \ldots, i_\mu$$

are identical.

Proof of claim 4. Recall that $j_1 \geq \cdots \geq j \geq \cdots \geq j_\mu$ and $i_1 \geq \cdots \geq i \geq \cdots \geq i_\mu$. We have that,

$$j \geq j_{k+1} = i_{k+1} \geq i_{k+2} = j_{k+2} \geq \cdots \geq j_{k'} = i_{k'} \geq i.$$

But $j = i$, so that $j = j_{k+1} = i_{k+1} = \cdots = j_{k'} = i_{k'} = i$. The induction hypothesis takes care of the rest.□

Claim 5: *The two vectors,*

$$a_1, \ldots, a_k, a, a_{k+1}, \ldots, a_\mu \text{ and } b_1, \ldots, b_{k'}, a, b_{k'+1}, \ldots, b_\mu$$

are in the same multiplicity class with respect to the expression

$$j_1 \cdot x_1 + \cdots + j_k \cdot x_k + j \cdot x_{k+1} + j_{k+1} \cdot x_{k+2} + \cdots + j_\mu \cdot x_{\mu+1} .$$

Proof of claim 5. Let k_0 be the least and k_1 the largest, such that $j = j_{k_0} = j_{k_0+1} = \cdots = j_{k_1}$, if such two indices exist. By the remark preceding claim 4 then there is a permutation from $\{a_{k_0}, \ldots, a_{k_1}\}$ to $\{b_{k_0}, \ldots, b_{k_1}\}$. Rearrange and extend this permutation to the index of a, which is possible since a is new.□

So (iii) is claim 5, (ii) is claim 4, and (i) follows from the remark preceding claim 4.□

This is a corrolary to the wide lemma, and it is the central argument for the proof of the circuit lemma.

Bit Lemma 4.2. *Assume that:*

(1) λ *is large, and*

(2) $t(x) = \sum_{i=1}^{n} q_i(x) \cdot 2^{p_i}$ *where:*

  (a) q_i *are polynomials in* x *with non-negative coefficients of size at most* $2^{\lceil \log_2 \lambda \rceil^m}$ *for some fixed* m*, and*

  (b) *with* d *the degree of* x *in* $t(x)$*, we have that,*

$$p_{i+1} - p_i > (d+1) \cdot \lambda .$$

Then:

(1′) *There are* A *and* B *such that:*

(a) $A = \lceil \log_2 \lambda \rceil^{m+1} \cdot B$;

(b) B *is wide for* d;

(c) $|B| = \left\lfloor \sqrt[2\cdot d]{\frac{\lambda}{\lceil \log_2 \lambda \rceil^{m+1} \cdot \rho(d)}} \right\rfloor$, *where* $\rho(d)$ *is as given in the wide lemma, and*

(d) $b \in B \Rightarrow b < \left\lfloor \lambda / \lceil \log_2 \lambda \rceil^{m+1} \right\rfloor$.

(2′) *If* $y \leq |t(2^\lambda)|$ *then there are* $a_1, \ldots, a_\mu \in A$, $\mu \leq d$, *and* $\Gamma_y(v, w_1, \ldots, w_\mu)$ *such that,*

$$\forall\, x \subseteq A\, [\mathbf{BIT}(t(x), y) \Longleftrightarrow \Gamma_y(x, a_1, \ldots, a_\mu)]\,,$$

where $\Gamma_y(v, w_1, \ldots, w_\mu)$ *is one of:*

(a) *false, or*

(b) $\mathbf{BIT}(v, w_1) \wedge \cdots \wedge \mathbf{BIT}(v, w_\mu)$.

Proof. (1′) follows by an application of the wide lemma, with d, $\left\lfloor \sqrt[2\cdot d]{\frac{\lambda}{\lceil \log_2 \lambda \rceil^{m+1} \cdot \rho(d)}} \right\rfloor$, and B in the rôles of d, λ, and A respectively. Let $A = \lceil \log_2 \lambda \rceil^{m+1} \cdot B$. Clearly, $a \in A \Rightarrow a < \lambda$. Call condition (2) of this lemma t *is in block form with spread* $(d+1)\cdot\lambda$ *and height* $2^{\lceil \log_2 \lambda \rceil^m}$. Call the binary expansion of $q_i(x) \cdot 2^{p_i}$ *the* i*-th block of* t. Now fix $y \leq |t(2^\lambda)|$. Certainly, there is i such that the y-th bit of $t(x)$ belongs to the i-th block of $t(x)$, irrespective of x. It follows from (2a) and (2b) that for $x = \sum_{a\in A} x(a) \cdot 2^a$ there is no block interference, because the bits of a block of $t(x)$ do not overflow into the next block of $t(x)$. Hence the y-th bit only depends on the binary expansion of $q_i(x)$. Suppressing the subscript i for simplicity, let

$$q(x) = q_i(x) = \sum_{j=0}^{d} x^j \cdot \varphi_j$$

where $\varphi_j < 2^{\lceil \log_2 \lambda \rceil^m}$. Now according to 4i, however,

$$x^j = \sum_{k=1}^{j} \sum_{\substack{j_1 \geq \cdots \geq j_k \\ j_1 + \cdots + j_k = j \\ a_1, \ldots, a_k \in A \\ \text{pairwise distinct}}}^{\bullet} \mu(j_1, \ldots, j_k) \cdot x(a_1) \cdots x(a_k) \cdot \frac{j!}{j_1! \cdots j_k!} \cdot 2^{j_1 \cdot a_1 + \cdots + j_k \cdot a_k}$$

and therefore,

$$q(x) = \varphi_0 + \sum_{\substack{1 \leq j \leq d \\ 1 \leq k \leq j}} \sum_{\substack{j_1 \geq \cdots \geq j_k \\ j_1 + \cdots + j_k = j \\ a_1, \ldots, a_k \in A, \\ \text{pairwise distinct}}}^{\bullet} \varphi_j \cdot \mu(j_1, \ldots, j_k) \cdot x(a_1) \cdots x(a_k) \cdot \frac{j!}{j_1! \cdots j_k!} \cdot 2^{j_1 \cdot a_1 + \cdots + j_k \cdot a_k}$$

recalling that the $\bullet$ means summation over only *one* member of the multiplicity class of $a_1, \ldots, a_k$ with respect to the expression $j_1 \cdot x_1 + \cdots + j_k \cdot x_k$.
For $j = 1, \ldots, d$ call the binary expansion of

$$\varphi_j \cdot \mu(j_1, \ldots, j_k) \cdot \frac{j!}{j_1! \cdots j_k!} \cdot 2^{j_1 \cdot a_1 + \cdots + j_k \cdot a_k}$$

the $(j_1 \cdot a_1 + \cdots + j_k \cdot a_k)$ *block of* q or just a *q-block.* The $(j_1 \cdot a_1 + \cdots + j_k \cdot a_k)$ block of q is called *relevant* if $\varphi_j \neq 0$.

Claim : *There is no bit-interference between relevant q-blocks of different multiplicity classes out of A.*

Proof of claim . Recall that:

(1) $A = \lceil \log_2 \lambda \rceil^{m+1} \cdot B$, and

(2) B is wide for d.

By (2) and (1) A is also wide for d. Hence, again by (1) the difference between any two numbers $j_1 \cdot a_1 + \cdots + j_\mu \cdot a_\mu$ and $i_1 \cdot b_1 + \cdots + i_\nu \cdot b_\nu$, with $j_1 \cdot x_1 + \cdots + j_\mu \cdot x_\mu$, $i_1 \cdot y_1 + \cdots + i_\nu \cdot y_\nu \in \mathcal{H}(d : x_1, \ldots, x_d)$, is at least $\lceil \log_2 \lambda \rceil^{m+1}$, unless $\mu = \nu$, $j_l = i_l$ for $l = 1, \ldots, \mu$, and both $a_1, \ldots, a_\mu$ and $b_1, \ldots, b_\mu$ belong to the same multiplicity class with respect to $j_1 \cdot x_1 + \cdots + j_\mu \cdot x_\mu$, in which case the difference of the abovementioned numbers is 0. By assumption (2a) in the statement, the length of the binary expansion of

$$\varphi_j \cdot \mu(j_1, \ldots, j_k) \cdot \frac{j!}{j_1! \cdots j_k!}$$

is at most $\lceil \log_2 \lambda \rceil^m + M_d$, where M_d is the length of the binary expansion of the largest amongst the coefficients,

$$\mu(j_1, \ldots, j_k) \cdot \frac{j!}{j_1! \cdots j_k!} ,$$

which only depend on d. Noticing that

$$\frac{\lceil \log_2 \lambda \rceil^{m+1}}{2} > \lceil \log_2 \lambda \rceil^m + M_d ,$$

a small calculation gives the claim.□By the *range* of the $j_1 \cdot a_1 + \cdots + j_k \cdot a_k$-th block of q we shall mean the interval,

$$\left[j_1 \cdot a_1 + \cdots + j_k \cdot a_k - \frac{\lceil \log_2 \lambda \rceil^{m+1}}{2}, j_1 \cdot a_1 + \cdots + j_k \cdot a_k + \frac{\lceil \log_2 \lambda \rceil^{m+1}}{2} \right] \bigcap \mathbf{N}.$$

So, it is clear from the claim that there is at most *one* q-block such that $y - p_i$ is in its range. Now, there are two possibilities: either such a block exists (call it then the *relevant* block for y) or such a block does not exist. In the latter case we are done because the y-th bit of $t(x)$ is 0 irrespective of x; we therefore set $\Gamma_y(\ldots) = 0$. In the former case, we consider the relevant q-block for y, say the $j_1 \cdot a_1 + \cdots + j_k \cdot a_k$-th block of q. We set,

$$\mathcal{R}^+ = \left\{ x \in 2^A \mid x(a_1) \cdots x(a_k) = 1 \right\}$$

and

$$\mathcal{R}^- = \left\{ x \in 2^A \mid x(a_1) \cdots x(a_k) = 0 \right\} .$$

$\mathbf{BIT}(q(x), y - p_i)$ is obviously oblivious in both $\mathcal{R}^-$ and $\mathcal{R}^+$; so let $\mathbf{bit}_+(y)$ and $\mathbf{bit}_-(y)$ denote the respective values of $\mathbf{BIT}(q(x), y - p_i)$ on the sets $\mathcal{R}^-$ and $\mathcal{R}^+$. We set,

$$\Gamma_y(v, \vec{w}) = \begin{cases} \mathit{false}, & \text{if } \mathbf{bit}_+(y) \text{ is } \mathit{false}; \\ \mathbf{BIT}(v, w_1) \wedge \cdots \wedge \mathbf{BIT}(v, w_\mu), & \text{if } \mathbf{bit}_+(y) \text{ is } \mathit{true}; \end{cases}$$

□

This lemma critically uses the length lemma. It will serve to fulfill the assumptions of the bit lemma so that it can be used in proving the forthcoming circuit lemma.

Block Lemma 4.3. *Let $t(x, \vec{z})$ be a term of bounded arithmetic and m a positive nonzero integer. Then there are positive nonzero integers m_t, κ_t and ℓ_t, such that with $\nu \geq \nu(t)$,*
$\forall \sigma < 2^\nu \, \forall \ell \geq \ell_t \, \forall \vec{c} < \ell^m$,
$\exists \mathcal{P}$ a partition of $[2^\ell, 2^{\ell+1}) \bigcap [\sigma]_\nu$ into at most κ_t many intervals,
$\forall P \in \mathcal{P}$ and for x varying in P
$t(x, \vec{c})$ is in block form with spread $\ell - \lceil \log_2 \ell \rceil^{m_t}$ and height $2^{\lceil \log_2 \ell \rceil^{m_t}}$.

Proof. Fix $\vec{c} < \ell^m$ and $\sigma < 2^\nu$. By the length lemma we get a partition P of $[2^\ell, 2^{\ell+1}) \bigcap [\sigma]_\nu$ with the properties mentioned there. For $P \in \mathcal{P}$ then as in the proof of the atomic truth lemma we get the expression $t_{P,\vec{c},\sigma}(x)$ which, over P, can be expressed as

$$\sum_{i=0}^{d} x^i \cdot \gamma_i(x, \vec{c}) ,$$

where:

(1) x can only occur in the γ_i within the scope of '$|\_|$' or '$\_\#\_$', and

(2) $\gamma_i(x,\vec{c}) = \sum_{j=1}^{n} r_{i,j}(x,\vec{c}) \geq 0$, with the $r_{i,j}$'s products of length and smash terms on the one hand and members of $\vec{c}$ on the other, possibly together with divisions by $2, 2^2, \ldots, 2^\nu$.

To show that $t_{P,\vec{c},\sigma}(x)$ is in block form with spread $\ell - \lceil \log_2 \ell \rceil^{m_t}$ and height $2^{\lceil \log_2 \ell \rceil^{m_t}}$ where $m_t - 1$ is the largest number of nested smash applications in $t(x,\vec{z})$, we need the following claim which depends on the proof of the length lemma,

Claim : *Let* $r = r_{i,j}$ *and* $s = r_{i',j'}$ *for some* $i, i' \leq d$ *and* $j, j' \leq n$. *Then, one of the following holds:*

(c1) $-\lceil \log_2 \ell \rceil^{m_t} \leq |r| - |s| \leq \lceil \log_2 \ell \rceil^{m_t}$;

(c2) $|r| - |s| > \ell - \lceil \log_2 \ell \rceil^{m_t}$; *or*

(c3) $|r| - |s| < -\ell + \lceil \log_2 \ell \rceil^{m_t}$.

Proof of claim . We work in P. The procedure in the *proof* of the length lemma of gives $\hat{r}(\ell+1,\vec{c},\vec{\epsilon_r}) = |r(x,\vec{c})|$ and $\hat{s}(\ell+1,\vec{c},\vec{\epsilon_s}) = |s(x,\vec{c})|$. Now, both $\hat{r}$ and $\hat{s}$ are constant because the vector $\vec{c}$ is fixed and because we are in P. Choose $\vec{\epsilon_r}, \vec{\epsilon_s}$ such that the polynomials $\hat{r}(\ell+1,\vec{c},\vec{\epsilon_r})$ and $\hat{s}(\ell+1,\vec{c},\vec{\epsilon_s})$ evaluate to $|r(x,\vec{c})|$ and $|s(x,\vec{c})|$. Consider $\hat{r} - \hat{s}$ as function of ℓ. This is a polynomial expression in ℓ of some fixed degree depending entirely on the *syntactic structure* of the terms $r(x,\vec{z})$ and $s(x,\vec{z})$, and with coefficients of size at most $\lceil \log_2 \ell \rceil^{m_t}$. Let j be the highest power of ℓ in $\hat{r} - \hat{s}$ with non-zero coefficient, then we conclude that:

(ℓ1) if $j > 0$ then the absolute value of $\hat{r} - \hat{s}$, is at least $\ell - \lceil \log_2 \ell \rceil^{m_t}$, and

(ℓ2) if $j = 0$ then the absolute value of $\hat{r} - \hat{s}$, is at most $\lceil \log_2 \ell \rceil^{m_t}$.

□By the claim, then, we can define an equivalence relation and classes among the $|r_{i,j}|$'s: put two of them in the same class just in case the absolute value of the difference of the lengths of their respective binary expansions is at most $\lceil \log_2 \ell \rceil^{m_t}$. It now follows that,

$$t_{P,\vec{c},\sigma}(x) = \sum_{i=1}^{l} 2^{p_i(x,\vec{c})} \cdot f_i(x,\vec{c}) \,,$$

where $2^{p_i(x,\vec{c})}$ is the minimum power of 2 amongst the powers of 2 that occur in the $i-$th equivalence class and the f_i are polynomial expressions in x with non-negative coefficients of size at most $2^{\lceil \log_2 \ell \rceil^{m_t}}$.□

This lemma is the obvious main step to the reduction of bit sharply bounded formulae to constant depth and polynomial size circuits.

Circuit Lemma 4.4. *Let $\Theta(x)$ be a bit sharply bounded formula. There are positive nonzero integers r, δ, ℓ_Θ, κ, m and d depending on the syntactic structure of $\Theta(x)$ such that with $k =$(number of quantified variables in $\Theta(x)$), $n = m \cdot k$, $\alpha = 2 \cdot \kappa \cdot \ell^n$ and*

$$\nu \geq \max_{t \text{ a term of } \Theta} \nu(t) .$$

we have that,

$\forall \sigma < 2^\nu$, $\forall \ell \geq \ell_\Theta$,

$\exists A$, *with* $|A| > \sqrt[2\cdot d+1]{\ell}$ *and* $A \subseteq [0, (\ell - \nu - \lceil \log_2 \alpha \rceil)/(d+2)) \bigcap \mathbf{N}$,

$\exists \beta < \alpha \exists \mathbf{C}_\ell^\Theta$, *a circuit of depth δ and size at most ℓ^r* ,

$\forall x \subseteq A \left[\Theta(\sigma + 2^\ell + \beta \cdot 2^{\ell - \lfloor \log_2 \alpha \rfloor} + 2^\nu \cdot x) \Longleftrightarrow \mathbf{C}_\ell^\Theta(x) \right]$.

Proof. Fix $\sigma < 2^\nu$. Let $\Psi(x, \vec{z})$ be the quantfier-free matrix of $\Theta(x)$. From the start let us replace by *false* all bit-extractors of which the bit, $s(x, \vec{z})$, depends on occurrences of x which are not within the scope of '$|\_|$' or '$\_\#\_$': this is justfied because for all large lengths ℓ, with $|x| = \ell + 1$, and any terms $t(x, \vec{z})$ and $s(x, \vec{z})$, with s of the above kind, we have that

$$|t(x, \vec{c})| < s(x, \vec{c}) ,$$

for any $\vec{c} < \ell^m$.

Superimposing the partitions given by the block lemma and the length lemma, we pass to an subinterval, say P, of $[2^\ell, 2^{\ell+1}) \bigcap [\sigma]_\nu$ of size at least $2^{\ell-\nu}/(\kappa \cdot \ell^m)$, and find positive integer M, with M, ν, κ and m depending solely on the syntactic structure of $\Psi(x, \vec{z})$ and its terms, such that for any $\vec{c} < \ell^n$:

(1) all positive atomic statements of $\Psi(x, \vec{c})$ which are not extractors, are oblivious in P;

(2) the bits of the bit-extractors in $\Psi(x, \vec{c})$ are constant in P, and

(3) the arguments of the bit-extractors in $\Psi(x, \vec{c})$ are in block form with spread ℓ and height $2^{\lceil \log_2 \lambda \rceil^M}$ in P.

From the above it follows that *in* P and for any $\vec{c} < \ell^n$, $\Psi(x, \vec{c})$ is reduced to a boolean combination of certain bit-extractors. The strategy now, is clear: *first* transform the domain, P, of the terms, and *then* apply the bit lemma simultaneously to all arguments of the bit-extractors. Now the details.

Fix a set Q such that:

(1) $\sigma + 2^\nu \cdot Q \subseteq P$, and

(2) $Q = \left[2^\lambda + \frac{\beta}{\alpha} \cdot 2^\lambda, 2^\lambda + \frac{\beta+1}{\alpha} \cdot 2^\lambda\right) \bigcap \mathbf{N}$, where $\alpha = 2 \cdot \kappa \cdot \lambda^m$ and $\beta < \alpha$.

Now,

$$Q = \left((2^\lambda + \beta \cdot \frac{2^\lambda}{\alpha}) + [0, 2^\lambda/\alpha)\right) \bigcap \mathbf{N}.$$

It follows that,

$$\left((2^\lambda + \beta \cdot 2^{\lambda - \lceil \log_2 \alpha \rceil}) + [0, 2^{\lambda - \lfloor \log_2 \alpha \rfloor})\right) \bigcap \mathbf{N} \subseteq Q.$$

Put $Q' = \left[0, 2^{\lambda - \lfloor \log_2 \alpha \rfloor}\right) \bigcap \mathbf{N}$.

Consider an extractor of $\Psi(x, \vec{z})$ with argument $t(x, \vec{z})$ and bit $s(x, \vec{z})$. By (2) above we can write $s_P(\vec{c})$ for the (unique) value of $s(x, \vec{c})$ in P. Define next,

$$\begin{aligned} t_\beta(x, \vec{z}) &= t(\sigma + 2^{\lambda+\nu} + \beta \cdot 2^{\lambda+\nu-\lfloor \log_2 \alpha \rfloor} + 2^\nu \cdot x, \vec{z}) \\ &= t(\sigma + 2^\ell + \beta \cdot 2^{\ell - \lfloor \log_2 \alpha \rfloor} + 2^\nu \cdot x, \vec{z}), \end{aligned}$$

and carry out the obvious extension to $\Psi_\beta(x, \vec{z})$ and $\Theta_\beta(x)$. It easy to see that $t_\beta(x, \vec{c})$ is in block form with spread ℓ and height $2^{\lceil \log_2 \lambda \rceil^{M+1}}$. Let

$$d = \max_{t \text{ a term of } \Theta} \{\text{degree of } x \text{ in } t\}.$$

We apply the bit lemma *simultaneously* to all $t_\beta(x, \vec{c})$, $\vec{c} < \ell^n$ and $t(x, \vec{z})$ an argument of an extractor in $\Psi(x, \vec{z})$, with

$$\lambda' = \frac{\ell - \nu - \lceil \log_2 \alpha \rceil}{d+2} = \frac{\lambda - \lceil \log_2 \alpha \rceil}{d+2}.$$

We get a set A such that:

(A1) $|A| > \sqrt[2 \cdot d + 1]{\ell}$;

(A2) $a \in A \Longrightarrow a < \lambda'$, and

(*A*3) for any $\vec{c} < \ell^n$, and any extractor in $\Psi(x, \vec{z})$ with argument $t(x, \vec{z})$ and bit $s(x, \vec{z})$ there are $\mu(s, \vec{c}) \leq d$ and ${}^{s_P(\vec{c})}a_1, \ldots, {}^{s_P(\vec{c})}a_{\mu(s,\vec{c})} \in A$ such that,

$$\forall\, x \subseteq A\ \mathbf{BIT}(t_\beta(x, \vec{c}), s_P(\vec{c})) \Longleftrightarrow \Gamma_{s_P(\vec{c})}\left(x, {}^{s_P(\vec{c})}a_1, \ldots, {}^{s_P(\vec{c})}a_{\mu(s_P(\vec{c}))}\right),$$

where the $\Gamma_{s_P(\vec{c})}$ *one* of (2'a) or (2'b), as found in the statement of the bit lemma.

We have transformed the domain of $\Theta(x)$ and produced the cylinder 2^A via the obvious extension of the bit lemma to several terms simultaneously . In 2^A the extractors are boolean combinations of certain bits of x. Next, we indicate the construction of the circuit. As remarked earlier on, in P and for any $\vec{c} < \ell^n$,

$$\Psi(x, \vec{c}) \Longleftrightarrow f_{bool}^{P,\vec{c}}\left(\mathbf{BIT}(t(x, \vec{c}), s(x, \vec{c})) \ : \text{ ranging over extractors in } \Psi\right),$$

where $f_{bool}^{P,\vec{c}}$ is a boolean combination. It follows that, $\forall\, x \subseteq A$,

$$\Psi_\beta(x, \vec{c}) \Longleftrightarrow f_{bool}^{P,\vec{c}}\left(\Gamma_{s_P(\vec{c})}(x, {}^{s_P(\vec{c})}a_1, \ldots, {}^{s_P(\vec{c})}a_{\mu(s_P(\vec{c}))}) : \ s \text{ is a bit in } \Psi\right).$$

For $\vec{c} < \ell^n$ collect all relevant points for all the $s_P(\vec{c})$, $s(x, \vec{z})$ a bit in $\Psi(x, \vec{z})$, say ${}^{\vec{c}}a_1, \ldots, {}^{\vec{c}}a_{\mu(\vec{c})}$, where $\mu(\vec{c}) \leq (d \times$ the number of extractors in $\Psi(x, \vec{z}))$. Then for $\vec{c} < \ell^n$,

$$\forall\, x \subseteq A\ \left[\Psi_\beta(x, \vec{c}) \Longleftrightarrow g_{bool}^{P,\vec{c}}\left(\mathbf{BIT}(x, {}^{\vec{c}}a_1), \ldots, \mathbf{BIT}(x, {}^{\vec{c}}a_{\mu(\vec{c})})\right)\right].$$

Finally, in Θ_β, turning the $\exists$'s to $\bigvee$'s and the $\forall$'s to $\bigwedge$'s we obtain the circuit $\mathbf{C}_\ell^\Theta$.□

Proof of the theorem

Assume that $\Theta(x)$ is a bit formula which defines **PARITY**. Apply the circuit lemma to $\Theta(x)$, getting constant depth, polynomial size circuitry $\left\{\mathbf{C}_\ell^\Theta\right\}_{\ell=\ell_\Theta}^\infty$, numbers $\{\beta_\ell\}_{\ell=\ell_\Theta}^\infty$, and sets $\{A_\ell\}_{\ell=\ell_\Theta}^\infty$, such that for any $\ell \geq \ell_\Theta$:

(1) $a \in A_\ell \Longrightarrow a < \ell$;

(2) $|A_\ell| > \sqrt[2\cdot d+1]{\ell}$, and

(3) $x \subseteq A_\ell\ \left[\Theta_{\beta_\ell}(x) \Longleftrightarrow \mathbf{C}_\ell^\Theta(x)\right]$.

Because Θ defines **PARITY** it follows that Θ_{β_ℓ} defines one of **PARITY** or $\neg$**PARITY** over 2^{A_ℓ}. Consider $\ell' = \ell^{2\cdot d+1}$. Certainly then, by fixing appropriately a few bits in $A_{\ell'}$ and hence passing to a subset B_ℓ of $A_{\ell'}$, with $|B_\ell| = \ell$, and a new circuit $\tilde{\mathbf{C}}_\ell^\Theta$ we ascertain,

$$\forall x \subseteq B_\ell \left[\tilde{\mathbf{C}}_\ell^\Theta(x) \Longleftrightarrow \mathbf{PARITY}(x)\right] .$$

It is now clear that the circuitry $\left\{\tilde{\mathbf{C}}_\ell^\Theta\right\}_{\ell,\,\text{large}}^{\infty}$ defines **PARITY** and is constant depth polynomial size. This contradicts the well known results in [Ajt83] and [FSS84]. ■

Chapter 5

INDEXPARITY is not $\mathcal{SB}_1^{\mathrm{N}}(\mathbf{BIT}, \mathbf{CARD})$ definable

The main result in this chapter is

Theorem. **INDEXPARITY** *is not bit symmetric sharply bounded definable.*

We enhance the language of the previous chapter by a binary predicate $\mathbf{CARD}(x, y)$, to be able to talk about the "cardinality" of terms of bounded arithmetic. The resulting sharply bounded formulae are then called bit symmetric sharply bounded. The name suggests that a variety of "symmetric" predicates are definable now. Indeed, symmetric predicates such as **PARITY** and **MAJORITY** are now definable. It is shown, however, that **INDEXPARITY**, *a polynomial time computable* predicate is not bit symmetric sharply bounded definable. This is a partial answer to the question of the power of nonuniform symmetric circuits, i.e. circuits which use gates which calculate symmetric predicates.

Notation and conventions

Previous notation and conventions apply, unless locally changed. Additionally, note the following:

5a. Odd_ℓ = odd numbers$< \ell$, Even_ℓ = even numbers$< \ell$.

5b. $\mathbf{card}(x)$ = number of $1's$ in the binary expansion of x.

5c. We add a new binary predicate symbol to the language of chapter 3, $\mathbf{CARD}(x, y)$, to be interpreted as follows,

$$\mathbf{CARD}(x, y) \text{ is true } \Longleftrightarrow \mathbf{card}(x) = y \ .$$

The new sharply bounded formulae are now called *bit symmetric sharply bounded.*

5d. We define **INDEXPARITY** as follows,

$$\mathbf{INDEXPARITY}(x) \text{ is true} \Longleftrightarrow \left(\sum_{i,\, x(i)=1} i \right) \bmod 2 = 1 \ .$$

5e. $A^{(\lambda)} = \{x \subseteq A \mid \mathbf{card}(x) = \lambda\}$.

5f. We say $\mathbf{C} = \{\mathbf{C}_\ell\}_{\ell=1}^{\infty}$ *is a multicircuit of constant depth and polynomial size* if there is δ, a positive integer, and two polynomials p_{size} and p_{width} such that for each ℓ:

(1) $\mathbf{C}_\ell = \{C_{\ell,j}\}_{j \leq p_{width}(\ell)}$, and

(2) the boolean circuits $C_{\ell,j}$ are of constant depth, δ, and size, at most $p_{size}(\ell)$.

Given δ, a positive integer, p_{size} and p_{width} polynomials, we shall say $\mathbf{C}$ *is a multicircuit of depth* δ, *size* p_{size}, *and width* p_{width}, if (1) and (2) above hold.

5g. Recall the relevant notation and definitions concerning restrictions and probabilistic restrictions from chapter 1. Furthermore, the notation $x \in \rho$, with ρ a restriction, say on X and $x \in \{0,1\}^X$, makes sense since restrictions are both sets of sequences and functions.

Probabilistic lemmata

This is the first and key lemma which guarantees that each time a random restriction is applied to the circuits, there are $*-$ed odd indices and $*-$ed even indices.

Two Set Lemma 5.1. *Let ϵ be a non-zero positive real number less than 1. Then there is positive integer ℓ_ϵ such that,*
$\forall \ell \geq \ell_\epsilon \ \forall A \subseteq \{1, \ldots, \ell\}$ *with* $|A| \geq \epsilon \cdot \ell$

$$\mathbf{Prob}\left(|\mathrm{free}(\hat{\rho}) \bigcap A| < \frac{\epsilon}{2} \cdot \sqrt{\ell} \right) = o(\ell^{-1/2}).$$

Proof. Let $\lambda = |A|$. Note that with $p = p_* = \frac{1}{\sqrt{\ell}}$ and $q = p_0 + p_1 = 1 - p$ we have that,

$$\mathbf{Prob}\left(|\text{free}(\rho) \bigcap A| < \tfrac{\epsilon}{2} \cdot \sqrt{\ell}\right) = \sum_{j=0}^{\frac{\epsilon}{2} \cdot \sqrt{\ell}} \binom{\lambda}{j} \cdot p^j \cdot q^{\lambda - j} .$$

We consider the λ Bernoulli trials with parameters p, q. The number of successes, S, has then Binomial distribution with parameters p, q and λ. Then,

$$\sum_{j=0}^{\frac{\epsilon}{2} \cdot \sqrt{\ell}} \binom{\lambda}{j} \cdot p^j \cdot q^{\lambda - j} = \mathbf{Prob}\left(S < \tfrac{\epsilon}{2} \cdot \sqrt{\ell}\right) \leq \mathbf{Prob}\left(|S - p \cdot \lambda| > \tfrac{\epsilon}{2} \cdot \sqrt{\ell}\right)$$

because

$$S - p \cdot \lambda < \tfrac{\epsilon}{2} \cdot \sqrt{\ell} - p \cdot \lambda < \tfrac{\epsilon}{2} \cdot \sqrt{\ell} - \tfrac{1}{\sqrt{\ell}} \cdot \epsilon \cdot \ell = -\tfrac{\epsilon}{2} \cdot \sqrt{\ell}.$$

By Chebychev's inequality we have that

$$\leq \mathbf{Prob}\left(|S - p \cdot \lambda| > \tfrac{\epsilon}{2} \cdot \sqrt{\ell}\right) \leq \tfrac{4 \cdot \lambda \cdot \frac{1}{\sqrt{\ell}} \cdot (1 - \frac{1}{\sqrt{\ell}})}{\epsilon^2 \cdot \ell} \leq \tfrac{4}{\epsilon^2} \cdot \tfrac{1}{\sqrt{\ell}}.$$

This finishes the proof.□

The following lemma concerning mlticircuits will also be referred to as the **Multicircuit Lemma** which has the same statement and proof as the lemma below but without the additional requirement on the nature of free(ρ).

Multicircuit-Odd-Even Lemma 5.2. *Let $\mathbf{C} = \{\mathbf{C}_\ell\}_{\ell=1}^{\infty}$ be a multicircuit of depth d, size p_{size} and width p_{width}. Then there are nonzero reals ϵ_0 and ϵ_1 less than 1, positive integers ℓ_0 and κ, and a polynomial r such that,*

$\forall \ell \geq \ell_0 \exists \rho \in \text{Restr}_{\epsilon_0}(\ell)$

$|\text{free}(\rho) \bigcap \text{Odd}_\ell|, |\text{free}(\rho) \bigcap \text{Even}_\ell| \geq \epsilon_1 \cdot \ell^{\epsilon_0}$,

$$\mathbf{C}_\ell|\rho \equiv \mathbf{C}'_\ell \text{ a } \kappa\text{-finite multicircuit of size} \leq r.$$

Proof. By induction on d. For $d = 0$, $\mathbf{C}_\ell$ is a set of literals and so $\mathbf{C}_\ell$ is 1-finite. By the two set lemma we get the appropriate restriction ρ. Assume the statement for d, $d \geq 0$. Let $\mathbf{C}$ be a multicircuit of depth $d+1$. Collect all d-children of $\mathbf{C}_\ell$ into a new multicircuit $\mathbf{D}_\ell$ and apply the induction hypothesis on $\mathbf{D}$. We get $\kappa', \epsilon'_0, \epsilon'_1$ and polynomial r' such that for large ℓ there is $\rho' \in \text{Restr}_{\epsilon'_0}(\ell)$ with

1. $|\text{free}(\rho') \bigcap \text{Odd}_\ell|, |\text{free}(\rho') \bigcap \text{Even}_\ell| \geq \epsilon_1' \cdot \ell^{\epsilon_0'}$, and

2. $\mathbf{D}_\ell|\rho' \equiv \mathbf{D}_\ell'$ a κ'-finite multicircuit of size $\leq r'$.

Evidently, $\mathbf{C}_\ell|\rho' \equiv \mathbf{C}_\ell'$ a κ'-short multicircuit of size r'. Proceed exactly as in the proof of the conversion lemma, incorporating the fact that

$$\mathbf{Prob}\left(|\text{free}(\rho) \bigcap \text{Odd}_\ell|, |\text{free}(\rho) \bigcap \text{Even}_\ell| < \frac{\epsilon_1'}{2} \cdot \ell^{\epsilon_0'/2}\right) = o(\ell^{-\epsilon_0'/2}).$$

We get $\kappa, \epsilon_0, \epsilon_1$ and a polynomial r such that for large ℓ there is a restriction $\rho \subseteq \rho', \rho \in \text{Restr}_\epsilon(\ell)$ with

1. $|\text{free}(\rho) \bigcap \text{Odd}_\ell|, |\text{free}(\rho) \bigcap \text{Even}_\ell| \geq \epsilon_1 \cdot \ell^{\epsilon_0}$, and

2. $\mathbf{C}_\ell'|\rho \equiv \mathbf{C}_\ell''$ a κ-finite multicircuit of size $\leq r$.

It is clear that $\mathbf{C}_\ell|\rho \equiv \mathbf{C}_\ell''$ which finishes the induction.□

Non-probabilistic lemmata

The analogue of the bit lemma is stated and proved here.

Cardinality Lemma 5.3. *Let m be a nonzero positive integer, ℓ be large, and assume that $t(x)$ be in block form with spread $(d+1)\cdot\ell$ and height $2^{\lceil \log_2 \ell \rceil^m}$. Then there is a set $A \subseteq \{0,\ldots,\ell-1\}$ with $|A| > \sqrt[2\cdot d+1]{\ell}$ such that,*

$$\forall \lambda \leq |A|\ \mathbf{card}(t(x)) \text{ is constant in } A^{(\lambda)}.$$

Proof. We proceed exactly as in the proof of the bit lemma. We get a set A such that there is no bit-interference between relevant q-blocks of different multiplicity classes out of A. Recall that $t(x) = \sum_{i=1}^n q_i(x) \cdot 2^{p_i}$, $q_i(x) = \sum_{j=0}^d x^j \cdot \varphi_{i,j}$ for $i, 1 \leq i \leq n$, and that according to 4i,

$$x^j = \sum_{k=1}^{j} \dot{\sum_{\substack{j_1 \geq \cdots \geq j_k \\ j_1+\cdots+j_k=j \\ a_1,\ldots,a_k \in A \\ \text{pairwise distinct}}}} \mu(j_1,\ldots,j_k) \cdot x(a_1)\cdots x(a_k) \cdot \frac{j!}{j_1!\cdots j_k!} \cdot 2^{j_1\cdot a_1+\cdots+j_k\cdot a_k},$$

recalling that the $\bullet$ means summation over only *one* member of the multiplicity class of $a_1,\ldots,a_k$ with respect to the expression $j_1\cdot x_1+\cdots+j_k\cdot x_k$.
The fact now, that the $\varphi_{i,j}\geq 0$ implies that for $x\subseteq A$,

$$\mathbf{card}(t(x))=\sum_{i=1}^{n}\mathbf{card}(\varphi_{i,0})+ \\ +\sum_{j=1}^{d}\sum_{k=1}^{j}\sum_{\substack{j_1\geq\cdots\geq j_k\\ j_1+\cdots+j_k=j}}\binom{\mathbf{card}(x)}{k}\cdot\frac{k!}{\mu(j_1,\ldots,j_k)}\cdot card\left(\xi_{i,j}(j_1,\ldots,j_k)\right)\quad ,$$

where, $\xi_{i,j}(j_1,\ldots,j_k)=\varphi_{i,j}\cdot\mu(j_1,\ldots,j_k)\cdot\frac{j!}{j_1!\cdots j_k!}$. So that for $x\subseteq A$, $\mathbf{card}(t(x))$ is a polynomial of $\mathbf{card}(x)$, with degree and coefficients depending only on the degree and the cardinality of the coefficients of $t(x)$. For further reference, call the polynomial $p_{card,t}$.$\square$

The relationship between the cardinality lemma and this lemma resembles closely the relationship between the bit lemma and the circuit lemma.

Symmetric Multicircuit Lemma 5.4. *Let $\Theta(x)$ be a bit symmetric sharply bounded formula. There are positive nonzero integers r, δ, ℓ_Θ, κ, m and d depending on the syntactic structure of $\Theta(x)$ such that with $k=$(number of quantified variables in $\Theta(x)$), $n=m\cdot k$, $\alpha=2\cdot\kappa\cdot\ell^n$ and*

$$\nu\geq\max_{t\text{ a term of }\Theta}\nu(t)$$

we have that,
$\forall\sigma<2^\nu\ \forall\ell\geq\ell_\Theta$,
$\exists A$ *with* $|A|>\sqrt[2\cdot d+1]{\ell}$ *and* $A\subseteq[0,(\ell-\nu-\lceil\log_2\alpha\rceil)/(d+2))\bigcap\mathbf{N}$,
$\exists\beta<\alpha\ \exists\mathbf{C}_\ell^\Theta$, *a multicircuit of depth δ size at most ℓ^r and width exactly* $|A|(\leq\ell)$,
$\forall\lambda\leq|A|\ \forall x\in A^{(\lambda)}\left[\Theta(\sigma+2^\ell+\beta\cdot 2^{\ell-\lfloor\log_2\alpha\rfloor}+2^\nu\cdot x)\Longleftrightarrow C_{\ell,\lambda}^\Theta(x)\right]$.

Proof. Once more, we proceed exactly as in the proof of the circuit lemma. In addition to applying the bit lemma to arguments of bit-extractors, here we apply the cardinality lemma *simultaneously* to terms $t_\beta(x,\vec{c})$ for $\vec{c}<\ell^n$, which occur as the arguments to instances of **CARD**getting polynomials $\vec{c}p_{card,t_\beta}$ such that for $x\subseteq A$ and $\vec{c}<\ell^n$,

$$\mathbf{CARD}(t_\beta(x,\vec{c}),s_P(\vec{c}))\Longleftrightarrow^{\vec{c}}p_{card,t_\beta}(\mathbf{card}(x))=s_P(\vec{c})\ .$$

$\Psi(x,\vec{z})$ again, being the quantifier free matrix of $\Theta(x,\vec{z})$, we conclude that for $x\subseteq A$ and

$\vec{c} < \ell^n$,

$$\Psi_\beta(x,\vec{c}) \Longleftrightarrow g^{\vec{c}}_{bool}\Big(\mathbf{BIT}(x,{}^{\vec{c}}a_1),\ldots\mathbf{BIT}(x,{}^{\vec{c}}a_{\mu(\vec{c})});\\ {}^{\vec{c}}p_{card,t_\beta}(\mathbf{card}(x)) = s_P(\vec{c}) : \mathbf{CARD}(t,s) \text{ occurring in } \Psi\Big) \ .$$

$\mathbf{CARD}(t_\beta(x,\vec{c}), s_P(\vec{c}))$ is oblivious in $A^{(\lambda)}$, so that for $x \in A^{(\lambda)}$ and $\vec{c} < \ell^n$,

$$\Psi_\beta(x,\vec{c}) \Longleftrightarrow^{(\lambda)} g^{\vec{c}}_{bool}\Big(\mathbf{BIT}(x,{}^{\vec{c}}a_1),\ldots,\mathbf{BIT}(x,{}^{\vec{c}}a_{\mu(\vec{c})})\Big) \ ,$$

where the boolean function ${}^{(\lambda)}g^{\vec{c}}_{bool}$ comes from $g^{\vec{c}}_{bool}$ after all literals " ${}^{\vec{c}}p_{card,t_\beta}(\lambda) = s_P(\vec{c})$ " have been replaced by their truth value in $g^{\vec{c}}_{bool}$. We obtain the multicircuit $\mathbf{C}^\Theta$ as follows:

(1) For $\lambda \leq |A|$ we obtain circuit $\mathbf{C}^\Theta_{\ell,\lambda}$ by converting $\exists$'s to $\bigvee$'s, $\forall$'s to $\bigwedge$'s in Θ_β, and appending to each branch created by $\vec{c} < \ell^n$,

$${}^{(\lambda)}g^{\vec{c}}_{bool}\Big(\mathbf{BIT}(x,{}^{\vec{c}}a_1),\ldots\mathbf{BIT}(x,{}^{\vec{c}}a_{\mu(\vec{c})})\Big) \ ; \text{and}$$

(2) we let $\mathbf{C}^\Theta_\ell = \{\mathbf{C}^\Theta_{\ell,\lambda}\}_{\lambda=0}^{|A|}$.

□

Proof of the theorem

Let $\Theta(x)$ be a bit symmetric sharply bounded formula in x. Let d be the highest degree of x in the terms occurring in Θ, m be the highest number of nested applications of smash in the terms of Θ, and $\nu \geq d+m+1+\max\{\nu(t) \mid t \text{ a term in } \Theta\}$. Put $\sigma = 2^d + 2^{d+m}$. $\sigma < 2^\nu$. By the symmetric multicircuit lemma then, and in its notation,

$$\forall\lambda \leq |A|\ \forall x \in A^{(\lambda)}\Big[\Theta(\sigma + 2^\ell + \beta\cdot 2^{\ell-\lfloor\log_2\alpha\rfloor} + 2^\nu\cdot x) \Longleftrightarrow C^\Theta_{\ell,\lambda}(x)\Big] \ ,$$

with ℓ large.

Remark. A can be chosen in such a way that $\min A$ is even, and if $a < b$, $a, b \in A$ and b is the next number after a in A, then one of a or b is even, and then the other is odd.

With $\ell_A = |A| = \lfloor \sqrt[2\cdot d+1]{\ell}\rfloor$ the multicircuit $\mathbf{C}^\Theta$ can be thought of as a multicircuit on ℓ_A by replacing elements of A with elements of $[\ell_A]$ in a fixed enumeration of A which sends even numbers to even and odd to odd. So we can apply the multicircuit-Odd-Even lemma to $\mathbf{C}^\Theta$ and get ϵ_0 and ϵ_1, a restriction $\rho \in \mathrm{Restr}_{\epsilon_0}(\ell_A)$ and a positive integer κ, such that:

(1) $\mathbf{C}^{\Theta}_{\ell_A}|\rho \equiv \mathbf{D}$ a κ-finite multicircuit;

(2) $|\text{free}(\rho)\bigcap \text{Odd}_{\ell_A}|, |\text{free}(\rho)\bigcap \text{Even}_{\ell_A}| \geq \epsilon_1 \cdot \ell_A^{\epsilon_0}$.

Identify ρ with the restriction it defines on A via the fixed enumeration and then extend it to $[\ell]$ by setting $\rho(z) = 0$ for $z \notin A$. We see that $x \in \rho \Longrightarrow x \subseteq A$. So, for $\lambda \leq |A|$ there are $\kappa(\lambda) \leq \kappa$ bits ${}^{\lambda}a_0, \ldots, {}^{\lambda}a_{\kappa(\lambda)} \in A$ such that for any $x \in \rho$

$$\begin{aligned}\Theta(\sigma + 2^{\ell} + \beta \cdot 2^{\ell - \lfloor \log_2 \alpha \rfloor} + 2^{\nu} \cdot x) &\iff C^{\Theta, A}_{\ell,\lambda}(x) \\ &\iff f_{\lambda}(x({}^{\lambda}a_0), \ldots, x({}^{\lambda}a_{\kappa(\lambda)}))\,.\end{aligned}$$

Take $\lambda = 2 + \kappa + \mathbf{card}(\rho|\text{unfree}(\rho))$. With $\bigoplus$ denoting the exclusive "or", we have that for $x \subseteq A$,

$$\begin{aligned}&\mathbf{INDEXPARITY}(\sigma + 2^{\ell} + \beta \cdot 2^{\ell - \lfloor \log_2 \alpha \rfloor} + 2^{\nu} \cdot x) \iff \\ &\iff \mathbf{INDEXPARITY}(x) \bigoplus \mathbf{INDEXPARITY}(\sigma + 2^{\ell} + \beta \cdot 2^{\ell - \lfloor \log_2 \alpha \rfloor}).\end{aligned}$$

Now, the truth of $\mathbf{INDEXPARITY}(\sigma + 2^{\ell} + \beta \cdot 2^{\ell - \lfloor \log_2 \alpha \rfloor})$ is fixed because it does not involve x. Consider the bits ${}^{\lambda}a_0, \ldots, {}^{\lambda}a_{\kappa(\lambda)}$. Certain of these bits may be members of unfree(ρ). Extend ρ to ρ' by letting $\rho'(a) = 1$ for $a \in \text{free}(\rho)$ and a amongst the ${}^{\lambda}a_0, \ldots, {}^{\lambda}a_{\kappa(\lambda)}$. So, f_{λ} is constant in $A^{(\lambda)}\bigcap \rho'$ and by the choice of λ, $A^{(\lambda)}\bigcap \rho' \neq \emptyset$. We conclude that $\Theta(\sigma + 2^{\ell} + \beta \cdot 2^{\ell - \lfloor \log_2 \alpha \rfloor} + 2^{\nu} \cdot x)$ is oblivious in $A^{(\lambda)} \cap \rho'$. Since $\kappa(\lambda) \leq \kappa$, however, we have at least one *extra* bit in free(ρ') which must be set to 1, so that x has cardinality λ. The extra bits, now, in free(ρ') can be chosen to be even or odd, while still remaining in $A^{(\lambda)}\bigcap \rho'$, so that $\mathbf{INDEXPARITY}(\sigma + 2^{\ell} + \beta \cdot 2^{\ell - \lfloor \log_2 \alpha \rfloor} + 2^{\nu} \cdot x)$ is not oblivious in $A^{(\lambda)} \cap \rho'$. ■

Chapter 6

PARITY is not $\mathcal{SB}_1^{\mathbf{N}}(\mathbf{trunc})$ definable

In this chapter the main result is ,

Theorem. **PARITY** *is not cut-and-paste sharply bounded definable.*

We are adding one new *term*, the *trunc*. In [Fer88] the *trunc* term is considered via "part of" quantification, and there the proof theoretic strength of a subsystem involving "part of" quantification is considered. The system Ferreira is considering consists of some basic axioms together with induction over sharply bounded formulae in *his* language enhanced by the binary relation "x is part of y". A word on the strategy for the proof of the main result: we employ the method of probabilistic restriction, the introduction of which can be found in [FSS84]; we understand restrictions as cylinders, found in [Ajt83], of narrow enough base.

Notation and conventions

Unless changed, all previous notation and conventions apply. Furthermore:

6a. The interpretation of the $\mathrm{trunc}(x\colon y, z)$ is,

$$\mathrm{trunc}(x\colon y,z) = \begin{cases} \lfloor x/2^z \rfloor - 2^{y-z} \cdot \lfloor x/2^y \rfloor, & \text{if } y > z; \\ 0, & \text{otherwise.} \end{cases}$$

We often shall use

$$\text{trunc}(x\colon \infty, y) = \text{trunc}(x\colon |x|, y) = \lfloor x/2^y \rfloor .$$

The sharply bounded formulae resulting from adjoining the function of truncation to the language of bounded arithmetic are called, *cut-and-paste* sharply bounded formulae. Note that the term $\lfloor t/2 \rfloor$ will be eliminated, because it is a special case of *trunc*. The **BIT** predicate is not in the language. However, we shall make use of it because it too is definable from truncation.

6b. For sequences of terms or integers $\vec{w}, \vec{v}$ we shall denote by $\vec{w} \bullet \vec{v}$ the dot product $\sum_{i<\kappa} w_i \cdot v_i$.

6c. Let $\vec{f}$, $\vec{g}$, $\vec{\varphi}$, and $\vec{\psi}$ such that $\vec{f}$ and $\vec{\varphi}$, and $\vec{g}$ and $\vec{\psi}$ have the same lengths, respectively. We agree then that,

$$\vec{\varphi}^\frown \vec{\psi} \bullet \vec{f}^\frown \vec{g} = \vec{\varphi} \bullet \vec{f} + \vec{\psi} \bullet \vec{g} ,$$

where $^\frown$ is concatenation of sequences.

6d. Let A be a set of positive integers. Then,

$$\text{spread}(A) = \begin{cases} min\{a' - a \mid a' > a,\ a, a' \in A\}, & \text{if } |A| > 1; \\ \infty , & \text{else.} \end{cases}$$

6e. Recall the definition of $\mathcal{H}(d\colon x_1, \ldots, x_d)$, for suitable integer d. Let d be an integer > 0, and A a set of positive integers. We then let,

$$\mathcal{H}(d\colon A) = \bigcup_{\substack{a_1,\ldots,a_d \in A \\ \text{pairwise distinct}}} \mathcal{H}(d\colon a_1, \ldots, a_d) .$$

Note: $\mathcal{H}(d\colon A) + \mathcal{H}(d'\colon A) \subseteq \mathcal{H}(d + d'\colon A)$ and $\mathcal{H}(d\colon A) \subseteq \mathcal{H}(d'\colon A)$, for $d \leq d'$.

6f. $\text{supp}(a, b, \ldots) = \text{rng}\, \vec{a} \cup \text{rng}\, \vec{b} \cup \ldots$, for $a = \vec{i} \bullet \vec{a}$, $b = \vec{j} \bullet \vec{b}, \ldots \in \mathcal{H}(d\colon A)$, where $\text{rng}\, \vec{a}$ is the set of entries in the sequence $\vec{a}$.

6g. For x a positive integer, we let:

1. $x[\vec{a}] = \langle x(a_0), \ldots, x(a_\mu) \rangle$ for $\vec{a} = a_0 \ldots a_\mu$, μ a positive integer, and
2. $\wedge x[Y] = \prod_{a \in Y} x(a)$ for Y a set of bits .

6h. Let $\vec{x} = x_0 \dots x_{\mu-1}$ be a sequence of μ objects, μ a positive integer, and g a unary function. We then let $g(\vec{x}) = \langle g(x_0), \dots, g(x_{\mu-1})\rangle$.

6i. For $\vec{a}, \vec{b}$ two finite sequences, we define

$$\vec{a} \cup \vec{b} = \vec{a} \frown \vec{b}_{-a}$$

where $\vec{b}_{-a}$ comes from $\vec{b}$ by deleting those entries from $\vec{b}$ that are in $\operatorname{rng} \vec{a}$.

6j. We say f *is a boolean function in* μ *variables* if $f\colon \{0,1\}^\mu \longrightarrow \{0,1\}$. It is clear that with μ fixed, a boolean function f in μ variables, $\epsilon_0, \dots, \epsilon_{\mu-1}$, can be written as the sum of products of ϵ_i's or $1 - \epsilon_j$'s, with $i, j < \mu$ (disjunctive normal form). With μ a positive integer, we let BOOL_μ =all boolean functions in at most μ variables. Likewise BOOL_μ^κ, for $\kappa > 0$ an integer, will denote the set of sequences of length $\leq \kappa$, with entries in BOOL_μ.

6k. Let κ, μ be positive nonzero integers. For $\vec{f} \in \mathrm{BOOL}_\mu^\kappa$ we say that $\vec{f}$ is a *partition* if for any sequence $\epsilon_0, \dots, \epsilon_{\mu-1}$ of 0's and 1's there is at most *one* member $f \in \vec{f}$ such that $f(\epsilon_0, \dots, \epsilon_\mu) = 1$. We say $\vec{f}$ is a *partition of unity* if $\vec{f}$ is a partition and $\bigvee_{f \in \vec{f}} f \equiv 1$.

6l. The unabbreviated form of quantification,

$$\mathrm{Bool}(Q) = \begin{cases} \bigvee, & \text{if } Q = \exists\,; \\ \bigwedge, & \text{if } Q = \forall\,. \end{cases}$$

With Ψ a statement we set

$$[\![\Psi]\!] = \begin{cases} 1, & \text{if } \Psi \text{ is true;} \\ 0, & \text{if } \Psi \text{ is false.} \end{cases}$$

6m. Let $t(x, \vec{z})$ be a cut-and-paste term. We then set:

1.

$$d(t) = \begin{cases} 1, & \text{if } t \text{ is a variable or a standard constant;} \\ d(s), & \text{if } t = \lfloor s/2 \rfloor \text{ or } t = |s|; \\ \max\{d(s), d(r)\}, & \text{if } t = r + s \text{ or } t = r\#s; \\ \max\{d(s), d(l), d(r)\}, & \text{if } t = \mathrm{trunc}(s\colon l, r); \\ e(s) + e(r), & \text{if } t = r \cdot s. \end{cases}$$

2.

$$e(t) = \begin{cases} 2, & \text{if } t \text{ is a variable or a standard constant;} \\ e(s), & \text{if } t = \lfloor s/2 \rfloor \text{ or } t = |s|; \\ \max\{e(s), e(r)\}, & \text{if } t = r + s; \\ \max\{e(s), e(r)\} + 1, & \text{if } t = s \cdot r; \\ \max\{e(s), e(l), e(r)\}, & \text{if } t = \text{trunc}(s{:}\, l, r); \\ e(s) + e(r) + 1, & \text{if } t = s\#r. \end{cases}$$

3. $||t||_{\ell,m} = \max\{|t^{-trunc}(x, \vec{c})| \mid 0 \leq x < 2^\ell - 1, \vec{c} < \ell^m\}$, where ℓ, m are positive integers, and t^{-trunc} comes from t by replacing occurrences of $\text{trunc}(s{:}\, l, r)$ in t with s.

6n. Let d, e, ν, and ℓ be positive integers. We then let:

1. $A_{d,e,\nu,\ell} = \lceil \log_2 \ell \rceil^e \cdot B$, where B is the set of positive integers wide for d obtained by applying the wide lemma to d and $\sqrt[2\cdot d+1]{\ell/(2 \cdot \nu \cdot \lceil \log_2 \ell \rceil^e)}$, and
2. $\text{Restr}_\epsilon(d, e, \nu, \ell) = \{\rho \mid \rho \text{ a restriction on } \ell, \text{ with } |\text{free}(\rho)| \geq \ell^\epsilon \text{ and } x \in \rho \Longrightarrow x \subset A_{d,e,\nu,\ell}\}$.

Restrictions are sets, and hence partially ordered by "$\subseteq$". $A_{d,e,\nu,\ell}$ is wide for d and $\text{spread}(A_{d,e,\nu,\ell}) > 2 \cdot \nu \cdot \lceil \log_2 \ell \rceil^e$.

6o. Let $t(x, \vec{z})$ be a cut-and-paste term, m a positive integer. We then say $t(x, \vec{z})$ *has the bit-finite property at* m if,

$\forall\, d \geq d(t), \epsilon{:}\, 0 < \epsilon \leq 1/(2 \cdot d + 2)$;

$\exists\, \mu, \nu_0, \delta{:}\, 0 < \delta \leq \epsilon$;

$\forall\, \nu \geq \nu_0, e \geq e(t) \exists\, \ell_0 \forall\, \ell \geq \ell_0$;

$\forall\, \rho \in \text{Restr}_\epsilon(d, e, \nu, \ell) \exists\, \rho' \in \text{Restr}_\delta(d, e, \nu, \ell), \rho' \subseteq \rho$;

$\forall\, \vec{c} < \ell^m$;

$\forall\, z < ||t||_{\ell,m}\ \exists\, \vec{a}^{(\vec{c},z)}$ a sequence of no more than μ bits $< \ell$, $f_{\vec{c},y} \in \text{BOOL}_\mu$;

$\forall\, x \in \rho'$,

$$t(x, \vec{c}) = \sum_{y < ||t||_{\ell,m}} 2^y \cdot f_{\vec{c},y}(x[\vec{a}^{(\vec{c},y)}])$$

or equivalently,

$$[\![\mathbf{BIT}(t(x, \vec{c}), y)]\!] = f_{\vec{c},y}(x[\vec{a}^{(\vec{c},y)}]) \,.$$

6p. Let $t(x, \vec{z})$ be a cut-and-paste term, m a positive integer and $\mathbf{Z}$ the set of all integers. We then say $t(x, \vec{z})$ *is in cluster form at* m if,

$\forall\, d \geq d(t), \epsilon\colon 0 < \epsilon \leq 1/(2 \cdot d + 2)$;

$\exists\, \mu, \nu_0, \delta\colon 0 < \delta \leq \epsilon$;

$\forall\, \nu \geq \nu_0, e \geq e(t) \exists\, \ell_0 \forall\, \ell \geq \ell_0$;

$\forall\, \rho \in \mathrm{Restr}_\epsilon(d, e, \nu, \ell) \exists\, \rho' \in \mathrm{Restr}_\delta(d, e, \nu, \ell), \rho' \subseteq \rho$;

$\forall\, \vec{c} < \ell^m$;

$\exists\, \vec{a}^{(\vec{c})}$ sequence of no more than μ bits;

$\exists\, \vec{f}^{(\vec{c})} \in \mathrm{BOOL}_\mu^{\nu_0}$;

$\forall\, j < \nu_0$;

$\exists\, \alpha_{\vec{c},j} \in \mathbf{Z}\,, \psi_{\vec{c},j}$ positive integer;

$\exists\, A_j^{\vec{c}} \subseteq \mathcal{H}(d(t)\colon A_{d,e,\nu,\ell})$;

$\forall\, b \in A_j^{\vec{c}} \exists\, \varphi_{\vec{c},j,b}$ positive integer;

such that:

1. $|\psi_{\vec{c},j} \cdot \varphi_{\vec{c},j,b}| \leq \lceil \log_2 \ell \rceil^{e(t)}$, and
2. $\forall\, x \in \rho'$,

$$t(x, \vec{c}) = \sum_{j < \nu_0} 2^{\alpha_{\vec{c},j}} \cdot \psi_{\vec{c},j} \cdot f_j^{\vec{c}}(x[\vec{a}^{(\vec{c})}]) \cdot \sum_{a \in A_j^{\vec{c}}} \varphi_{\vec{c},j,a} \cdot \wedge x[\mathrm{supp}(a)] \cdot 2^a.$$

6q. The following two conventions are used throughout:

1. A sum taken over the *empty set* will always equal to 1, contrary to the common practice of putting such a sum equal to 0.
2. $\emptyset + A = A$, for A a set of numbers.

The lemmata

This lemma shows why we want to prove the cluster lemma. It is used in the proofs of the cluster length and cluster smash sublemmata, and the boolean cut-and-paste lemma.

Cluster-to-finite Lemma 6.1. *Let m be a positive integer. Then, any cut-and-paste term $t(x, \vec{z})$ in cluster form at m has the bit-finite property at m.*

Proof. Fix ϵ with $0 < \epsilon \leq 1/(2 \cdot d + 2)$, $d \geq d(t)$. Get ν_0, μ, δ with $0 < \delta \leq \epsilon$. Fix $\nu \geq \nu_0$, $e \geq e(t)$, get ℓ_0, and fix $\ell \geq \ell_0$, $\rho \in \mathrm{Restr}_\epsilon(d, e, \nu, \ell)$ and get $\rho' \in \mathrm{Restr}_\delta(d, e, \nu, \ell)$, $\rho' \subseteq \rho$ such that for each $\vec{c} < \ell^m$ and $x \in \rho'$ we have that

$$t(x, \vec{c}) = \sum_{j < \nu_0} 2^{\alpha_{\vec{c},j}} \cdot \psi_{\vec{c},j} \cdot f_j^{\vec{c}}(x[\vec{a}^{(\vec{c})}]) \cdot \sum_{a \in A_j^{\vec{c}}} \varphi_{\vec{c},j,a} \cdot \wedge x[\mathrm{supp}(a)] \cdot 2^a \ .$$

We now define "cluster". Consider $G = \cup_{j<\nu_0}(\alpha_{\vec{c},j} + A_j^{\vec{c}}) \times \{j\}$. The following order defined on G is a strict linear order: for $(a + \alpha_{\vec{c},j}, j), (a' + \alpha_{\vec{c},j'}, j') \in G$, $(a + \alpha_{\vec{c},j}, j) \prec (a' + \alpha_{\vec{c},j'}, j')$ if $a + \alpha_{\vec{c}} < a' + \alpha_{\vec{c},j'}$ or $j < j'$. Call an $\prec$-interval of G, $\vec{g} = \langle g_0, \ldots, g_{\nu'-1} \rangle$, *narrow* if for $i < \nu' - 2$ we have that $d(g_{i+1}, g_i) \leq 2 \cdot \lceil \log_2 \ell \rceil^e$, where $d(g, g') = |a + \alpha_{\vec{c},j} - (a' + \alpha_{\vec{c},j'})|$, $g = (a + \alpha_{\vec{c},j}, j)$, $g' = (a' + \alpha_{\vec{c},j'}, j')$ some $j, j' < \nu_0$. An $\prec$-interval of G, $\vec{g}$, is called *precluster* if $\vec{g}$ is a maximal narrow $\prec$-interval. Preclusters are linearly ordered in the obvious way. Finally, we call the sum,

$$t_{\vec{g}}(x) = \sum_{(a+\alpha_{\vec{c},j},j) \in \vec{g}} 2^{a+\alpha_{\vec{c},j}} \cdot \psi_{\vec{c},j} \cdot f_j^{\vec{c}}(x[\vec{a}^{(\vec{c})}]) \cdot \varphi_{\vec{c},j,a} \cdot \wedge x[\mathrm{supp}(a)]$$

the *cluster based on* $\vec{g}$ with $\vec{g}$ a precluster. Clearly,

$$t(x, \vec{c}) = \sum_{\vec{g} \text{ a precluster}} t_{\vec{g}}(x).$$

Let $\vec{g} = g_0 \ldots g_{\nu'-1}$ be a precluster in G. We want to show that $\nu' \leq \nu_0$. Assume that $\nu' \geq \nu_0$. Recall that $\vec{g}$ consists of pairs. We claim that the first ν_0 members in $\vec{g}$ have distinct second entries: since $\vec{g}$ is a precluster, it follows that $d(g_{\nu_0-1}, g_0) \leq 2 \cdot (\nu_0 - 1) \cdot \lceil \log_2 \ell \rceil^e$; if $i < i' < \nu_0$, $g_i = (a + \alpha_{\vec{c},j}, j)$ and $g_{i'} = (a' + \alpha_{\vec{c},j}, j)$ then $a + \alpha_{\vec{c},j} < a' + \alpha_{\vec{c},j}$ because $g_i \prec g_{i'}$; but then,

$$\mathrm{spread}(A_j^{\vec{c}}) \leq a' - a = a' + \alpha_{\vec{c},j} - (a + \alpha_{\vec{c},j}) \leq d(g_{\nu_0-1}, g_0) \leq 2 \cdot (\nu_0 - 1) \cdot \lceil \log_2 \ell \rceil^e \ ,$$

a contradiction. If $\nu_0' > \nu_0$ then $g_{\nu_0} = (a + \alpha_{\vec{c},j'}, j')$, where the j' appears earlier in the precluster because there are at most ν_0 sets $A_j^{\vec{c}}$. Again, $d(g_{\nu_0}, g_0) \leq 2 \cdot \nu_0 \cdot \lceil \log_2 \ell \rceil^e$. The same argument as above leads to a contradiction. It is now clear that distinct clusters have no bit-interference. Let $\vec{g}$ be a precluster. Denote by $\vec{g}^+$ the next precluster after $\vec{g}$, if it exists. The *range* of the cluster based on $\vec{g}$ is the interval $[a_0 + \alpha_{\vec{c},j}, a_0^+ + \alpha^+_{\vec{c},j^+}) \bigcap \mathbf{N}$, where $g_0 = (a_0 + \alpha_{\vec{c},j}, j)$ and $g_0^+ = (a_0^+ + \alpha^+_{\vec{c},j^+}, j^+)$ for some $j, j^+ < \nu_0$. When $\vec{g}$ is the leftmost

precluster the range is the interval $[a_0 + \alpha_{\vec{c},j}, \infty) \bigcap \mathbf{N}$. Fix $y < \|t\|_{\ell,m}$. Clearly, there is a unique cluster with range that contains y. Let

$$\sum_{(a+\alpha_{\vec{c},j},j)\in\vec{g}} 2^{a+\alpha_{\vec{c},j}} \cdot \psi_{\vec{c},j} \cdot f_j^{\vec{c}}(x[\vec{a}^{(\vec{c})}]) \cdot \varphi_{\vec{c},j,a} \cdot \wedge x[\mathrm{supp}(a)]$$

be that (unique) cluster of which the range contains y. Let $\vec{a}^{(\vec{c},y)} = \bigcup_{(a+\alpha_{\vec{c},j},j)\in\vec{g}} \vec{a}^{(\vec{c},j,a)}$ and note that $\vec{a}^{(\vec{c},y)}$ has at most $\nu_0 \cdot \mu$ entries. We say x *is* $(t{:}\,\xi)$*-good* for $\xi \in \{0,1\}^{\nu_{t,0}}$, $\nu_{t,0} = \nu_0$ iff

$$1 = [\![\bigwedge_{j<\nu_{t,0}} [\xi(j) = f_j^{\vec{c}}(x[\vec{a}^{(\vec{c})}])]]\!].$$

It is now easy to see that,

$$\mathbf{BIT}(t(x,\vec{c}),y) \Longleftrightarrow \bigvee_{\xi\in\{0,1\}^{\nu_{t,0}}} [(x \text{ is } (t{:}\,\xi)\text{-good}) \wedge \wedge \mathbf{BIT}(\sum_{(a+\alpha_{\vec{c},j},j)\in\vec{g}} 2^{a+\alpha_{\vec{c},j}} \cdot \psi_{\vec{c},j} \cdot \xi(j) \cdot \varphi_{\vec{c},j,a} \cdot \wedge x([\mathrm{supp}(a)]), y)].$$

□

Cut-and-paste Cluster Lemma 6.2. *Let m be a positive integer. Then, all cut-and-paste terms $t(x,\vec{z})$ are in cluster form at m.*

Proof. By induction on the complexity of the cut-and-paste term $t(x,\vec{z})$. All cases are treated by the sublemmata below:

Base Cases Sublemma. *Let m be a positive integer. If $t(x,\vec{z}) = z_i$, $i = 1,\ldots,k$, $t(x,\vec{z}) = c$ a constant, or $t(x,\vec{z}) = x$ then $t(x,\vec{z})$ is in cluster form at m.*

Proof. In all of the above cases $d(t) = 1$ and $e(t) = 2$. So, fix $d \geq 1$, ϵ with $0 < \epsilon \leq 1/(d+1)$. Set $\mu = 0, \nu_0 = 1, \delta = \epsilon$. Fix $e, \nu \geq 1$. Find $\ell_0 > 0$ such that $\lceil \log_2 \ell \rceil^2 > m \cdot \lceil \log_2 \ell \rceil$ for any $\ell \geq \ell_0$. Fix $\ell \geq \ell_0$ and $\rho \in \mathrm{Restr}_\epsilon(d,e,\nu,\ell)$. Set $\rho' = \rho$. For $\vec{c} < \ell^m$ we set:

1. $\vec{a}^{(\vec{c})} = \emptyset$ and $\vec{f}^{(\vec{c})} = \langle 1 \rangle$;
2. $\alpha_{\vec{c},0} = 0$ and $\psi_{\vec{c},0} = \begin{cases} c_i, & \text{if } t(x,\vec{z}) = z_i; \\ c, & \text{if } t(x,\vec{z}) = c; \\ 1, & \text{if } t(x,\vec{z}) = x. \end{cases}$;

3. $A_0^{\vec{c}} = \begin{cases} \emptyset, & \text{if } t(x,\vec{z}) = z_i \text{ or } t(x,\vec{z}) = c; \\ A_{d,e,\nu,\ell}, & \text{if } t(x,\vec{z}) = x. \end{cases}$ and $\varphi_{\vec{c},0,a} = 1$ for any $a \in A_0^{\vec{c}}$.

After the appropriate substitutions and observing the convention $\sum \emptyset = 1$ we see that for any $x \in \rho'$,

$$t(x,\vec{c}) = 2^{\alpha_{\vec{c},0}} \cdot \psi_{\vec{c},0} \cdot f_0^{\vec{c}}(x[\vec{a}^{(\vec{c})}]) \cdot \sum_{a \in A_0^{\vec{c}}} \varphi_{\vec{c},0,a} \cdot \wedge x[\mathrm{supp}(a)] \cdot 2^a.$$

Finally, for any $b \in A_0^{\vec{c}}$, $|\psi_{\vec{c},0} \cdot \varphi_{\vec{c},0,b}| \leq m \cdot \lceil \log_2 \ell \rceil < \lceil \log_2 \ell \rceil^{e(t)}$, by the choice of ℓ_0.□

Cluster Length Sublemma. *Let m be a positive integer, and $t(x,\vec{z})$ a cut-and-paste term. If $t(x,\vec{z})$ is in cluster form at m then,*

$\forall\, d \geq d(t), \epsilon\colon 0 < \epsilon \leq 1/(2 \cdot d + 2)$;

$\exists\, \mu, \nu_0, \delta\colon 0 < \delta \leq \epsilon$;

$\forall\, \nu \geq \nu_0, e \geq e(t) \exists\, \ell_0 \forall\, \ell \geq \ell_0$;

$\forall \rho \in \mathrm{Restr}_\epsilon(d,e,\nu,\ell) \exists \rho' \in \mathrm{Restr}_\delta(d,e,\nu,\ell), \rho' \subseteq \rho$;

$\forall \vec{c} < \ell^m$;

$\exists \vec{f}^{(\vec{c})} \in \mathrm{BOOL}_\mu^{\nu_0}$, *partition of unity;*

$\exists \vec{a}^{(\vec{c})}$ *a sequence of no more than μ bits $< \ell$;*

$\exists \vec{\varphi}^{(\vec{c})}$ *sequence of no more than κ positive integers $< \|t\|_{\ell,m}$;*

$\forall x \in \rho'$,

$$|t(x,\vec{c})| = \vec{\varphi}^{(\vec{c})} \bullet \vec{f}^{(\vec{c})}(x[\vec{a}^{(\vec{c})}]) = \sum_{j < \nu_0} \varphi_j^{\vec{c}} \cdot f_j^{\vec{c}}(x[\vec{a}^{(\vec{c})}]).$$

Hence if $t(x,\vec{z})$ is in cluster form at m, then $|t(x,\vec{c})|$ is in cluster form at m.

Proof. Since $t(x,\vec{z})$ is in cluster form at m, it has the bit-finite property at m. Fix ϵ with $0 < \epsilon \leq 1/(2 \cdot d + 2)$, $d \geq d(t)$. Get ν_0, μ, δ with $0 < \delta \leq \epsilon$. Fix $\nu \geq \nu_0$, $e \geq e(t)$, get ℓ_0, fix $\ell \geq \ell_0$, $\rho \in \mathrm{Restr}_\epsilon(d,e,\nu,\ell)$ and get $\rho' \in \mathrm{Restr}_\delta(d,e,\nu,\ell)$, $\rho' \subseteq \rho$ such that for each $\vec{c} < \ell^m$ and $y < \|t\|_{\ell,m}$, $\mathbf{BIT}(t(x,\vec{c}),y)$ is equivalent to a circuit of at most μ inputs. For each $\vec{c} < \ell^m$ and $y < \|t\|_{\ell,m}$ define,

$$\mathrm{boolean}(|t(x,\vec{c})| = y+1) = \mathbf{BIT}(t(x,\vec{c}),y) \cdot \bigwedge_{\substack{y' < \|t\|_{\ell,m} \\ y < y'}} \neg \mathbf{BIT}(t(x,\vec{c}),y').$$

We identify the predicate boolean$(|t(x,\vec{c})| = y+1)$ with the circuit that calculates its truth value. Clearly then,

$$|t(x,\vec{c})| = \sum_{y<\|t\|_{\ell,m}} (y+1)\cdot[\![\text{boolean}(|t(x,\vec{c})| = y+1)]\!].$$

For $\vec{c} < \ell^m$ the family $\{\text{boolean}(|t(x,\vec{c})| = y+1) \mid y < \|t\|_{\ell,m}\}$ is a partition of unity. Consider the multicircuit $\mathbf{C}$ consisting of,

$$\mathbf{C}_\ell = \{\text{boolean}(|t(x,\vec{c})| = y+1) \mid \vec{c} < \ell^m, y < \|t\|_{\ell,m}\} .$$

$\mathbf{C}$ has constant depth, polynomial size and width. By an easy variant of the multicircuit-Odd-Even lemma there are ℓ_1, μ', and δ' with $0 < \delta' \leq \delta$ such that for $\ell \geq \ell_1$,
$\forall \tau \in \text{Restr}_\delta(d,e,\nu,\ell) \exists \tau' \in \text{Restr}_{\delta'}(d,e,\nu,\ell), \tau' \subseteq \tau$;
$\forall \vec{c} < \ell^m$, $y < \|t\|_{\ell,m}$;
$\exists f_y^{\vec{c}} \in \text{BOOL}_{\mu'}$;
$\exists \vec{a}^{(\vec{c},y)}$ a sequence of no more than μ' bits $\in \text{free}(\tau')$;
$\forall x \in \tau'$,

$$\text{boolean}(|t(x,\vec{c})| = y+1) = f_y^{\vec{c}}(x[\vec{a}^{(\vec{c},y)}]).$$

Using the above get a restriction $\rho'' \in \text{Restr}_{\delta'}(d,e,\nu,\ell)$, with $\rho'' \subseteq \rho'$, $f_y^{\vec{c}} \in \text{BOOL}_{\mu'}$, and sequences $\vec{a}^{(\vec{c},y)}$ of $\leq \mu'$ bits $< \ell$ such that for any $\vec{c} < \ell^m$ and any $x \in \rho''$,

$$\text{boolean}(|t(x,\vec{c})| = y+1) = f_y^{\vec{c}}(x[\vec{a}^{(\vec{c},y)}]) .$$

μ' depends on m, d, μ, and δ, i.e. the parameters for the size of the multicircuit, and the depth of the multicircuit which is < 5. Fix $\vec{c} < \ell^m$ and consider the set of functions,

$$\mathbf{F}_{\vec{c}} = \{f_y^{\vec{c}} \mid y < \|t\|_{\ell,m} \text{ and } f_y^{\vec{c}} \not\equiv 0\} .$$

We claim two things:

1. that $\mathbf{F}_{\vec{c}}$ has at most $2^{\mu'}$ members, and
2. that $\vec{a}^{(\vec{c})} = \bigcup_{y: f_y^{\vec{c}} \in \mathbf{F}_{\vec{c}}} \vec{a}^{(\vec{c},y)}$ has at most $\mu' \cdot 2^{\mu'}$ bits.

First we prove (1). Fix $f = f_y^{\vec{c}} \in \mathbf{F}_{\vec{c}}$, where $\vec{a} = \vec{a}^{(\vec{c},y)}$ are the bits on which f is applied. We identify f with the set $\{x \in \rho'' \mid f(x[\vec{a}]) = 1\}$. As a boolean function $f = \bigvee_{i<\gamma_f} f_i$

(disjunctive normal form) and as a set $f = \bigcup_{i<\gamma_f} f_i$, where $\gamma_f < 2^{\mu'}$ and the f_i are pairwise disjoint cylinders with a base of size at most μ'. Now,

$$\rho'' = \bigcup_{f\in\mathbf{F}_{\vec{c}}}\bigcup_{i<\gamma_f} f_i \text{ and } |\rho''| = \sum_{f\in\mathbf{F}_{\vec{c}}}\sum_{i<\gamma_f} |f_i|$$

the second equality following from the fact that $\mathbf{F}_{\vec{c}}$ is a partition of unity. $|\rho''| = 2^{\lambda}$, where $\lambda = |\text{free}(\rho'')|$. For $f \in \mathbf{F}_{\vec{c}}$, $|f_i| \geq 2^{\lambda-\mu'}$ for $i < \gamma_f$. So, $2^{\lambda} = |\rho''| \geq \sum_{f\in\mathbf{F}_{\vec{c}}} \gamma_f \cdot 2^{\lambda-\mu'}$ from which it follows that $2^{\mu'} \geq \sum_{f\in\mathbf{F}_{\vec{c}}} \gamma_f$. Clealry, now (1) is established. (2) is trivial, except that the known lower bound to me is $bound(\mu') \geq 1 + 2 \cdot bound(\mu' - 1)$, and clearly this is disappointingly exponential, though it hardly matters here. In the *statement* of this lemma now, replace the μ by $\hat{\mu} = \mu' \cdot 2^{\mu'}$ and ν_0 by $\hat{\nu}_0 = 2^{\mu'}$. For $\vec{c} < \ell^m$ then and $x \in \rho''$ we have that

$$|t(x,\vec{c})| = \vec{\varphi}^{(\vec{c})} \bullet \vec{f}^{(\vec{c})},$$

where $\vec{\varphi}^{(\vec{c})} = \langle y \mid f_y^{\vec{c}} \in \mathbf{F}_{\vec{c}}\rangle$ and $\vec{f}^{(\vec{c})} = \langle f_y^{\vec{c}} \mid f_y^{\vec{c}} \in \mathbf{F}_{\vec{c}}\rangle$. Note that both $\hat{\nu}_0$ and $\hat{\mu}$ only depend on μ, ν_0 from the induction hypothesis on t, ϵ, and d, but not on e nor on ν, because they do not influence the size of free(ρ').□

Cluster Smash Sublemma. *Let m be a positive integer, and $s(x,\vec{z})$, $t(x,\vec{z})$ two cut-and-paste terms in cluster form at m. Then,*

$\forall\, d \geq d(t\#s), \epsilon\colon 0 < \epsilon \leq 1/(2\cdot d+2)$;

$\exists\, \mu, \nu_0, \delta\colon 0 < \delta \leq \epsilon$;

$\forall\, \nu \geq \nu_0, e \geq e(t\#s) \exists\, \ell_0 \forall\, \ell \geq \ell_0$;

$\forall \rho \in \text{Restr}_{\epsilon}(d,e,\nu,\ell) \exists \rho' \in \text{Restr}_{\delta}(d,e,\nu,\ell), \rho' \subseteq \rho$;

$\forall \vec{c} < \ell^m$;

$\exists \vec{f}^{(\vec{c})} \in \text{BOOL}_{\mu}^{\nu_0}$, *partition of unity;*

$\exists \vec{a}^{(\vec{c})}$ *a sequence of no more than μ bits $< \ell$;*

$\exists \vec{\varphi}^{(\vec{c})}$ *sequence of no more than κ positive integers $< \|t\|_{\ell,m}$;*

$\forall x \in \rho'$,

$$s(x,\vec{z})\#t(x,\vec{c}) = \vec{f}^{(\vec{c})}(x[\vec{a}^{(\vec{c})}]) \bullet 2^{\vec{\varphi}^{(\vec{c})}}.$$

Hence if $s(x,\vec{z})$ and $t(x,\vec{z})$ are in cluster form at m, then $r(x,\vec{z}) = s(x,\vec{z})\#t(x,\vec{z})$ is in cluster form at m.

Proof. The parameters to follow are from the cluster length lemma applied to s and t. Fix ϵ with $0 < \epsilon \leq 1/(2 \cdot d + 2)$, $d \geq \max\{d(t), d(s)\}$. Get $\nu_{t,0}$, μ_t, δ with $0 < \delta \leq \epsilon$ such that the property after the first $\forall\exists$ in the statement of the cluster length sublemma on t holds. Now, with δ in place of ϵ in the first $\forall\exists$ of the statement of the cluster length sublemma for s, get $\nu_{s,0}$, μ_s, and δ', $0 < \delta' \leq \delta$ so that the property after the first holds $\forall\exists$ for s. Fix $\nu \geq \max\{\nu_{t,0}, \nu_{s,0}\}$, $e \geq \max\{e(t), e(s)\}$, get ℓ_0 for t, ℓ_1 for s and fix $\ell \geq \max\{\ell_0, \ell_1\}$, $\rho \in \mathrm{Restr}_\epsilon(d, e, \nu, \ell)$. First get $\rho' \in \mathrm{Restr}_\delta(d, e, \nu, \ell)$, $\rho' \subseteq \rho$ such that for $\vec{c} < \ell^m$ and $x \in \rho'$ we have that

$$|t(x, \vec{c})| = \vec{\varphi}^{(\vec{c},t)} \bullet \vec{f}^{(\vec{c},t)}$$

and then get $\rho'' \in \mathrm{Restr}_{\delta'}(d, e, \nu, \ell)$, $\rho'' \subseteq \rho'$ such that for $\vec{c} < \ell^m$ and $x \in \rho''$ we have that

$$|s(x, \vec{c})| = \vec{\varphi}^{(\vec{c},s)} \bullet \vec{f}^{(\vec{c},s)} .$$

Both of the above equations hold for $\vec{c} < \ell^m$ and $x \in \rho''$. So,

$$|t(x, \vec{c})| \cdot |s(x, \vec{c})| = \sum_{\substack{i < \nu_{s,0} \\ j < \nu_{t,0}}} \varphi^{(\vec{c})}_{<i,j>} \cdot f^{(\vec{c})}_{<i,j>} = \vec{\varphi}^{(\vec{c})} \bullet \vec{f}^{(\vec{c})} ,$$

where $\varphi^{(\vec{c})}_{<i,j>} = \varphi^{(\vec{c},s)}_i \cdot \varphi^{(\vec{c},t)}_j$ and $f^{(\vec{c})}_{<i,j>} = f^{(\vec{c},s)}_i \cdot f^{(\vec{c},t)}_j$, for $i < \nu_{s,0}$ and $j < \nu_{t,0}$. Furthermore:

1. $\vec{f}^{(\vec{c})} \in BOOL^{\nu_{s,0} \cdot \nu_{t,0}}_{\mu_s + \mu_t}$, because boolean functions are closed under multiplication, and
2. $\vec{f}^{(\vec{c})}$ is a partition of unity in $\vec{a}^{(\vec{c})} = \vec{a}^{(\vec{c},s)} \cup \vec{a}^{(\vec{c},t)}$, because the superimposition of partitions is again a partition and because the superimposition of coverings is again a covering.

By (1) and (2) above, we conclude that $2^{\vec{\varphi}^{(\vec{c})} \bullet \vec{f}^{(\vec{c})}} = 2^{\vec{\varphi}^{(\vec{c})}} \bullet \vec{f}^{(\vec{c})}$. Finally, put $\nu_0 = \nu_{t,0} \cdot \nu_{s,0}$, and $\mu = \mu_s + \mu_t$.□

Cluster + Sublemma. *Let m be a positive integer, and $s(x, \vec{z})$, $t(x, \vec{z})$ two cut-and-paste terms in cluster form at m. Then, $r(x, \vec{z}) = s(x, \vec{z}) + t(x, \vec{z})$ is also in cluster form at m.*

Proof. Whenever we have two or more cut-and-paste terms we shall resort to the first 4 lines of definitions such as bit-finite and cluster form at m, in order to get a suitable restriction in which the terms have the claimed properties simultaneously. It is for this

reason that the mentioned definitions are headed by these four lines. What follows is typical of the situation when two terms are involved. It will serve as a blueprint for all the later cases. Fix ϵ with $0 < \epsilon \leq 1/(2 \cdot d + 2)$, $d \geq d(t+s)$. Get $\nu_{t,0}$, μ_t and δ with $0 < \delta \leq \epsilon$ such that the property after the first $\forall\exists$ in the statement of cluster form at m holds for t. Now, with δ in place of ϵ in the first $\forall\exists$ of the statement of cluster form at m for s, get $\nu_{s,0}$, μ_s and δ', $0 < \delta' \leq \delta$ so that the property after the first alternation holds for s. Fix $\nu \geq \max\{\nu_{t,0}, \nu_{s,0}\}$, $e \geq e(t+s)$, get ℓ_0 for t, ℓ_1 for s and fix $\ell \geq \max\{\ell_0, \ell_1\}$, $\rho \in \mathrm{Restr}_\epsilon(d, e, \nu, \ell)$. First get $\rho' \in \mathrm{Restr}_\delta(d, e, \nu, \ell)$, $\rho' \subseteq \rho$ such that for $\vec{c} < \ell^m$ we have that

$\exists\, \vec{a}^{(s,\vec{c})}$ sequence of no more than μ_s bits;

$\exists\, \vec{f}^{(s,\vec{c})} \in \mathrm{BOOL}_{\mu_s}^{\nu_{s,0}}$;

$\forall\, j < \nu_{s,0}$;

$\exists\, \alpha_{s,\vec{c},j} \in \mathbf{Z}$, $\psi_{s,\vec{c},j}$ positive integer;

$\exists\, A_j^{s,\vec{c}} \subseteq \mathcal{H}(d(s)\colon A_{d,e,\nu,\ell})$;

$\forall\, b \in A_j^{s,\vec{c}}\ \exists\, \varphi_{s,\vec{c},j,b}$ positive integer;

such that:

1. $|\psi_{s,\vec{c},j} \cdot \varphi_{s,\vec{c},j,b}| \leq \lceil \log_2 \ell \rceil^{e(s)}$, and

2. $\forall\, x \in \rho'$,

$$s(x, \vec{c}) = \sum_{j < \nu_{s,0}} 2^{\alpha_{s,\vec{c},j}} \cdot \psi_{s,\vec{c},j} \cdot f_j^{s,\vec{c}}(x[\vec{a}^{(s,\vec{c})}]) \cdot \sum_{a \in A_j^{s,\vec{c}}} \varphi_{s,\vec{c},j,a} \cdot \wedge x[\mathrm{supp}(a)] \cdot 2^a$$

and then get $\rho'' \in \mathrm{Restr}_{\delta'}(d, e, \nu, \ell)$, $\rho'' \subseteq \rho'$ such that for $\vec{c} < \ell^m$ we have that

$\exists\, \vec{a}^{(t,\vec{c})}$ sequence of no more than μ_t bits;

$\exists\, \vec{f}^{(t,\vec{c})} \in \mathrm{BOOL}_{\mu_t}^{\nu_{t,0}}$;

$\forall\, j < \nu_{t,0}$;

$\exists\, \alpha_{t,\vec{c},j} \in \mathbf{Z}$, $\psi_{t,\vec{c},j}$ positive integer;

$\exists\, A_j^{t,\vec{c}} \subseteq \mathcal{H}(d(t)\colon A_{d,e,\nu,\ell})$;

$\forall\, b \in A_j^{t,\vec{c}}\ \exists\, \varphi_{t,\vec{c},j,b}$ positive integer;

such that:

1. $|\psi_{t,\vec{c},j} \cdot \varphi_{t,\vec{c},j,b}| \leq \lceil \log_2 \ell \rceil^{e(t)}$, and

2. $\forall\, x \in \rho''$,

$$t(x,\vec{c}) = \sum_{j<\nu_{t,0}} 2^{\alpha_{t,\vec{c},j}} \cdot \psi_{t,\vec{c},j} \cdot f_j^{t,\vec{c}}(x[\vec{a}^{(t,\vec{c})}]) \cdot \sum_{a\in A_j^{t,\vec{c}}} \varphi_{t,\vec{c},j,a} \cdot \wedge x[\mathrm{supp}(a)] \cdot 2^a.$$

Since both of the above hold in the finer restriction ρ'' we have that

$$s(x,\vec{c}) + t(x,\vec{c}) = \sum_{j<\nu_{s,0}} 2^{\alpha_{s,\vec{c},j}} \cdot \psi_{s,\vec{c},j} \cdot f_j^{s,\vec{c}}(x[\vec{a}^{(s,\vec{c})}]) \cdot \sum_{a\in A_j^{s,\vec{c}}} \varphi_{s,\vec{c},j,a} \cdot \wedge x[\mathrm{supp}(a)] \cdot 2^a +$$
$$+ \sum_{j<\nu_{t,0}} 2^{\alpha_{t,\vec{c},j}} \cdot \psi_{t,\vec{c},j} \cdot f_j^{t,\vec{c}}(x[\vec{a}^{(t,\vec{c})}]) \cdot \sum_{a\in A_j^{t,\vec{c}}} \varphi_{t,\vec{c},j,a} \cdot \wedge x[\mathrm{supp}(a)] \cdot 2^a$$

holds for $\vec{c} < \ell^m$ and $x \in \rho''$. Put $\nu_0 = \nu_{s,0} + \nu_{t,0}$, $\mu = \max\{\mu_s, \mu_t\}$. It is now clear that for $\vec{c} < \ell^m$ and $x \in \rho''$ we have that

$$r(x,\vec{c}) = \sum_{j<\nu_0} 2^{\alpha_{\vec{c},j}} \cdot \psi_{\vec{c},j} \cdot f_j^{\vec{c}}(x[\vec{a}^{(\vec{c})}]) \cdot \sum_{a\in A_j^{\vec{c}}} \varphi_{\vec{c},j,a} \cdot \wedge x[\mathrm{supp}(a)] \cdot 2^a$$

with:

1. $A_j^{\vec{c}}, j < \nu_0$ running, with possible repetitions, over all $A_j^{\vec{c},s}, j < \nu_{0,s}$ and $A_j^{\vec{c},t}, j < \nu_{0,t}$, and
2. $\alpha_j^{\vec{c}}, j < \nu_0$ running through all of $\alpha_j^{\vec{c},s}, j < \nu_{0,s}$ and $\alpha_j^{\vec{c},t}, j < \nu_{0,t}$.

□

Cluster × Sublemma. *Let m be a positive integer, and $s(x,\vec{z})$, $t(x,\vec{z})$ two cut-and-paste terms in cluster form at m. Then, $r(x,\vec{z}) = s(x,\vec{z}) \cdot t(x,\vec{z})$ is also in cluster form at m.*

Proof. Recall that $e(t \cdot s) = \max\{e(t), e(s)\} + 1$ and $d(t \cdot s) = d(t) + d(s)$. Proceed as in the blueprint of the cluster + sublemma. We get appropriate parameters, restriction ρ'', and expressions for s and t so that for $\vec{c} < \ell^m$ and $x \in \rho''$,

$$t(x,\vec{c}) \cdot s(x,\vec{c}) = \sum_{\langle i,j\rangle \in \nu_{t,0}\times\nu_{s,0}} 2^{\alpha_{\langle i,j\rangle}} \cdot \psi_{\langle i,j\rangle} \cdot f_{\langle i,j\rangle}(x[\vec{a}]) \cdot \sum_{a\in A_{\langle i,j\rangle}} \varphi_{\langle i,j\rangle,a} \cdot 2^a,$$

where (recalling the convention $\emptyset + A = A$):

1. $A_{\langle i,j\rangle} = A_i^{t,\vec{c}} + A_j^{s,\vec{c}}$, $\vec{a} = \vec{a}^{(t,\vec{c})} \cup \vec{a}^{(s,\vec{c})}$, $\psi_{\langle i,j\rangle} = \psi_{t,\vec{c},i} \cdot \psi_{s,\vec{c},j}$;
2. $\varphi_{\langle i,j\rangle,a} = \sum\{\varphi_{t,\vec{c},i,b} \cdot \varphi_{s,\vec{c},j,a} \mid b \in A_i^{t,\vec{c}}, b' \in A_j^{s,\vec{c}}, a = b + b'\}$, and

3. $f_{\langle i,j\rangle} = f_i^{t,\vec{c}} \cdot f_j^{s,\vec{c}}$.

We now claim that,

$$|\psi_{\langle i,j\rangle} \cdot \varphi_{\langle i,j\rangle,a}| \leq \lceil \log_2 \ell \rceil^{\max\{e(t),e(s)\}+1}.$$

It suffices to show that the bound on the cardinality of

$$\overset{\circ}{a} = \{(b,b') \mid \ b \in A_i^{t,\vec{c}}, b' \in A_j^{s,\vec{c}}, a = b + b'\}$$

depends only on $d(r)$ $(= d(s) + d(t))$. Now,

$$A_i^{t,\vec{c}} + A_j^{s,\vec{c}} \subseteq \mathcal{H}(d(s) + d(t)\colon A_{d,e,\nu,\ell}).$$

So, if $(b,b') \in \overset{\circ}{a}$ then $\mathrm{supp}(b,b') = \mathrm{supp}(a)$. $|\mathrm{supp}(a)| \leq d(r)$ and the coeficients in b sum to $\leq d(t)$, and the coefficients in b' sum to $\leq d(s)$. The claim on the bound of the cardinality of the set $\overset{\circ}{a}$ is now evident.□

Cluster Truncate Sublemma. *Let m be a positive integer, and $t(x,\vec{z})$, $l(x,\vec{z})$, $r(x,\vec{z})$ three cut-and-paste terms. Suppose $t(x,\vec{z})$ is in cluster form at m. Then,*

$$s(x,\vec{z}) = \mathrm{trunc}(t(x,\vec{z})\colon l(x,\vec{z}), r(x,\vec{z})$$

is also in cluster form at m.

Proof. We are going to split the proof into two cases: $\mathrm{trunc}(t\colon s, 0)$ and $\mathrm{trunc}(t\colon \infty, s) = \mathrm{trunc}(t\colon ||t||_{m,\ell}, s)$. We need some preliminary work and notation before we tackle the cases separately. Proceed as in the blueprint of the cluster + sublemma and use the cluster length sublemma to get appropriate parameters and a restriction ρ'' such that for $\vec{c} < \ell^m$ we have that

$\exists\, \vec{a}^{(s,\vec{c})}$ sequence of no more than μ_s bits;

$\exists\, \vec{f}^{(s,\vec{c})} \in \mathrm{BOOL}_{\mu_s}^{\nu_{s,0}}$;

$\forall\, j < \nu_{s,0}$;

$\exists\, \alpha_{s,\vec{c},j} \in \mathbf{Z}$, $\psi_{s,\vec{c},j}$ positive integer;

$\exists\, A_j^{s,\vec{c}} \subseteq \mathcal{H}(d(s)\colon A_{d,e,\nu,\ell})$;

$\forall\, b \in A_j^{s,\vec{c}}\ \exists\, \varphi_{s,\vec{c},j,b}$ positive integer;

such that:

1. $|\psi_{s,\vec{c},j} \cdot \varphi_{s,\vec{c},j,b}| \leq \lceil \log_2 \ell \rceil^{e(s)}$, and

2. $\forall\, x \in \rho''$,

$$s(x,\vec{c}) = \sum_{j<\nu_{s,0}} 2^{\alpha_{s,\vec{c},j}} \cdot \psi_{s,\vec{c},j} \cdot f_j^{s,\vec{c}}(x[\vec{a}^{(s,\vec{c})}]) \cdot \sum_{a \in A_j^{s,\vec{c}}} \varphi_{s,\vec{c},j,a} \cdot \wedge x[\mathrm{supp}(a)] \cdot 2^a$$

and

$$|s(x,\vec{c})| = \vec{\chi}^{(\vec{c})} \bullet \vec{h}^{(\vec{c})}(x[\vec{a}^{(s,\vec{c})}]) \quad \text{from the cluster length sublemma,}$$

and

$\exists\, \vec{a}^{(t,\vec{c})}$ sequence of no more than μ_t bits;

$\exists\, \vec{f}^{(t,\vec{c})} \in \mathrm{BOOL}_{\mu_t}^{\nu_{t,0}}$;

$\forall\, j < \nu_{t,0}$;

$\exists\, \alpha_{t,\vec{c},j} \in \mathbf{Z}$, $\psi_{t,\vec{c},j}$ positive integer;

$\exists\, A_j^{t,\vec{c}} \subseteq \mathcal{H}(d(t)\colon A_{d,e,\nu,\ell})$;

$\forall\, b \in A_j^{t,\vec{c}}\ \exists\, \varphi_{t,\vec{c},j,b}$ positive integer;

such that:

1. $|\psi_{t,\vec{c},j} \cdot \varphi_{t,\vec{c},j,b}| \leq \lceil \log_2 \ell \rceil^{e(t)}$, and

2. $\forall\, x \in \rho''$,

$$t(x,\vec{c}) = \sum_{j<\nu_{t,0}} 2^{\alpha_{t,\vec{c},j}} \cdot \psi_{t,\vec{c},j} \cdot f_j^{t,\vec{c}}(x[\vec{a}^{(t,\vec{c})}]) \cdot \sum_{a \in A_j^{t,\vec{c}}} \varphi_{t,\vec{c},j,a} \cdot \wedge x[\mathrm{supp}(a)] \cdot 2^a.$$

Recall that there is m_t such that $||t||_{\ell,m} \leq \ell^{m_t}$ for all large ℓ. Since clusters are involved, recall all definitions in the proof of the cluster-to-finite lemma, regarding clusters, preclusters, $\prec$ and $d(g,g')$. We now let,

$$\mathrm{small}(x) = \sum_{i,\ \text{with}\ \chi_i^{\vec{c}} < m_t \cdot \lceil \log_2 \ell \rceil} h_i^{\vec{c}}(x[\vec{a}^{(s,\vec{c})}]).$$

small(x) is boolean because $\vec{h}^{(\vec{c})}$ is a partition (cf. cluster length sublemma), so that the sum coincides with disjunction. Put large$(x) = 1 -$ small(x).

Claim : *Let* $\vec{g}_s$ *be the first precluster in* $s(x,\vec{c})$. *Then,*

$$s(x,\vec{c}) = s(x,\vec{c}) \cdot \mathrm{large}(x) + \mathrm{small}(x) \cdot$$
$$\cdot \sum_{(a+\alpha_{s,\vec{c},j},j) \in \vec{g}_s} 2^{a+\alpha_{s,\vec{c},j}} \cdot \psi_{s,\vec{c},j} \cdot f_j^{s,\vec{c}}(x[\vec{a}^{(s,\vec{c})}]) \cdot \varphi_{s,\vec{c},j,a} \cdot \wedge x([\mathrm{supp}(a)]).$$

Proof of claim . By definition, a new precluster begins only when

$$d((a+\alpha_{s,\vec{c},j},j),(a'+\alpha_{s,\vec{c},j'},j'),j')) > 2\cdot\lceil\log_2\ell\rceil^e.$$

Therefore the second precluster (if it exists) starts no earlier than $2\cdot\lceil\log_2\ell\rceil^e$. We conclude that for $x\in\rho''$ with small$(x)=1$, we have that $|s(x,\vec{c})| < m_t\cdot\lceil\log_2\ell\rceil < \lceil\log_2\ell\rceil^2 \leq \lceil\log_2\ell\rceil^{e(s)} \leq \lceil\log_2\ell\rceil^e$. This, certainly shows that the binary expansion of $s(x,\vec{c})$ stops well before the second cluster begins.□

Let $\mathrm{v}=\langle\tau_s,\xi_s,\tau_t,\xi_t\rangle$, where $\tau_s,\xi_s\in\{0,1\}^{\nu_{s,0}}$ and $\tau_t,\xi_t\in\{0,1\}^{\nu_{t,0}}$. These quadruples will be called *relevant.* For relevant v we let:

1. $s(\mathrm{v}) = \sum\limits_{(a+\alpha_{s,\vec{c},j},j)\in\vec{g}_s} 2^{a+\alpha_{s,\vec{c},j}}\cdot\psi_{s,\vec{c},j}\cdot\xi_s(j)\cdot\varphi_{s,\vec{c},j,a}\cdot\tau_s(j)$, the value of the first cluster of s relative to v;

2. $\vec{g}(\mathrm{v})=$ the (unique) precluster of t the range of which contains $s(\mathrm{v})$;

3. $t_{\vec{g}(\mathrm{v})}(\mathrm{v}) = \sum\limits_{(a+\alpha_{t,\vec{c},j},j)\in\vec{g}(\mathrm{v})} 2^{a+\alpha_{t,\vec{c},j}}\cdot\psi_{t,\vec{c},j}\cdot\xi_t(j)\cdot\varphi_{t,\vec{c},j,a}\cdot\tau_t(j)$, the value of the *critical cluster* of t relative to v;

4. $\mathrm{left}(x,\mathrm{v}) = \sum\limits_{\substack{\vec{g}\text{ a precluster of }t\\ \vec{g}\succ\vec{g}(\mathrm{v})}} t_{\vec{g}}(x)\cdot 2^{-s(\mathrm{v})}$, the sum of all clusters strictly to the right of the critical cluster of t relative to v;

5. $\mathrm{right}(x,\mathrm{v}) = \sum\limits_{\substack{\vec{g}\text{ a precluster of }t\\ \vec{g}\prec\vec{g}(\mathrm{v})}} t_{\vec{g}}(x)$, the sum of all clusters strictly to the left of the critical cluster of t relative to v;

the truth values of the properties below are necessary to offset the convention $\sum_\emptyset = 1$ when what is needed is that $\sum_\emptyset = 0$; these properties only depend on the sets $A_j^{t,\vec{c}}, j<\nu_{t,0}$, the order $\prec$, and the critical cluster $\vec{g}(\mathrm{v})$, which in turn depends on v, $\vec{c}$ and ρ'' all of which are fixed:

6. Not-first$(\mathrm{v}) = [\![\vec{g}(\mathrm{v})$ is not the $\prec$-minimum precluster in $t]\!]$;

7. Not-last$(\mathrm{v}) = [\![\vec{g}(\mathrm{v})$ is not the $\prec$-maximum precluster in $t]\!]$;

8. Not-first$(\mathrm{v},j) = [\![$there is $a\in A_j^{t,\vec{c}}$ such that $(a+\alpha_{t,\vec{c},j},j)\prec\vec{g}(\mathrm{v})]\!]$, for $j<\nu_{t,0}$;

9. Not-last$(\mathrm{v}, j) = [\![$there is $a \in A_j^{t,\vec{c}}$ such that $(a + \alpha_{t,\vec{c},j}, j) \succ \vec{g}(\mathrm{v})]\!]$, for $j < \nu_{t,0}$.

With $x \in \rho''$, say that x *is* v*-good* if the following hold:

1. x is both (t, ξ_t)-good and (s, ξ_s)-good (cf. cluster-to-finite lemma);
2. $1 = \bigwedge_{(a+\alpha_{s,j},j)\in\vec{g}_s} [\![\tau_s(j) = \wedge x[\mathrm{supp}(a)]]\!]$, and
3. $1 = \bigwedge_{(a+\alpha_{t,j},j)\in\vec{g}(\mathrm{v})} [\![\tau_t(j) = \wedge x[\mathrm{supp}(a)]]\!]$.

$[\![x$ is v-good$]\!]$ is a partition of unity. Given a relevant v, for any v-good $x \in \rho''$ we have that

$$\mathrm{trunc}(t(x,\vec{c})\colon \infty, s(x,\vec{c})) = \mathrm{small}(x) \cdot [\![x \text{ is v-good}]\!] \cdot$$
$$\cdot [\mathrm{Not\text{-}last}(\mathrm{v}) \cdot \mathrm{left}(x,\mathrm{v}) + + \mathrm{trunc}(t_{\vec{g}(\mathrm{v})}(\mathrm{v})\colon \infty, s(\mathrm{v}))] \quad (\mathbf{R})$$

and

$$\mathrm{trunc}(t(x,\vec{c})\colon s(x,\vec{c}), 0) = \mathrm{large}(x) \cdot [\![x \text{ is v-good}]\!] \cdot t(x,\vec{c}) +$$
$$+ \mathrm{small}(x) \cdot [\![x \text{ is v-good}]\!] \cdot [\mathrm{Not\text{-}first}(\mathrm{v}) \cdot \mathrm{right}(x,\mathrm{v}) + \mathrm{trunc}(t_{\vec{g}(\mathrm{v})}(\mathrm{v})\colon s(\mathrm{v}), 0)]. \quad (\mathbf{L})$$

For fixed relevant v we shall need to partition the critical cluster of t. To this end we set for $j < \nu_{t,0}$,

$$\eta(j) = \begin{cases} a + \alpha_{t,\vec{c},j}, & \text{if } a \in A_j^{t,\vec{c}} \text{ and } (a + \alpha_{t,\vec{c},j}, j) \in \vec{g}(\mathrm{v}); \\ \text{undefined}, & \text{if } a \in A_j^{t,\vec{c}} \Longrightarrow (a + \alpha_{t,\vec{c},j}, j) \notin \vec{g}(\mathrm{v}). \end{cases}$$

and call (j) η*-defined* if $\eta(j)$ is defined. Next, we order the $(j), j < \nu_{t,0}$ as follows:

$$(j) \prec (j') \text{ if both } (j),(j') \text{ are } \eta\text{-defined and } \eta(j) < \eta(j') \text{ or else if } j < j'.$$

$\eta(j)$ and $\prec$ are well defined because there is at most one $a \in A_j^{t,\vec{c}}$ with $(a + \alpha_{t,\vec{c},j}, j) \in \vec{g}(\mathrm{v})$ since $\mathrm{spread}(\mathcal{H}(d\colon A_{d,e,\nu,\ell})) > 2 \cdot \nu_{t,0} \cdot \lceil \log_2 \ell \rceil^e$. For (j) η-defined we set $(j)^+$ the η-defined $\prec$-successor of (j), and if (j) is the η-defined $\prec$-maximum then $(j)^+$ is undefined but we let $\eta(j)^+ = \min\{\eta(j) + \lceil \log_2 \ell \rceil^{e(t)}, s(\mathrm{v})\}$. We now distinguish the cases as promised for a fixed relevant v:

Case: R. Put,

$$\begin{aligned} t_{\mathrm{v}} &= \mathrm{trunc}(t_{\vec{g}(\mathrm{v})}(\mathrm{v})\colon \infty, s(\mathrm{v})) \\ &= \mathrm{trunc}\Big(\sum_{\substack{(a+\alpha_{t,j},j) \\ \in\vec{g}(\mathrm{v})}} 2^{a+\alpha_{t,j}} \cdot \xi_t(j) \cdot \psi_{t,j} \cdot \varphi_{t,j,a} \cdot \tau_t(j)\colon \infty, s(\mathrm{v})\Big). \end{aligned}$$

Clearly,

$$|\sum_{\substack{(a+\alpha_{t,j},j)\\ \in\vec{g}(\mathrm{v})}} 2^{a+\alpha_{t,j}}\cdot\xi_t(j)\cdot\psi_{t,j}\cdot\varphi_{t,j,a}\cdot\tau_t(j)|\le 2\cdot\nu_{t,0}\cdot\lceil\log_2\ell\rceil^{e(t)}+\lceil\log_2\nu_{t,0}\rceil$$

but because $|\psi_{t,j}\cdot\varphi_{t,j,a}|\le\lceil\log_2\ell\rceil^{e(t)}$ we actually have that

$$|\sum_{\substack{(a+\alpha_{t,j},j)\\ \in\vec{g}(\mathrm{v})}} 2^{a+\alpha_{t,j}}\cdot\xi_t(j)\cdot\psi_{t,j}\cdot\varphi_{t,j,a}\cdot\tau_t(j)|\le 2\cdot\nu_{t,0}\cdot\lceil\log_2\ell\rceil^{e(t)}.$$

It follows that with

$$\tilde{\eta}(j)=\max\{s(\mathrm{v}),\eta(j)\},\text{ for }(j)\ \eta\text{-defined, and}$$

$$t_{\mathrm{v},j}=\begin{cases}\mathrm{trunc}(t_{\mathrm{v}}{:}\,\eta(j)^+,\tilde{\eta}(j)), & \text{if }(j)\text{ is }\eta\text{-defined}\\ 0, & \text{if }(j)\text{ in not }\eta\text{-defined.}\end{cases}$$

we have that

$$t_{\mathrm{v}}=\sum_{j<\nu_{t,0}} t_{\mathrm{v},j}.$$

We now give the parameters for the cluster form at m. For $j<\nu_{t,0}$, the part responsible for the correct expansion of left(x,v) in $\mathbf{R}$:

1. $A_{\vec{c},\mathrm{v},j}=A_j^{t,\vec{c}}\setminus\{a\mid(a+\alpha_{t,\vec{c},j},j)\in\vec{g}\text{ with }\vec{g}\preceq\vec{g}(\mathrm{v})\}$, $\alpha_{\vec{c},\mathrm{v},j}=\alpha_{t,\vec{c},j}-s(\mathrm{v})$, $\psi_{\vec{c},\mathrm{v},j}=\psi_{t,\vec{c},j}$, $\varphi_{\vec{c},\mathrm{v},j,a}=\varphi_{t,\vec{c},j,a}$ for $a\in A_{\vec{c},\mathrm{v},j}$, and
2. $f_{\vec{c},\mathrm{v},j}=\mathrm{small}(x)\cdot\text{Not-last}(\mathrm{v},j)\cdot[\![x\text{ is v-good}]\!]\cdot f_j^{t,\vec{c}}$

for $\nu_{t,0}\le j<2\cdot\nu_{t,0}$, the part responsible for the correct expansion of t_{v}:

3. $A_{\vec{c},\mathrm{v},j}=\emptyset$, $\alpha_{\vec{c},\mathrm{v},j}=0$, $\psi_{\vec{c},\mathrm{v},j}=t_{\mathrm{v},(j\bmod\nu_{t,0})}$, $\varphi_{\vec{c},\mathrm{v},j,a}=1$, and
4. $f_{\vec{c},\mathrm{v},j}=\mathrm{small}(x)\cdot[\![x\text{ is v-good}]\!]$,

and finally for $2\cdot\nu_{t,0}\le j<3\cdot\nu_{t,0}$ (which are really reserved for the next case):

5. $f_{\vec{c},\mathrm{v},j}=0$.

It is a matter of unraveling the settings and observing the convention $\sum \emptyset = 1$ to see that for any $x \in \rho''$,

$$\mathrm{trunc}(t(x,\vec{c})\colon \infty, s(x,\vec{c})) =$$
$$= \sum_{\mathrm{v\ relevant}} \sum_{j<2\cdot\nu_{t,0}} 2^{\alpha_{\vec{c},\mathrm{v},j}} \cdot \psi_{\vec{c},\mathrm{v},j} \cdot f_{\vec{c},\mathrm{v},j}(x[\vec{a}^{(\vec{c})}]) \cdot \sum_{a\in A_{\vec{c},\mathrm{v},j}} \varphi_{\vec{c},\mathrm{v},j,a} \cdot \wedge x[\mathrm{supp}(a)] \cdot 2^a.$$

This finishes the "right" subcase.

Case: L. Put,

$$\begin{aligned} t_{\mathrm{v}} &= \mathrm{trunc}(t_{\vec{g}(\mathrm{v})}(\mathrm{v})\colon s(\mathrm{v}), 0) \\ &= \mathrm{trunc}\Big(\sum_{\substack{(a+\alpha_{t,j},j)\\ \in\vec{g}(\mathrm{v})}} 2^{a+\alpha_{t,j}} \cdot \xi_t(j) \cdot \psi_{t,j} \cdot \varphi_{t,j,a} \cdot \tau_t(j)\colon s(\mathrm{v}), 0\Big). \end{aligned}$$

Here, with

$$\tilde{\eta}(j) = \min\{\eta(j)^+, s(\mathrm{v})\} \text{ for } (j)\ \eta\text{-defined and}$$
$$t_{\mathrm{v},j} = \begin{cases} \mathrm{trunc}(t_{\mathrm{v}}\colon \tilde{\eta}(j), \eta(j)), & \text{if } (j) \text{ is } \eta\text{-defined;} \\ 0, & \text{if } (j) \text{ is not } \eta\text{-defined.} \end{cases}$$

we have that

$$t_{\mathrm{v}} = \sum_{j<\nu_{t,0}} t_{\mathrm{v},j}.$$

We now fix the parameters for the cluster form at m. For $j < \nu_{t,0}$, the part responsible for the correct expansion of $\mathrm{right}(x,\mathrm{v})$ in **L**:

1. $A_{\vec{c},\mathrm{v},j} = A_j^{t,\vec{c}} \setminus \{a \mid (a+\alpha_{t,\vec{c},j}, j) \in \vec{g} \text{ with } \vec{g} \succeq \vec{g}(\mathrm{v})\}$, $\alpha_{\vec{c},\mathrm{v},j} = \alpha_{t,\vec{c},j}$, $\psi_{\vec{c},\mathrm{v},j} = \psi_{t,\vec{c},j}$, $\varphi_{\vec{c},\mathrm{v},j,a} = \varphi_{t,\vec{c},j,a}$ for $a \in A_{\vec{c},\mathrm{v},j}$, and
2. $f_{\vec{c},\mathrm{v},j} = \text{Not-first}(\mathrm{v},j) \cdot [\![x \text{ is v-good}]\!] \cdot f_j^{t,\vec{c}}$,

for $\nu_{t,0} \le j < 2\cdot\nu_{t,0}$, the part responsible for the correct expansion of t_{v}:

3. $A_{\vec{c},\mathrm{v},j} = \{a\}$, $\varphi_{\vec{c},\mathrm{v},j,a} = 1$ for $(j \mod \nu_{t,0})$ η-defined with $(a+\alpha_{t,\vec{c},(j \bmod \nu_{t,0})}, (j \mod \nu_{t,0})) \in \vec{g}(\mathrm{v})$, and $A_{\vec{c},\mathrm{v},j} = \emptyset$ for $(j \mod \nu_{t,0})$ not η-defined,
4. $\psi_{\vec{c},\mathrm{v},j} = t_{\mathrm{v},(j \bmod \nu_{t,0})}$, $\alpha_{\vec{c},\mathrm{v},j} = \alpha_{t,\vec{c},(j \bmod \nu_{t,0})}$, $f_{\vec{c},\mathrm{v},j} = \mathrm{small}(x) \cdot [\![x \text{ is v-good}]\!]$,

and finally for $2\cdot\nu_{t,0} \le j < 3\cdot\nu_{t,0}$, this part being intended for $x \in \rho''$ with $\mathrm{large}(x) = 1$:

5. $A_{\vec{c},\mathbf{v},j} = A^{t,\vec{c}}_{(j \bmod \nu_{t,0})}$, $\alpha_{\vec{c},\mathbf{v},j} = \alpha_{t,\vec{c},(j \bmod \nu_{t,0})}$,

6. $\psi_{\vec{c},\mathbf{v},j} = \psi_{t,\vec{c},(j \bmod \nu_{t,0})}$, $\varphi_{\vec{c},\mathbf{v},j,a} = \varphi_{t,\vec{c},(j \bmod \nu_{t,0}),a}$ for $a \in A_{\vec{c},\mathbf{v},j}$, and

7. $f_{\vec{c},\mathbf{v},j} = \mathrm{large}(x) \cdot [\![x \text{ is v-good}]\!] \cdot f^{t,\vec{c}}_{(j \bmod \nu_{t,0})}$,

It is a matter of unraveling the settings and observing the convention $\sum \emptyset = 1$ to see that for any $x \in \rho''$,

$$\mathrm{trunc}(t(x,\vec{c})\colon s(x,\vec{c}),0) =$$
$$= \sum_{\mathbf{v} \text{ relevant}} \sum_{j<2\cdot\nu_{t,0}} 2^{\alpha_{\vec{c},\mathbf{v},j}} \cdot \psi_{\vec{c},\mathbf{v},j} \cdot f_{\vec{c},\mathbf{v},j}(x[\vec{a}^{(\vec{c})}]) \cdot \sum_{a \in A_{\vec{c},\mathbf{v},j}} \varphi_{\vec{c},\mathbf{v},j,a} \cdot \wedge x[\mathrm{supp}(a)] \cdot 2^a.$$

This finishes the "left" subcase.

We can now give the first parameters in the definition of cluster form at m. First we determine $\nu_{r,0}$. We set

$$\nu_{r,0} = \max\{3 \cdot \nu_{t,0} \cdot 2^{\kappa}, \nu_{s,0}\}$$

where κ is such that 2^{κ} is a bound for the number of relevant $\mathbf{v}$'s. In fact,

$$\kappa = 4 \cdot (\nu_{t,0} + \nu_{s,0})$$

will do. Next, μ_r. It suffices to find an upper bound to the number of bits that are needed to determine the truth value

$$[\![x \text{ is v-good}]\!] \cdot \mathrm{small}(x).$$

Clearly, the following is such an upper bound

$$\mu_t + \mu_s + d \cdot (\nu_{t,0} + \nu_{s,0}).$$

□

The induction on the complexity of a cut-and-paste term is now complete.□

Boolean Cut-and-paste Circuit Lemma 6.3. *Let m and n be nonzero positive integers, $t_i(x,\vec{z})$, $i < n$, cut-and-paste terms and Ψ a boolean combination of predicates in the terms t_i, $i < n$. Then,*

$\exists \nu, m_0, d, \ell_0, \delta\colon 0 < \delta \leq 1$;

$\forall \ell \geq \ell_0$;

$\exists \rho \in \mathrm{Restr}_\delta(d, e, \nu, \ell)$;
$\forall \vec{c} < \ell^m$;
$\exists C^{\Psi}_{\ell,\vec{c},\rho}$ *a circuit of size* $\leq \ell^{m_0}$ *and depth* ≤ 6;
$\forall x \in \rho$,

$$[\![\Psi(x, \vec{c})]\!] = C^{\Psi}_{\ell,\vec{c},\rho}(x).$$

Proof. Create a new cut-and-paste term $s(x, \vec{z})$ by concatenating the terms of Ψ at appropriate distances so that they do not have any bit-interference, i.e. with $t_{-1} = 0$ put

$$s(x, \vec{z}) = \sum_{i<n} t_i(x, \vec{z}) \cdot 2^{\sum_{-1 \geq j < i} |2 \cdot t_j(x, \vec{z})|}.$$

s has the bit-finite property at m. Also note that there is no bit-interference between the terms t_i. The bits of a term t_i are bits of s. The boolean combination Ψ can be expressed as a circuit of depth 2 and constant size in the atomic relations '=' and '<' (disjunctive normal form). Each of the relations $t_i < t_j$ and $t_i = t_j$ can be expressed as a circuit of depth 2 in the bits of t_i, t_j; its size is evidently polynomial in the lengths of t_i, t_j. And finally, the bits of each t_i are boolean functions of no more than μ bits of the input x, $x \in \rho'$ where ρ' is the restriction we get from the bit-finite property at m applied to s. These boolean functions written in disjunctive normal form, again have depth 2 and size $\leq 2^\mu$. So for $x \in \rho'$ we have that $[\![\Psi(x, \vec{c})]\!] = C^{\Psi}_{\ell,\vec{c},\rho'}(x)$ a circuit of depth at most 6 and size polynomial in $|x|$.□

Sharp Cut-and-paste Circuit Lemma 6.4. *Let* $\Theta(x)$ *be a cut-and-paste sharply bounded formula. Then,*
$\exists\, N, m_0, \nu, d, \ell_0, \delta : 0 < \delta \leq 1$;
$\forall \ell \geq \ell_0$;
$\exists \rho \in \mathrm{Restr}_\delta(d, e, \nu, \ell)$;
$\exists C^{\Theta}_{\ell,\rho}$ *a circuit of size* $\leq \ell^{m_0}$ *and depth* $\leq N + 6$;
$\forall x \in \rho$,

$$[\![\Theta(x)]\!] = C^{\Theta}_{\ell,\rho}(x).$$

Proof. $\Theta(x)$ is *logically* equivalent to

$$Q_0 z_0 < |x|^m \ldots Q_{N-1} z_{N-1} < |x|^m \Psi(x, \vec{z})$$

for some nonzero positive integers N and m, where Q_i are quantifiers, for $i < N$, $\vec{z} = z_0 \dots z_{N-1}$, and a quantifier free cut-and-paste formula $\Psi(x, \vec{z})$. Apply the boolean cut-and-paste circuit lemma to m and $\Psi(x, \vec{z})$. Get ν, m_0, d, ℓ_0, and δ with $0 < \delta \leq 1$. For $\ell \geq \ell_0$, get a restriction $\rho \in \mathrm{Restr}_\delta(d, e, \nu, \ell)$ such that
$\forall \vec{c} < \ell^m$
$\exists C^{\Psi}_{\ell,\vec{c},\rho}$ a circuit of size $\leq \ell^{m_0}$ and depth ≤ 6;
$\forall x \in \rho$,

$$[\![\Psi(x, \vec{c})]\!] = C^{\Psi}_{\ell,\vec{c},\rho}(x).$$

Let

$$C^{\Theta}_{\ell,\rho}(x) = \underset{c_0 < \ell^m}{\mathrm{Bool}(Q_0)} \cdots \underset{c_{N-1} < \ell^m}{\mathrm{Bool}(Q_{N-1})}\, C^{\Psi}_{\ell,\vec{c},\rho}(x, \vec{c})$$

with $\vec{c}$ substituted for $\vec{z}$. For $x \in \rho$,

$$[\![\Theta(x)]\!] = C^{\Theta}_{\ell,\rho}(x).$$

□

Proof of the theorem

Let $\Theta(x)$ be a cut-and-paste sharply bounded formula in x, defining **PARITY**(x). Apply the sharp cut-and-paste circuit lemma to Θ. Get $N, m_0, \nu, d, \ell_0, \delta: 0 < \delta \leq 1$. Consider $\lambda = \ell^{1/\delta} \geq \ell_0$. Find $\rho \in \mathrm{Restr}_\delta(d, e, \nu, \lambda)$ and a circuit $C^{\Theta}_{\rho,\lambda}$ of size λ^{m_0} and depth $\leq N + 6$, such that

$$[\![\Theta(x)]\!] = C^{\Theta}_{\rho,\lambda}(x)$$

holds for any $x \in \rho$. But $|\mathrm{free}(\rho)| \geq \lambda^\delta = \ell$. By (possibly) fixing α many bits in free(ρ) so that $|\mathrm{free}(\rho)| - \alpha = \ell$ we get a new circuit C^{Θ}_{ℓ} from $C^{\Theta}_{\rho,\lambda}$ and a new restriction ρ' such that for $x \in \rho'$ we have that

$$[\![\mathbf{PARITY}(x)]\!] = C^{\Theta}_{\ell}(x).$$

But this is impossible because the depth of C^{Θ}_{ℓ} is bounded by $N + 6$ and its size is bounded by $\ell^{m_0/\delta}$. ■

Chapter 7

INDEXPARITY is not $\mathcal{SB}_1^{\mathrm{N}}(\mathbf{trunc}, \mathbf{CARD})$ definable

The main result in this chapter

Theorem. **INDEXPARITY** *is not symmetric cut-and-paste sharply bounded definable.*

An analogous result was proved back in chapter 5. There we could build a feasible multicircuit because we had to deal only with the cardinality of x. The present situation, however, forces us to consider cardinalities of polynomially many *intervals* of the binary expansion of cut-and-paste terms in x. Therefore, we introduce cylinders of *symmetries* and their *volume*, i.e. the set of binary words belonging to the orbit of a fixed binary word, where the orbit is taken with respect to the action of a cylinder on binary words.

Notation, conventions, and remarks

For items without local definition consult their *most recent* definition in previous chapters, unless the reference for their definition is explicit:

7a. We add to the cut-and-paste language the binary predicate symbol $\mathbf{CARD}(x, y)$ which is the same predicate symbol with the same interpretation as in chapter 5. The new sharply bounded formulae are called *symmetric cut-and-paste sharply bounded.*

7b. Fix $\ell \in \mathbf{N}, \ell > 0$ and consider $\mathbf{S}_\ell$, the group of permutations on ℓ elements. Let G_ℓ be the subgroup of $\mathbf{S}_\ell$ consisting of two elements the identity id and α the product of all transpositions of the form $(a, a+1)$, for even $a < \ell$, i.e. $a \in \text{Even}_\ell$, with the convention that if ℓ is odd we take $(\ell-1, \ell-1)$ instead of $(\ell-1, \ell)$. We now say,

1. π *is a* G_ℓ*-word* if π: $\text{Even}_\ell \longrightarrow G_\ell$, and
2. ρ *is a* G_ℓ*-restriction (*G_ℓ*-cylinder)* if ρ: $\text{Even}_\ell \longrightarrow G_\ell \bigcup \{*\}$ and both free(ρ) and unfree(ρ) are defined as before.

Convention. We shall drop the subscript ℓ from G_ℓ since for $\ell \geq 2$ these groups are isomorphic and have order 2. With X a set of numbers G^X is the set of all functions from X into G. For $a \in \text{Even}_\ell$ and π a G-word, we shall write π_a (ρ_a) for the element in G which is assigned to a by π (ρ).

Remark. The set of G-words forms a group with respect to the operation

$$\pi \cdot \pi' = \langle \pi_a \cdot \pi'_a \mid a \in \text{Even}_\ell \rangle,$$

where the '$\cdot$' appearing in the righthand expression is the group operation of $\mathbf{S}_\ell$. A G-restriction (cylinder) does not necessarily form a group in the above sense.

7c. Given ρ, a G-restriction, the *cylinder based on* ρ is the set

$$\{\pi \mid \pi \text{ a } G\text{-word such that } \forall a \in \text{unfree}(\rho)\ \pi_a = \rho_a\}\ .$$

The cylinder based on ρ will also be denoted by ρ.

7d. Let $d, \ell > 0$, be integers and $A \subseteq [\ell]$. We need refinements of the definition of "wide for d", because we have to deal with permuted clusters. We say A *is* $\Delta(G)$*-wide for* d if A is wide for d and whenever:

1. $a_1, \ldots, a_\mu$ $(= \vec{a})$, $a'_1, \ldots, a'_{\mu'}$ $(= \vec{a}')$, $b_1, \ldots, b_\nu$ $(= \vec{b})$, and $b'_1, \ldots, b'_{\nu'}$ $(= \vec{b}')$ are sequences each with distinct entries out of A, with $\mu, \mu', \nu, \nu' \leq d$, and
2. $i_1 \cdot x_1 + \cdots + i_\mu \cdot x_\mu$ $(= \vec{\imath} \bullet \vec{x})$, $i'_1 \cdot x_1 + \cdots + i'_{\mu'} \cdot x_{\mu'}$ $(= \vec{\imath}' \bullet \vec{x})$, $j_1 \cdot x_1 + \cdots + j_\nu \cdot x_\nu$ $(= \vec{\jmath} \bullet \vec{x})$, $j'_1 \cdot x_1 + \cdots + j'_{\nu'} \cdot x_{\nu'}$ $(= \vec{\jmath}' \bullet \vec{x}) \in \mathcal{H}(d\colon x_1, \ldots, x_d)$

then

$$\vec{\imath} \bullet \vec{a} - \vec{\jmath} \bullet \vec{b} \neq \vec{\imath}' \bullet \vec{a}' - \vec{\jmath}' \bullet \vec{b}'$$

unless

$$\vec{\imath} \bullet \vec{a}_{-\mathbf{a}} - \vec{\jmath} \bullet \vec{b}_{-\mathbf{a}} = \vec{\imath}' \bullet \vec{a}'_{-\mathbf{a}} - \vec{\jmath}' \bullet \vec{b}'_{-\mathbf{a}} \ ,$$

for some $\mathbf{a} \in \mathrm{supp}(a, b, a', b')$ where $\vec{\imath} \bullet \vec{a}_{-\mathbf{a}}$ results from $\vec{\imath} \bullet \vec{x}$ by substituting into $\vec{\imath} \bullet \vec{x}$ the numbers in $\vec{a}$ *except* that instead of $\mathbf{a}$ we substitute 0.

7e. Let d, e, ν, and ℓ be positive nonzero integers, and let $A_{d,e,\nu,\ell} \subset \mathrm{Even}_\ell$ the $\Delta(G)$-wide set for d of width $> 2 \cdot \nu \cdot \lceil \log_2 \ell \rceil^e$ (the existence of which follows from the $\Delta(G)$-wide lemma below). We then set:

1. $\sigma_{d,e,\nu,\ell}(a) = \begin{cases} 1, & \text{if } a \in A_{d,e,\nu,\ell}; \\ 0, & \text{if } a \notin A_{d,e,\nu,\ell} \end{cases}$;
2. $\mathrm{vol}(\rho) = \mathrm{vol}_{d,e,\nu,\ell}(\rho) = \{\pi(\sigma_{d,e,\nu,\ell}) \mid \pi \in \rho\}$, for ρ a G-restriction, and
3. as in chapter 6 and for $\epsilon \in (0, 1]$,

$$G\text{-Restr}_\epsilon(d, e, \nu, \ell) = \{\rho \mid \rho \text{ a } G\text{-restriction with } |\mathrm{free}(\rho) \bigcap A_{d,e,\nu,\ell}| \geq \ell^\epsilon\}.$$

7f. Given is a G-word, π. We now describe the actions of π on various objects:

1. the action on $x < 2^\ell$ is described by setting,

$$y = \pi(x) \iff \forall\, a < \ell,\ y(a) = x(\pi_{a-(a \bmod 2)}(a));$$

2. the action on $\vec{a} < \ell$ with distinct entries:

$$\pi\vec{a} = \pi_{a_1-(a_1 \bmod 2)}(a_1) \ldots \pi_{a_k-(a_k \bmod 2)}(a_k)$$

where $a - a \bmod 2$ is needed since G-words are only defined on Even_ℓ, and finally

3. the action on $a = \vec{\imath} \bullet \vec{a} \in \mathcal{H}(d\colon A_{d,e,\nu,\ell})$:

$$\pi\vec{\imath} \bullet \vec{a} = \vec{\imath} \bullet \pi\vec{a}.$$

Remark. The above action is well defined because, as it is easily seen, for $a = \vec{\imath} \bullet \vec{a}, b = \vec{\jmath} \bullet \vec{b}$ with $a, b \in \mathcal{H}(d\colon A_{d,e,\nu,\ell})$ we have that

$$a = b \Longrightarrow \pi a = \pi b.$$

In general, however, if $\mathrm{rng}\,\vec{a}, \mathrm{rng}\,\vec{b}$ are not chosen from a set which is wide or $\Delta(G)$-wide for d then

$$\vec{\imath} \bullet \vec{a} = \vec{\jmath} \bullet \vec{b} \not\Longrightarrow \pi\vec{\imath} \bullet \vec{a} = \pi\vec{\jmath} \bullet \vec{b}$$

and the action is not well defined.

7g. Let $t(x, \vec{z})$ be a cut-and-paste term and m a positive integer. We say $t(x, \vec{z})$ *is in symmetric cluster form at* m if,

$\forall\, d \geq d(t), \epsilon\colon 0 < \epsilon \leq 1/(2 \cdot d + 2)$;

$\exists\, \mu, \nu_0, \delta\colon 0 < \delta \leq \epsilon$;

$\forall\, \nu \geq \nu_0, e \geq e(t) \exists\, \ell_0 \forall\, \ell \geq \ell_0$;

$\forall\, \rho \in G\text{-Restr}_\epsilon(d, e, \nu, \ell) \exists\, \rho' \in G\text{-Restr}_\delta(d, e, \nu, \ell), \rho' \subseteq \rho$;

$\forall\, \vec{c} < \ell^m$;

$\exists\, \vec{a}^{(\vec{c})}$ sequence of no more than μ bits;

$\exists\, \vec{f}^{(\vec{c})} \in \text{BOOL}_\mu^{\nu_0}$;

$\forall\, j < \nu_0$;

$\exists\, \alpha_{\vec{c},j} \in \mathbf{Z}\,, \psi_{\vec{c},j}$ positive integer;

$\exists\, A_j^{\vec{c}} \subseteq \mathcal{H}(d(t)\colon A_{d,e,\nu,\ell})$;

$\forall\, b \in A_j^{\vec{c}} \exists\, \varphi_{\vec{c},j,b}$ positive integer;

such that:

1. $|\psi_{\vec{c},j} \cdot \varphi_{\vec{c},j,b}| \leq \lceil \log_2 \ell \rceil^{e(t)}$, and
2. $\forall\, x \in \text{vol}(\rho')$

$$t(x, \vec{c}) = \sum_{j < \nu_0} 2^{\alpha_{\vec{c},j}} \cdot \psi_{\vec{c},j} \cdot f_j^{\vec{c}}(x[\vec{a}^{(\vec{c})}]) \cdot \sum_{a \in A_j^{\vec{c}}} \varphi_{\vec{c},j,a} \cdot 2^{\pi a},$$

where $x = \pi \sigma_{d,e,\nu,\ell}, \pi \in \rho'$.

7h. Let $t(x, \vec{z})$ be a cut-and-paste term and m a positive integer. We say $t(x, \vec{z})$ *has the symmetric bit-finite property at* m if,

$\forall\, d \geq d(t), \epsilon\colon 0 < \epsilon \leq 1/(2 \cdot d + 2)$;

$\exists\, \mu, \nu_0, \delta\colon 0 < \delta \leq \epsilon$;

$\forall\, \nu \geq \nu_0, e \geq e(t) \exists\, \ell_0 \forall\, \ell \geq \ell_0$;

$\forall\, \rho \in G\text{-Restr}_\epsilon(d, e, \nu, \ell) \exists\, \rho' \in G\text{-Restr}_\delta(d, e, \nu, \ell), \rho' \subseteq \rho$;

$\forall\, \vec{c} < \ell^m$;

$\forall\, y < \|t\|_{\ell,m}\ \exists\, \vec{a}^{(\vec{c},y)}$ a sequence of no more than μ bits $< \ell$, $f_{\vec{c},y} \in \text{BOOL}_\mu$;

$\forall\, x \in \text{vol}(\rho')$,

$$t(x, \vec{c}) = \sum_{y < \|t\|_{\ell,m}} 2^y \cdot f_{\vec{c},y}(x[\vec{a}^{(\vec{c},y)}]).$$

7i. A *random* G-restriction, $\hat{\rho}$, is a vector of $\lambda = \lfloor \ell/2 \rfloor$, independent random variables, one for each pair $(a, a+1)$ with $a \in \mathrm{Even}_\ell$. So,

$$\hat{\rho} = (X_{(a,a+1)} : a \in \mathrm{Even}_\ell) \ .$$

These random variables are trials with three outcomes: the members of G and $*$, with $p_* = 1/\sqrt{\lambda}$, $p_\alpha = p_{id} = 1/2 \cdot (1 - p_*)$.

Probabilistic lemmata

The purpose of these lemmata is to establish an almost exact analogy between classical random restrictions and polynomial size constant, depth circuits as found in [FSS84] on the one hand and G-restrictions and polynomial size, constant depth circuits. In what follows d, e, and ν are fixed nonzero positive integers.

G-Lemma 7.1. *Let C be a polynomial size, constant depth circuit. Then,*
$\forall\, \epsilon, 0 < \epsilon \leq 1;$
$\exists\, \ell_0, \delta : 0 < \delta \leq \epsilon;$
$\forall\, \ell \geq \ell_0;$
$\forall\, \rho \in G\text{-}Restr_\epsilon(d, e, \nu, \ell) \exists\, \rho' \in G\text{-}Restr_\delta(d, e, \nu, \ell), \rho' \subseteq \rho,$
C *is constant in* $\mathrm{vol}(\rho')$.

Proof. Define the function,

$$\iota : G\text{-words} \xrightarrow{\text{onto}} 2^{A_{d,e,\nu,\ell}}$$

by setting

$$\begin{aligned} \iota(\pi) = y &\iff \forall\, a < \ell,\ y(a) = \pi\sigma_{d,e,\nu,\ell}(a) \cdot \sigma_{d,e,\nu,\ell}(a) \\ &\iff \forall\, a < \ell, y(a) = \begin{cases} 0, & \text{if } a \notin A_{d,e,\nu,\ell}; \\ 1, & \text{if } a \in A_{d,e,\nu,\ell} \text{ and } \pi_a = id; \\ 0, & \text{if } a \in A_{d,e,\nu,\ell} \text{ and } \pi_a = \alpha \end{cases} . \end{aligned}$$

Let C_ℓ^G be the circuit obtained from C_ℓ by altering the bits of input in *two* stages:

i. first change $a+1$ to $\neg a$ for $a \in A_{d,e,\nu,\ell}$ to obtain an intermediate circuit, and then

ii. in the intermidiate circuit from (i) change a to 0 ($\neg a$ to 1) for $a \notin A_{d,e,\nu,\ell}$ to obtain C_ℓ^G.

Clearly now,

$$\forall \pi, \text{ a } G\text{-word}, C_\ell^G(\iota(\pi)) = C_\ell(\pi(\sigma_{d,e,\nu,\ell})) .$$

We now argue on C_ℓ^G as in [FSS84] to produce a cylinder (restriction), $\tilde{\rho} \subseteq 2^{A_{d,e,\nu,\ell}}$, such that C_ℓ^G is constant in $\tilde{\rho}$. Set $\rho = \iota^{-1}(\tilde{\rho})$. ρ is a G-cylinder (restriction) with $|\text{free}(\rho) \bigcap A_{d,e,\nu,\ell}| = |\text{free}(\tilde{\rho})|$. Finally, let $x \in \text{vol}(\rho)$ and let $\pi \in \rho$ be the G-word such that $x = \pi(\sigma_{d,e,\nu,\ell})$, and put $y = \iota(\pi)$. Then $y \in \tilde{\rho}$ and

$$\begin{aligned} C_\ell(x) &= C_\ell(\pi(\sigma_{d,e,\nu,\ell})) \\ &= C_\ell^G(\iota(\pi)) \\ &= C_\ell^G(y). \end{aligned}$$

It follows that C_ℓ is constant in $\text{vol}(\rho)$.□

Multicircuit G-Lemma 7.2. *Let* $\mathbf{C} = \{\mathbf{C}_\ell\}_{\ell=0}^{\infty}$ *be a polynomial size and width, constant depth multicircuit. Then,*

$\forall \epsilon, 0 < \epsilon \leq 1;$

$\exists \mu, \ell_0, \delta: 0 < \delta \leq \epsilon;$

$\forall \ell \geq \ell_0;$

$\forall \rho \in G\text{-}Restr_\epsilon(d, e, \nu, \ell) \exists \rho' \in G\text{-}Restr_\delta(d, e, \nu, \ell), \rho' \subseteq \rho;$

$\forall \lambda < p_{width}(\ell);$

$\exists f_\lambda \in \text{BOOL}_\mu, \vec{a}^{(\lambda)}$ *of* $\leq \mu$ *bits;*

$\forall x \in \text{vol}(\rho'),$

$$C_{\ell,\lambda}(x) = f_\lambda(x[\vec{a}^{(\lambda)}]) .$$

Proof. We apply the multicircuit lemma of chapter 5 to the multicircuit $\mathbf{C}^G$ which is obtained from $\mathbf{C}$ by replacing the circuits $C_{\ell,\lambda}$ with the circuits $C_{\ell,\lambda}^G$, for $\lambda < p_{width}$.□

Refinements and corrolaries of the wide lemma

Δ(G)-wide Lemma 7.3. *Fix* $d \in \mathbf{Z}$, $d > 0$. *There is a positive integer* $\rho(d)$ *such that for every positive integer* λ *there is a set of positive integers* A *such that,*

$\lambda 1$. $|A| = \lambda$;

$\lambda 2$. $a \in A \Longrightarrow a \leq \rho(d) \cdot \lambda^{4 \cdot d}$, *and*

$\lambda 3$. A *is* $\Delta(G)$*-wide for* d .

Proof. Given d consider arbitrary λ. By the wide lemma applied to $2 \cdot d$ we get a set A which is wide for $2 \cdot d$, has λ elements, and $a \in A \Longrightarrow a < \rho(2 \cdot d) \cdot \lambda^{4 \cdot d}$, where $\rho(2 \cdot d)$ is as in the wide lemma. If we only want even or only odd numbers then we consider $2 \cdot \rho(2 \cdot d)$ as the upper bound for $\max A$. Suppose now that an equation, **E**, of the form

$$\vec{\imath}' \bullet \vec{a}\,' - \vec{\jmath}' \bullet \vec{b}\,' = \vec{\imath}'' \bullet \vec{a}\,'' - \vec{\jmath}'' \bullet \vec{b}\,'' \qquad \text{(E)}$$

holds, with $\sum \vec{\imath}'$, $\sum \vec{\jmath}'$, $\sum \vec{\imath}''$, $\sum \vec{\jmath}'' \leq d$ and the sequences $\vec{a}', \vec{b}', \vec{a}''$, and $\vec{b}''$, each have distinct entries from A. Certainly then $\vec{\imath} \bullet \vec{a} = \vec{\imath}' \bullet \vec{a}\,' + \vec{\jmath}'' \bullet \vec{b}\,''$ and $\vec{\jmath} \bullet \vec{b} = \vec{\imath}'' \bullet \vec{a}\,'' + \vec{\jmath}' \bullet \vec{b}\,'$ for some $\vec{\imath} \bullet \vec{a}$, $\vec{\jmath} \bullet \vec{b} \in \mathcal{H}(2 \cdot d\colon A)$. Equation **E** is equivalent to

$$\vec{\imath} \bullet \vec{a} = \vec{\jmath} \bullet \vec{b} \ .$$

Now, since A is wide for $2 \cdot d$ we get that $\vec{\imath} = \vec{\jmath}$ and the sequences $\vec{a}, \vec{b}$ belong to the same multiplicity class. In particular we have that $\operatorname{rng} \vec{a} = \operatorname{rng} \vec{b}$. Choose any $\mathbf{a} \in \operatorname{rng} \vec{a}$. We conclude that

$$\vec{\imath} \bullet \vec{a}_{-\mathbf{a}} = \vec{\jmath} \bullet \vec{b}_{-\mathbf{a}}$$

and hence

$$\vec{\imath}' \bullet \vec{a}\,'_{-\mathbf{a}} - \vec{\jmath}' \bullet \vec{b}\,'_{-\mathbf{a}} = \vec{\imath}'' \bullet \vec{a}\,''_{-\mathbf{a}} - \vec{\jmath}'' \bullet \vec{b}\,''_{-\mathbf{a}} \ .$$

□

A cluster is created by an integer shift. Roughly speaking, the corrolary below tells us that G-words which fix a prescribed set, S, will move clusters without changing their binary expansion and hence in particular their cardinality.

Finite Support Corollary. *Fix* $d \in \mathbf{N}$, $d > 0$, *and* A *a* $\Delta(G)$*-wide set for* d. *Then,*
$\forall r > 0 \, \forall a, b, a', b' \in \mathcal{H}(d\colon A) \, \exists \, a_r, b_r, a'_r, b'_r, r(a,b), r(a',b') \in \mathcal{H}(d\colon A)$

$$r = a - b = a' - b' \ \Longrightarrow a = a_r + r(a,b), b = b_r + r(a,b), a' = a'_r + r(a',b'), b' = b'_r + r(a',b')$$

and

$$\mathrm{supp}(a_r, b_r) = \mathrm{supp}(a'_r, b'_r) \ .$$

Proof. Denote by Δ_r the set of pairs (a, b) in $\mathcal{H}(d\colon A)$ with $a - b = r$. By a *sub-expression* of $a = \vec{\imath} \bullet \vec{a}$ we mean the resriction of $\vec{\imath} \bullet \vec{x}$ to a certain subsequence of $\vec{a}$. For every pair $(a, b) \in \Delta_r$ let $r(a, b)$ be *the* maximal sub-expression (if it exists, otherwise set $r(a, b) = 0$) of both a and b, such that for some $a_r, b_r \in \mathcal{H}(d\colon A) \bigcup \{0\}$ we have that

$$a = a_r + r(a, b),\ b = b_r + r(a, b).$$

Certainly then $a_r - b_r = a - b = r$, so that $(a_r, b_r) \in \Delta_r$. Let Δ_r^- be the set of (a_r, b_r), $(a, b) \in \Delta_r$, and $\gamma_r = \min_{(x,y)\in\Delta_r^-} |\mathrm{supp}(x, y)|$. Note that for any r, $\gamma_r \leq 2 \cdot d$. We now prove the statement

$$\forall\, r > 0\, \forall (x, y), (z, w) \in \Delta_r^-\ \mathrm{supp}(x, y) = \mathrm{supp}(z, w)$$

by induction on γ_r.

Fix arbitrary $r > 0$ and assume that $\Delta_r^- \neq \emptyset$. Clearly then, $1 \leq \gamma_r \leq 2 \cdot d$.

When $\gamma_r = 1$: Consider arbitrary $(a_r, b_r), (a'_r, b'_r) \in \Delta_r^-$ such that $|\mathrm{supp}(a_r)| = 1$. Since A is $\Delta(G)$-wide for d, we can find $\mathbf{a} \in \mathrm{supp}(a_r, b_r, a'_r, b'_r)$ such that

$$(a_r)_{-\mathbf{a}} - (b_r)_{-\mathbf{a}} = (a'_r)_{-\mathbf{a}} - (b'_r)_{-\mathbf{a}} \ .$$

We claim that

$$\mathbf{a} \in \mathrm{supp}(a_r, b_r) \cap \mathrm{supp}(a'_r, b'_r) \ .$$

For if $\mathbf{a} \in \mathrm{supp}(a'_r, b'_r) \setminus \mathrm{supp}(a_r, b_r)$ then

$$\begin{aligned} a_r - b_r &= (a_r)_{-\mathbf{a}} - (b_r)_{-\mathbf{a}} \\ &= (a'_r)_{-\mathbf{a}} - (b'_r)_{-\mathbf{a}} \\ &= a'_r - b'_r \end{aligned}$$

and so the coefficient of $\mathbf{a}$ in $a'_r - b'_r$ is 0. But $\mathbf{a}$ has nonzero coefficient in at least one of a'_r and b'_r. So we can nontrivially extend $r(a', b')$, contradicting its maximality. We conclude that $\mathbf{a} \in \mathrm{supp}(a_r, b_r)$. Actually, $\mathbf{a} \in \mathrm{supp}(a_r) \cap \mathrm{supp}(b_r)$ because

$$1 \leq |\mathrm{supp}(a_r)|, |\mathrm{supp}(b_r)| \leq |\mathrm{supp}(a_r, b_r)| = 1 \ .$$

Hence, $0 = (a_r)_{-\mathbf{a}} - (b_r)_{-\mathbf{a}} = (a'_r)_{-\mathbf{a}} - (b'_r)_{-\mathbf{a}}$ which gives that $(a'_r)_{-\mathbf{a}} = (b'_r)_{-\mathbf{a}}$. Again because A is wide for d, if it were the case that $0 \neq (a'_r)_{-\mathbf{a}} = (b'_r)_{-\mathbf{a}}$, then we could nontrivially extend $r(a', b')$, contradicting its maximality. So, finally, $0 = (a'_r)_{-\mathbf{a}} = (b'_r)_{-\mathbf{a}}$ which implies that $\{\mathbf{a}\} = \mathrm{supp}(a'_r) = \mathrm{supp}(b'_r)$ and we are done.

When $\gamma_r > 1$: Here we assume that for any r' such that $1 \leq \gamma_{r'} < \gamma_r$ we have that

$$\forall (x, y), (z, w) \in \Delta^-_{r'}\ \mathrm{supp}(x, y) = \mathrm{supp}(z, w)\ .$$

We choose $(a_r, b_r), (a'_r, b'_r) \in \Delta^-_r$ such that $|\mathrm{supp}(a_r, b_r)| = \gamma_r$. Since A is $\Delta(G)$-wide for d, we can find $\mathbf{a} \in \mathrm{supp}(a_r, b_r, a'_r, b'_r)$ be such that

$$(a_r)_{-\mathbf{a}} - (b_r)_{-\mathbf{a}} = (a'_r)_{-\mathbf{a}} - (b'_r)_{-\mathbf{a}}\ .$$

As before we have that $\mathbf{a} \in \mathrm{supp}(a_r, b_r) \cap \mathrm{supp}(a'_r, b'_r)$. Set $r' = (a_r)_{-\mathbf{a}} - (b_r)_{-\mathbf{a}}$, $\tilde{a}_r = (a_r)_{-\mathbf{a}}$, $\tilde{b}_r = (b_r)_{-\mathbf{a}}$, $\tilde{a}'_r = (a'_r)_{-\mathbf{a}}$, and $\tilde{b}'_r = (b'_r)_{-\mathbf{a}}$. Clearly now,

1. $(\tilde{a}_r, \tilde{b}_r), (\tilde{a}'_r, \tilde{b}'_r) \in \Delta^-_{r'}$, and
2. $1 \leq \gamma_{r'} \leq \gamma_r - 1 < \gamma_r$.

So by the induction hypothesis applied to $\gamma_{r'}$, it follows that

$$\begin{aligned} \mathrm{supp}(a_r, b_r) &= \mathrm{supp}(\tilde{a}_r, \tilde{b}_r) \cup \{\mathbf{a}\} \\ &= \mathrm{supp}(\tilde{a}'_r, \tilde{b}'_r) \cup \{\mathbf{a}\} \\ &= \mathrm{supp}(a'_r, b'_r). \end{aligned}$$

The induction is complete and the lemma follows.□

The Cluster Shift Argument.

We describe an argument and its notation. We agree to invoke this argument by saying "by the cluster shift argument" and then use its notation and its conclusions whenever we invoke the argument: let m be a positive integer. Assume that $t(x, \vec{z})$ is a cut-and-paste term in symmetric cluster form at m. The procedure below will be called *quantifier tracing* and will be used whenever we have a term in symmetric cluster form at m or whenever we want to use a statement about the term which involves many quantifiers. Fix $d \geq d(t), \epsilon: 0 < \epsilon \leq 1/(2 \cdot d + 2)$. Get $\mu, \nu_0, \delta: 0 < \delta \leq \epsilon$. Fix $\nu \geq \nu_0, e \geq e(t)$ and get ℓ_0. Fix

$\ell \geq \ell_0$, $\rho \in G\text{-Restr}_\epsilon(d, e, \nu, \ell)$ and get $\rho' \in G\text{-Restr}_\delta(d, e, \nu, \ell), \rho' \subseteq \rho$. Fix $\vec{c} < \ell^m$ and get $\vec{a}^{(\vec{c})}$ a sequence of no more than μ bits, $\vec{f}^{(\vec{c})} \in \text{BOOL}_\mu^{\nu_0}$. For each $j < \nu_0$ get $\alpha_{\vec{c},j} \in \mathbf{Z}$, $\psi_{\vec{c},j}$ a positive integer, and $A_j^{\vec{c}} \subseteq \mathcal{H}(d(t)\colon A_{d,e,\nu,\ell})$ and for each $b \in A_j^{\vec{c}}$ get $\varphi_{\vec{c},j,b}$ a positive integer such that:

1. $|\psi_{\vec{c},j} \cdot \varphi_{\vec{c},j,b}| \leq \lceil \log_2 \ell \rceil^{e(t)}$, and
2. $\forall\, x \in \text{vol}(\rho')$

$$t(x, \vec{c}) = \sum_{j<\nu_0} 2^{\alpha_{\vec{c},j}} \cdot \psi_{\vec{c},j} \cdot f_j^{\vec{c}}(x[\vec{a}^{(\vec{c})}]) \cdot \sum_{a \in A_j^{\vec{c}}} \varphi_{\vec{c},j,a} \cdot 2^{\pi a},$$

where $x = \pi(\sigma_{d,e,\nu,\ell})$.

Notice that in the case above we traced the quantifiers to the end. Occasionally, however, we only trace up to a certain quantifier. If such is the case we shall trace quantifiers as we did in the above, stop at the required quantifier and then write down the rest of the property *with its quantification intact.* We indicate the depth of the trace by saying that *we trace quantifiers upto* $\mathbf{z}$, where $\mathbf{z}$ is the variable(s) after which we preserve quantification. Next, we let:

$A = A_{d,e,\nu,\ell}/(2 \cdot \nu \cdot \lceil \log_2 \ell \rceil^e)$, which is the $\Delta(G)$-wide set for d obtained from the $\Delta(G)$-wide lemma;

$R = \{|\alpha_i - \alpha_j| \mid i, j < \nu_0\}$, and

$\tilde{R} = \{\lfloor r/(2 \cdot \nu \cdot \lceil \log_2 \ell \rceil^e) \rfloor + \beta \mid r \in R,\ \beta = 0, 1\}$.

By the finite support corrolary and noting that $|\tilde{R}| \leq 2 \cdot |R| \leq 2 \cdot (\nu_0)^2$, we can find a set $\tilde{S} \subset A$ with $|\tilde{S}| \leq 2 \cdot d \cdot 2 \cdot (\nu_0)^2$ such that if $a, b \in \mathcal{H}(d\colon A)$ and $a - b = r \in \tilde{R}$ then $\text{supp}(a_r, b_r) \subseteq \tilde{S}$. Note further that:

$\mathcal{H}(d\colon A_{d,e,\nu,\ell}) = 2 \cdot \nu \cdot \lceil \log_2 \ell \rceil^e \cdot \mathcal{H}(d\colon A)$, and

$\text{supp}(a) = 2 \cdot \nu \cdot \lceil \log_2 \ell \rceil^e \cdot \text{supp}(a/(2 \cdot \nu \cdot \lceil \log_2 \ell \rceil^e))$, for $a \in \mathcal{H}(d\colon A_{d,e,\nu,\ell})$.

Put $S = 2 \cdot \nu \cdot \lceil \log_2 \ell \rceil^e \cdot \tilde{S} \subset A_{d,e,\nu,\ell}$. We call S the *critical set for* t, the various parameters being understood. Call $x \in \text{vol}(\rho')$ $(t\colon \tau, \xi)$-*good*, with $\xi \in \{0,1\}^{\nu_0}$ and $\tau \in G^S$, iff

1. $1 = [\![\bigwedge_{j<\nu_0}[\xi(j) = f_j(x[\vec{a}])]]\!]$, and

2. $\tau(\sigma_{d,e,\nu,\ell}) \subset x$.

For $\tau \in G^S$ let $\rho'|\tau = \{\pi \in \rho' \mid \tau \subset \pi\}$. Then call $\pi \in \rho'|\tau$, $(t{:}\,\tau, \xi)$-*good* if $x = \pi(\sigma_{d,e,\nu,\ell})$ is $(t{:}\,\tau, \xi)$-good. Recall the action og G-words on members of $\mathcal{H}(d{:}\,A_{d,e,\nu,\ell})$. We show that for $a, b \in \mathcal{H}(d{:}\,A_{d,e,\nu,\ell})$ if $\pi a + \alpha_i - (\pi b + \alpha_j) \leq 2 \cdot \lceil \log_2 \ell \rceil^e$ for *some* $\pi \in \rho'|\tau$ then $\pi a - \pi b = \pi' a - \pi' b$ for *any* other $\pi' \in \rho'|\tau$. Choose $\pi \in \rho'|\tau$ with the property of the assumption, if possible. Note that

$$\begin{aligned} & \pi a + \alpha_i - (\pi b + \alpha_j) \leq 2 \cdot \lceil \log_2 \ell \rceil^e \\ \Longrightarrow \quad & \Big| |\pi a - \pi b| - |\alpha_i - \alpha_j| \Big| \leq 2 \cdot \lceil \log_2 \ell \rceil^e \\ \Longrightarrow \quad & \Big| |a - b| - |\alpha_i - \alpha_j| \Big| \leq d + 2 \cdot \lceil \log_2 \ell \rceil^e \\ \Longrightarrow \quad & \Big| |\tilde{a} - \tilde{b}| - (r' + \eta) \Big| \leq 1/\nu + d/(2 \cdot \nu \cdot \lceil \log_2 \ell \rceil^e) \end{aligned}$$

where $0 \leq \eta < 1$, $r' = \lfloor |\alpha_i - \alpha_j|/(2 \cdot \nu \cdot \lceil \log_2 \ell \rceil^e) \rfloor$, $r' + \eta = |\alpha_i - \alpha_j|/(2 \cdot \nu \cdot \lceil \log_2 \ell \rceil^e)$, $\tilde{a} = a/(2 \cdot \nu \cdot \lceil \log_2 \ell \rceil^e)$, and $\tilde{b} = b/(2 \cdot \nu \cdot \lceil \log_2 \ell \rceil^e)$. It follows that

$$(r' + \eta) - (1/\nu + d/(2 \cdot \nu \cdot \lceil \log_2 \ell \rceil^e)) \leq |\tilde{a} - \tilde{b}| \leq (r' + \eta) + (1/\nu + d/(2 \cdot \nu \cdot \lceil \log_2 \ell \rceil^e))$$

so

$$r' - \frac{1}{2} \leq |\tilde{a} - \tilde{b}| \leq r' + 1 + \frac{1}{2} .$$

We conclude that because $r = |\tilde{a} - \tilde{b}|$ is an integer, $r \in \tilde{R}$. So $\mathrm{supp}(\tilde{a}_r, \tilde{b}_r) \subseteq \tilde{S}$. Recall that $\tilde{a} = \tilde{a}_r + r(\tilde{a}, \tilde{b})$ and $\tilde{b} = \tilde{b}_r + r(\tilde{a}, \tilde{b})$, so that with $r(a, b) = 2 \cdot \nu \cdot \lceil \log_2 \ell \rceil^e \cdot r(\tilde{a}, \tilde{b})$, $a_r = 2 \cdot \nu \cdot \lceil \log_2 \ell \rceil^e \cdot \tilde{a}_r$ and $b_r = 2 \cdot \nu \cdot \lceil \log_2 \ell \rceil^e \cdot \tilde{b}_r$ we have that $a = a_r + r(a, b)$, $b = b_r + r(a, b)$, and $\mathrm{supp}(a_r, b_r) = 2 \cdot \nu \cdot \lceil \log_2 \ell \rceil^e \cdot \mathrm{supp}(\tilde{a}_r, \tilde{b}_r) \subseteq S = \mathrm{dom}(\tau) \subset \mathrm{unfree}(\rho'|\tau)$. Finally, let π' be any other G-word, $\pi' \in \rho'|\tau$.

$$\pi' a - \pi' b = \pi' a_r - \pi' b_r = \tau a_r - \tau b_r = \pi a_r - \pi b_r = \pi a - \pi b.$$

With π a G-word and $\vec{g}$ a sequence of pairs $(a + \alpha_j, j)$, we denote by $\pi\vec{g}$ the sequence obtained by replacing in $\vec{g}$ the pair $(a + \alpha_j, j)$ by the pair $(\pi a + \alpha_j, j)$. Write t_π for $t(\pi(\sigma_{d,e,\nu,\ell}), \vec{c})$, π a G-word. Put $\mathrm{precl}(\pi) = \{\vec{g} \mid \pi\vec{g} \text{ is a precluster in } t_\pi\}$ and $t_{\pi,\vec{g}} = \sum_{(a+\alpha_j, j) \in \vec{g}} \varphi_{j,a} \cdot \psi_j \cdot \xi(j) \cdot 2^{\pi a + \alpha_j}$. The following are the conclusions of the argument for $(t{:}\,\tau, \xi)$-good G-words $\pi, \pi' \in \rho'$:

$1(\tau,\xi)$. $\pi\vec{g}$ is a precluster of $t_{\pi,\vec{g}}$ if and only if $\pi'\vec{g}$ is a precluster of $t_{\pi',\vec{g}}$;

$2(\tau,\xi)$. there are positive integers β and β' such that $t_{\pi,\vec{g}}\cdot 2^{\beta} = t_{\pi',\vec{g}}\cdot 2^{\beta'}$, and

$3(\tau,\xi)$. $\mathrm{precl}(\pi) = \mathrm{precl}(\pi')$.

Set $\mathrm{precl}(t{:}\,\tau,\xi) = \mathrm{precl}(\pi)$ for some $(t{:}\,\tau,\xi)$-good G-word in ρ'. By $(3(\tau,\xi))$, $\mathrm{precl}(t{:}\,\tau,\xi)$ is well defined for any $\tau \in G^{\tilde{S}(t)}$ and $\xi \in \{0,1\}^{\nu_0}$. Call $\mathrm{precl}(t{:}\,\tau,\xi)$ the *cluster set for* t, with the parameters understood. Clearly then

$4(\tau,\xi)$. $t_\pi = \sum\limits_{\vec{g}\in \mathrm{precl}(t:\tau,\xi)} t_{\pi,\vec{g}}$

for any $(t{:}\,\tau,\xi)$-good $\pi \in \rho'$.

Structural lemmata

The lemma below should constitute the two corresponding cases in the induction taking place in the symmetric cut-and-paste cluster lemma. However, it is proved here because it will be used in the forthcoming symmetric cut-and-paste cardinality lemma. The proof uses the symmetric cluster-to-finite and G-multicircuit lemmata.

Symmetric cut-and-paste length&smash cluster Lemma 7.4. *Let m be a positive integer and $t(x,\vec{z})$, $s(x,\vec{z})$ two cut-and-paste term in symmetric cluster form at m. Then,*
$\forall\, d \geq d(t\#s), \epsilon{:}\, 0 < \epsilon \leq 1/(2\cdot d+2)$;
$\exists\, \mu, \nu_0, \delta{:}\, 0 < \delta \leq \epsilon$;
$\forall\, \nu \geq \nu_0, e \geq e(t\#s) \exists\, \ell_0 \forall\, \ell \geq \ell_0$;
$\forall\, \rho \in G\text{-}Restr_\epsilon(d,e,\nu,\ell)\, \exists\, \rho' \in \mathrm{Restr}_\delta(d,e,\nu,\ell)$
$\forall\, \vec{c} < \ell^m$
$\exists\, \vec{g}^{(\vec{c})}, \vec{f}^{(\vec{c})} \in \mathrm{BOOL}_\mu^{\nu_0}$;
$\exists\, \vec{a}^{(\vec{c})} < \ell$, *of* $\leq \mu$ *bits*;
$\exists\, \vec{\phi}^{(\vec{c})}, \vec{\psi}^{(\vec{c})} < \|t\|_{\ell,m}$;
$\forall\, x \in \mathrm{vol}(\rho')$,

$$|t(x,\vec{c})| = \vec{\phi}^{(\vec{c})} \bullet \vec{f}^{(\vec{c})} = \sum_{j<\nu_0} \phi_j^{(\vec{c})} \cdot f_j^{(\vec{c})}(x[\vec{a}^{(\vec{c})}])$$

and

$$t(x,\vec{c})\#s(x,\vec{c}) = \sum_{j<\nu_0} 2^{\psi_j^{(\vec{c})}} \cdot g_j^{(\vec{c})}(x[\vec{a}^{(\vec{c})}]) \ .$$

Proof. Apply the symmetric cluster-to-finite lemma (which is proved below amongst the circuit lemmata) to t and s simultaneously. Now, the proof proceeds in exactly the same way as the cluster length and smash sublemmata of the cut-and-paste cluster lemma of the previous chapter. The only difference is that here we use the multicircuit G-lemma in place of the multicircuit lemma.□

Here is the analog of the cluster lemma of chapter 6. It gives the cut-and-paste terms their desired form.

Symmetric Cut-and-paste Cluster Lemma 7.5. *Let m be a positive integer. Then, all cut-and-paste terms, $t(x,\vec{z})$, are in symmetric cluster form at m.*

Proof. The proof proceeds by induction on the complexity of the cut-and-paste term $r(x,\vec{z})$.

Case: x, z_i, c. In all of these cases $d(t) = 1$ and $e(t) = 2$. So, fix $d \geq 1$, ϵ with $0 < \epsilon \leq 1/(d+1)$. Set $\mu = 0, \nu_0 = 1, \delta = \epsilon$. Fix $e, \nu \geq 1$. Find $\ell_0 > 0$ such that $\lceil \log_2 \ell \rceil^2 > m \cdot \lceil \log_2 \ell \rceil$ for any $\ell \geq \ell_0$. Fix $\ell \geq \ell_0$ and $\rho \in G\text{-Restr}\epsilon d, e, \nu, \ell$. Set $\rho' = \rho$. For $\vec{c} < \ell^m$ we set:

1. $\vec{a}^{(\vec{c})} = \emptyset$ and $\vec{f}^{(\vec{c})} = \langle 1 \rangle$;
2. $\psi_{\vec{c},0} = \begin{cases} c_i, & \text{if } t(x,\vec{z}) = z_i; \\ c, & \text{if } t(x,\vec{z}) = c; \text{ and } \alpha_{\vec{c},0} = 0; \\ 1, & \text{if } t(x,\vec{z}) = x. \end{cases}$
3. $A_0^{\vec{c}} = \begin{cases} \emptyset, & \text{if } t(x,\vec{z}) = z_i \text{ or } t(x,\vec{z}) = c; \\ A_{d,e,\nu,\ell}, & \text{if } t(x,\vec{z}) = x. \end{cases}$, and $\varphi_{\vec{c},0,a} = 1$ for any $a \in A_0^{\vec{c}}$.

After the appropriate substitutions and observing the convention $\sum \emptyset = 1$ we see that for any $x \in \mathrm{vol}(\rho')$,

$$t(x,\vec{c}) = 2^{\alpha_{\vec{c},0}} \cdot \psi_{\vec{c},0} \cdot f_0^{\vec{c}}(x[\vec{a}^{(\vec{c})}]) \cdot \sum_{a \in A_0^{\vec{c}}} \varphi_{\vec{c},0,a} \cdot 2^{\pi a},$$

where $x = \pi \sigma_{d,e,\nu,\ell}$, $\pi \in \rho'$. Finally, for any $b \in A_0^{\vec{c}}$, $|\psi_{\vec{c},0} \cdot \varphi_{\vec{c},0,b}| \leq m \cdot \lceil \log_2 \ell \rceil < \lceil \log_2 \ell \rceil^{e(t)}$, by the choice of ℓ_0.

Case: $|t(x,\vec{z})|$. cf. symmetric cut-and-paste length&smash lemma.

Case: $t(x,\vec{z}) \# s(x,\vec{z})$. cf. symmetric cut-and-paste length&smash lemma.

Case: $t(x,\vec{z}) + s(x,\vec{z})$. Proceed exactly as in the cluster + sublemma.

Case: $t(x,\vec{z})\cdot s(x,\vec{z})$. Proceed exactly as in the cluster $\times$ sublemma. Then,

$$t(x,\vec{c})\cdot s(x,\vec{c}) = \sum_{\langle i,j\rangle\in\nu_{t,0}\times\nu_{s,0}} 2^{\alpha_{\langle i,j\rangle}}\cdot\psi_{\langle i,j\rangle}\cdot f_{\langle i,j\rangle}(x[\vec{a}])\cdot\sum_{a\in A_{\langle i,j\rangle}}\varphi_{\langle i,j\rangle,a}\cdot 2^{\pi a},$$

where:

1. $A_{\langle i,j\rangle} = A_i^{t,\vec{c}} + A_j^{s,\vec{c}}$, $\psi_{\langle i,j\rangle} = \psi_{t,\vec{c},i}\cdot\psi_{s,\vec{c},j}$, $\vec{a} = \vec{a}^{\,(t,\vec{c})}\cup\vec{a}^{\,(s,\vec{c})}$;
2. $\varphi_{\langle i,j\rangle,a} = \sum\{\varphi_{t,\vec{c},i,b}\cdot\varphi_{s,\vec{c},j,a} \mid b\in A_i^{t,\vec{c}}, b'\in A_j^{s,\vec{c}}, a = b+b'\}$, and
3. $f_{\langle i,j\rangle} = f_i^{t,\vec{c}}\cdot f_j^{s,\vec{c}}$; .

We now claim that,

$$|\varphi_{\langle i,j\rangle,a}| \le \lceil\log_2\ell\rceil^{e(t)+e(s)+1}.$$

It suffices to show that the bound on the cardinality of

$$\overset{\circ}{a} = \{(b,b') \mid b\in A_i^{t,\vec{c}}, b'\in A_j^{s,\vec{c}}, a = b+b'\}$$

depends only on $d(r)$ $(= d(s)+d(t))$. Now,

$$A_i^{t,\vec{c}} + A_j^{s,\vec{c}} \subseteq \mathcal{H}(d(s)+d(t)\colon A_{d,e,\nu,\ell}).$$

So, if $(b,b')\in\overset{\circ}{a}$ then $\mathrm{supp}(b,b') = \mathrm{supp}(a)$. $|\mathrm{supp}(a)|\le d(r)$ and the coeficients in b sum to $\le d(t)$, and the coefficients in b' sum to $\le d(s)$. The claim on the bound of the cardinality of the set $\overset{\circ}{a}$ is now evident. Finally, by the definition of the action of G-words and the fact that $b+b' = a$ implies that $\mathrm{supp}(b,b') = \mathrm{supp}(a)$ we conclude that

$$\pi a = \pi b + \pi b'.$$

Case: $\mathrm{trunc}(t(x,\vec{z})\colon l(x,\vec{z}), r(x,\vec{z}))$. We proceed as in the cluster truncate sublemma. Here, however, in order to partition the critical cluster we need to condition $x\in\mathrm{vol}(\rho'')$ further. This arises because clusters shift as π varies in ρ''. We redefine the local notions of *relevance* and v-*goodness*, for quadruples v. Let:

1. $S = \bigcup_{(a+\alpha_{s,\vec{c},j},j)\in\vec{g}_s}\mathrm{supp}(a)$;
2. $\tau_s\in G^S$, $\xi_s\in\{0,1\}^{\nu_{s,0}}$;

3. $\tau_t \in G^X$, X a set of no more than $\mu_t \cdot \nu_{t,0}$ bits, $\xi_t \in \{0,1\}^{\nu_{t,0}}$;

4. $\mathrm{v} = \langle \tau_s, \xi_s, \tau_t, \xi_t \rangle$

5. $s(\mathrm{v}) = \sum_{(a+\alpha_{s,\vec{c},j},j)\in\vec{g}_s} \varphi_{s,\vec{c},j,a} \cdot \psi_{s,\vec{c},j} \cdot \xi_s(j) \cdot 2^{\tau_s a+\alpha_{s,\vec{c},j}}$;

6. $\vec{g}(\mathrm{v}) =$ the (unique) precluster of t, $\vec{g}$, of which the range contains $s(\mathrm{v})$ if such a precluster exists, and any precluster of t otherwise, and

7. $T = \bigcup_{(a+\alpha_{t,\vec{c},j},j)\in\vec{g}(\mathrm{v})} \mathrm{supp}(a)$.

Call quadruple $\mathrm{v} = \langle \tau_s, \xi_s, \tau_t, \xi_t \rangle$ *relevant* if:

1. $\xi_s \in \{0,1\}^{\nu_{s,0}}$, $\xi_t \in \{0,1\}^{\nu_{t,0}}$, $\tau_s \in G^S$, $\tau_t \in G^T$, and

2. $\tau_s \cup \tau_t \in G^{S \bigcup T}$, i.e. the two functions are compatible.

Call $x \in \mathrm{vol}(\rho'')$ v-*good* if x is both $(s{:}\,\tau_s, \xi_s)$-good and $(t{:}\,\tau_t, \xi_t)$-good. To verify v-goodness requires

$$\kappa = \mu_t \cdot (1 + \nu_{t,0}) + \mu_s \cdot (1 + \nu_{s,0})$$

bits, call them $\vec{a}^{(\vec{c},\mathrm{v})}$, so that the verification of all properties requires

$$\mu = \kappa \cdot 2^{\kappa}$$

bits; this is because the number of relevant v's is at most 2^{κ}.
The generating cases have now been considered and the induction is complete.□

Circuit lemmata

The lemma below is the analogue of the cluster-to-finite lemma of chapter 6.

Symmetric Cluster-to-finite Lemma 7.6. *Let m be a positive integer. Every cut-and-paste term $t(x, \vec{z})$ in symmetric cluster form at m has the symmetric bit-finite property at m.*

Proof. Let m be a positive integer and $t(x, \vec{z})$ be a cut-and-paste term in symmetric cluster form at m. Trace the quantifiers in the definition of symmetric cluster form at m so

that with fixed $\vec{c} < \ell^m$ we have that for any $x \in \mathrm{vol}(\rho')$,

$$t(x, \vec{c}) = \sum_{j<\nu_0} 2^{\alpha_{\vec{c},j}} \cdot \psi_{\vec{c},j} \cdot f_j^{\vec{c}}(x[\vec{a}^{(\vec{c})}]) \cdot \sum_{a \in A_j^{\vec{c}}} \varphi_{\vec{c},j,a} \cdot 2^{\pi a},$$

where $x = \pi\sigma_{d,e,\nu,\ell}$. Let $\mathrm{precl}(t)$ be the set of preclusters in $t(\sigma_{d,e,\nu,\ell}, \vec{c})$. Preclusters will shift with $x \in \mathrm{vol}(\rho)'$ but not far enough to cause bit interference amongst adjacent preclusters. We then have that for $(t{:}\,\xi)$-good (cf. cluster-to-finite lemma) $x \in \mathrm{vol}(\rho)$,

$$\begin{aligned} t(x, \vec{c}) &= \sum_{\substack{\vec{g} \in \mathrm{precl}(t) \\ (a+\alpha_j, j) \in \vec{g}}} \varphi_{j,a} \cdot \psi_j \cdot \xi(j) \cdot 2^{\pi a + \alpha_j} \\ &= \sum_{\vec{g} \in \mathrm{precl}(t)} t_{\pi,\vec{g}} \end{aligned}$$

with $x = \pi(\sigma_{d,e,\nu,\ell})$ and the $t_{\pi,\vec{g}}$ non-interfering, at a distance $> 2 \cdot \lceil \log_2 \ell \rceil^e - 2 \cdot d$ from each other. Furthermore, with $\mathrm{supp}(\vec{g}) = \bigcup_{(a+\alpha_j,j)} \mathrm{supp}(a)$, $\tau \in G^S$, $\xi \in \{0,1\}^{\nu_0}$, and $(t{:}\,\tau, \xi)$-good $x \in \mathrm{vol}(\rho)$

$$\mathbf{BIT}(t_{\pi,\vec{g}}, y) \iff \mathbf{BIT}\Big(\sum_{(a+\alpha_j,j) \in \vec{g}} \varphi_{j,a} \cdot \psi_j \cdot \xi(j) \cdot 2^{\tau a + \alpha_j}, y\Big).$$

We conclude that

$$\begin{aligned} &\mathbf{BIT}(t(x, \vec{c}), y) \iff \\ &\iff \bigvee_{\substack{\vec{g} \in \mathrm{precl}(t) \\ \tau \in G^{\mathrm{supp}(\vec{g})} \\ \xi \in \{0,1\}^{\nu_0}}} [(x \text{ is } (t{:}\,\tau, \xi)\text{-good}) \wedge \mathbf{BIT}\Big(\sum_{(a+\alpha_j,j) \in \vec{g}} \varphi_{j,a} \cdot \psi_j \cdot \xi(j) \cdot 2^{\tau a + \alpha_j}, y\Big)]. \end{aligned}$$

Since the $\varphi_{j,a}$ and ψ_j are fixed independently of $x \in \mathrm{vol}(\rho')$, the circuit on the righthand side of the last equivalence has depth ≤ 4 and size $\leq \|t\|_{\ell,m}$.□

The analogue for terms of bounded arithmetic was proved in chapter 5. It is for the proof of *this* lemma that we established the cluster shift argument.

Cut-and-paste Cardinality Lemma 7.7. *Let m be a fixed positive integer. If two cut-and-paste terms $t(x, \vec{z})$ and $s(x, \vec{z})$ are in symmetric cluster form at m then,*

$\forall\, d \geq d(t+s), \epsilon{:}\, 0 < \epsilon \leq 1/(2 \cdot d + 2)$;

$\exists\, \mu, \nu_0, \delta{:}\, 0 < \delta \leq \epsilon$;

$\forall\, \nu \geq \nu_0, e \geq e(t+s) \exists\, \ell_0 \forall\, \ell \geq \ell_0$;

$\forall\, \rho \in G\text{-}Restr_\epsilon(d,e,\nu,\ell) \exists\, \rho' \in G\text{-}Restr_\delta(d,e,\nu,\ell), \rho' \subseteq \rho,$

$\forall\, \vec{c} < \ell^m;$

$\exists\, \vec{a}^{(\vec{c})}$ *a sequence of* $\leq \mu$ *bits in* $A_{d,e,\nu,\ell};$

$\exists\, f^{(\vec{c})} \in \mathrm{BOOL}_\mu;$

$\forall\, x \in \mathrm{vol}(\rho'),$

$$[\![\mathbf{CARD}(t(x,\vec{c}), s(x,\vec{c}))]\!] = f^{(\vec{c})}(x[\vec{a}^{(\vec{c})}])\ .$$

Proof. Following the blueprint for two terms (cf. cluster + sublemma) and using the symmetric cut-and-paste length&smash lemma find $\rho'' \in G\text{-}\mathrm{Restr}_{\delta'}(d,e,\nu,\ell)$ such that for $\vec{c} < \ell^m$ it is true that

$\exists\, \vec{a}^{(t,\vec{c})}, \vec{a}^{(s,\vec{c})}$, sequences of no more than μ_t and μ_t bits, respectively;

$\exists\, \vec{f}^{(t,\vec{c})} \in \mathrm{BOOL}_{\mu_t}^{\nu_{t,0}}$ and $\vec{f}^{(s,\vec{c})} \in \mathrm{BOOL}_{\mu_s}^{\nu_{s,0}}$;

$\forall\, j < \nu_{t,0}\,, j' < \nu_{s,0};$

$\exists\, \alpha_{t,\vec{c},j}\,, \alpha_{s,\vec{c},j'} \in \mathbf{Z};$

$\exists\, \psi_{t,\vec{c},j}\,, \psi_{s,\vec{c},j'}$ positive integers;

$\exists\, A_j^{t,\vec{c}} \subseteq \mathcal{H}(d(t)\colon A_{d,e,\nu,\ell}), A_{j'}^{s,\vec{c}} \subseteq \mathcal{H}(d(s)\colon A_{d,e,\nu,\ell});$

$\forall\, b \in A_j^{t,\vec{c}}, b' \in A_{j'}^{s,\vec{c}};$

$\exists\, \varphi_{t,\vec{c},j,b}\,, \varphi_{s,\vec{c},j',b'}$ positive integers;

such that:

1. $|\psi_{t,\vec{c},j} \cdot \varphi_{t,\vec{c},j,b}| \leq \lceil \log_2 \ell \rceil^{e(t)};$
2. $|\psi_{s,\vec{c},j'} \cdot \varphi_{s,\vec{c},j',b'}| \leq \lceil \log_2 \ell \rceil^{e(s)}$, and
3. $\forall\, x \in \mathrm{vol}(\rho'')$,

$$t(x,\vec{c}) = \sum_{j<\nu_{t,0}} 2^{\alpha_{t,\vec{c},j}} \cdot \psi_{t,\vec{c},j} \cdot f_j^{t,\vec{c}}(x[\vec{a}^{(t,\vec{c})}]) \cdot \sum_{a \in A_j^{t,\vec{c}}} \varphi_{t,\vec{c},j,a} \cdot 2^{\pi a}$$

and

$$s(x,\vec{c}) = \sum_{j'<\nu_{s,0}} 2^{\alpha_{s,\vec{c},j'}} \cdot \psi_{s,\vec{c},j'} \cdot f_{j'}^{s,\vec{c}}(x[\vec{a}^{(s,\vec{c})}]) \cdot \sum_{a' \in A_{j'}^{s,\vec{c}}} \varphi_{s,\vec{c},j',a'} \cdot 2^{\pi a'}$$

where $x = \pi\sigma_{d,e,\nu,\ell}, \pi \in \rho''$

and

$$|s(x,\vec{c})| = \vec{\chi}^{(\vec{c})} \bullet \vec{h}^{(\vec{c})}(x[\vec{b}^{(\vec{c})}]),$$

where $\vec{\chi}^{(\vec{c})}$ has no more than $\nu_{s,0}$ members $< \|s\|_{\ell,m}$ and $\vec{h}^{(\vec{c})} \in \mathrm{BOOL}_{\mu_s}^{\nu_{s,0}}$ is a partition of unity. Set:

1. $\mathrm{small}(x) = \bigvee_{\chi_i < \lceil \log_2 \|t\|_{\ell,m} \rceil} h_i^{\vec{c}}(x[\vec{b}^{(\vec{c})}])$;

2. $\vec{g}_s$ the first precluster of $s(x, \vec{c})$;

3. $S(s) = \bigcup_{(a+\alpha_{s,\vec{c},j},j) \in \vec{g}_s} \mathrm{supp}(a)$, and

4. $s(\tau_s, \xi_s) = \sum_{(a+\alpha_{s,\vec{c},j},j) \in \vec{g}_s} \varphi_{s,\vec{c},j,a} \cdot \psi_{s,\vec{c},j} \cdot \xi_s(j) \cdot 2^{\tau_s a + \alpha_{s,\vec{c},j}}$, for $\tau_s \in G^{S(s)}$ and $\xi_s \in \{0,1\}^{\nu_{s,0}}$

$s(x, \vec{c}) = s(\tau_s, \xi_s)$ holds for any $(s{:}\,\tau_s, \xi_s)$-good(cf. cluster shift argument) $x \in \mathrm{vol}(\rho'')$ such that $\mathrm{small}(x)$ is true. It is now clear that for $x \in \mathrm{vol}(\rho'')$,

$$\mathbf{CARD}(t(x, \vec{c}), s(x, \vec{c})) \Longleftrightarrow$$
$$\Longleftrightarrow \bigvee_{\substack{\tau_s \in G^{S(s)} \\ \xi_s < \{0,1\}^{\nu_{s,0}}}} \mathrm{small}(x) \wedge (x \text{ is } (s{:}\,\tau_s, \xi_s)\text{-good}) \wedge \mathbf{CARD}(t(x, \vec{c}), s(\tau_s, \xi_s)).$$

Next we reduce $\mathbf{CARD}(t(x, \vec{c}), s(\tau_s, \xi_s))$ to boolean function of a bounded number of bits of x, given that x is $(s{:}\,\tau_s, \xi_s)$-good. Put $\nu_0 = \nu_{t,0}$, $A_j = A_j^{(t,\vec{c})}$, $\alpha_j = \alpha_{t,\vec{c},j}$, $\vec{a} = \vec{a}^{(t,\vec{c})}$, $f_j = f_j^{(t,\vec{c})}$, $\varphi_{j,a} = \varphi_{t,\vec{c},j,a}$, $\psi_j = \psi_{t,\vec{c},j}$ for $j < \nu_0$. Now use the cluster shift argument on t with $S(t)$ the critical set for t. By conclusions $(1(\tau,\xi))$, $(2(\tau,\xi))$, and $(3(\tau,\xi))$ of the cluster shift argument the following hold for $(t{:}\,\tau, \xi)$-good $\pi, \pi' \in \rho''$:

1′. $t_\pi = \sum_{\vec{g} \in \mathrm{precl}(\pi)} t_{\pi,\vec{g}}$;

2′. $\mathrm{card}(t_{\pi,\vec{g}}) = \mathrm{card}(t_{\pi',\vec{g}})$, and

3′. $\mathrm{card}(t_\pi) = \mathrm{card}(t_{\pi'})$.

So, given $\tau \in G^{S(t)}, \xi \in \{0,1\}^{\nu_0}$ it now makes sense to talk about $\mathbf{c}(\tau, \xi)$ =the cardinality of t_π, for $(t{:}\,\tau, \xi)$-good $\pi \in \rho'$. So for $x \in \mathrm{vol}(\rho'')$ we have,

$$\mathbf{CARD}(t(x, \vec{c}), s(\tau_s, \xi_s)) \Longleftrightarrow \bigvee_{\substack{\tau \in G^{S(t)} \\ \xi < \{0,1\}^{\nu_0}}} [(x \text{ is } (t{:}\,\tau, \xi)\text{-good}) \wedge (\mathbf{c}(\tau, \xi) = s(\sigma, \xi_s))].$$

Clearly now the lemma follows.□

This final lemma is the key for establishing the theorem. The proof uses the symmetric cut-and-paste cluster, cluster-to-finite, and cut-and-paste cardinality lemmata.

Symmetric Cut-and-paste Circuit Lemma 7.8. *Let* $\Theta(x)$ *be a symmetric cut-and-paste sharply bounded formula. Then,*
$\exists\, N, m_0, \nu, d, \ell_0, \delta : 0 < \delta \leq 1;$
$\forall \ell \geq \ell_0;$
$\exists \rho \in G\text{-}Restr_\delta(d, e, \nu, \ell);$
$\exists C^{\Theta}_{\ell,\rho}$ *a circuit of size* $\leq \ell^{m_0}$ *and depth* $\leq N + 6;$
$\forall x \in \mathrm{vol}(\rho),$

$$[\![\Theta(x)]\!] = C^{\Theta}_{\ell,\rho}(x).$$

Proof. $\Theta(x)$ is *logically* equivalent to

$$Q_0 z_0 < |x|^m \ldots Q_{N-1} z_{N-1} < |x|^m \Psi(x, \vec{z})$$

for some nonzero positive integers N and m, where Q_i are quantifiers, for $i < N$, $\vec{z} = z_0 \ldots z_{N-1}$, and $\Psi(x, \vec{z})$ a quantifier free symmetric cut-and-paste formula. By the symmetric cut-and-paste cluster lemma, all terms in Ψ are in symmetric cluster form at m. We therefore can proceed exactly as in the proof of the boolean cut-and-paste circuit lemma. By the cut-and-paste cardinality lemma all occurrences of **CARD** have the property in the cut-and-paste cardinality lemma. By the properties of the first five quantifiers in the definition of the symmetric bit-finite property and in the property of the cut-and-paste cardinality lemma we get that,
$\exists \nu, m_0, d, \ell_0, \delta : 0 < \delta \leq 1;$
$\forall \ell \geq \ell_0;$
$\exists \rho \in G\text{-}\mathrm{Restr}_\delta(d, e, \nu, \ell);$
$\forall \vec{c} < \ell^m$
$\exists C^{\Psi}_{\ell,\vec{c},\rho}$ a circuit of size $\leq \ell^{m_0}$ and depth $\leq 6;$
$\forall x \in \mathrm{vol}(\rho),$

$$[\![\Psi(x, \vec{c})]\!] = C^{\Psi}_{\ell,\vec{c},\rho}(x).$$

So get the appropriate δ and G-restriction ρ so that
$\forall \vec{c} < \ell^m$

$\exists C^{\Psi}_{\ell,\vec{c},\rho}$ a circuit of size $\leq \ell^{m_0}$ and depth ≤ 6;
$\forall x \in \mathrm{vol}(\rho)$,

$$[\![\Psi(x,\vec{c})]\!] = C^{\Psi}_{\ell,\vec{c},\rho}(x).$$

Let

$$C^{\Theta}_{\ell,\rho}(x) = \underset{c_0<\ell^m}{\mathrm{Bool}(Q_0)} \cdots \underset{c_{N-1}<\ell^m}{\mathrm{Bool}(Q_{N-1})}\, C^{\Psi}_{\ell,\vec{c},\rho}(x,\vec{c})$$

with $\vec{c}$ substituted for $\vec{z}$. For $x \in \mathrm{vol}(\rho)$,

$$[\![\Theta(x)]\!] = C^{\Theta}_{\ell,\rho}(x).$$

□

Proof of the theorem

Assume that $\Theta(x)$ is a symmetric cut-and-paste symmetric cut-and-paste. Apply the symmetric cut-and-paste circuit lemma to $\Theta(x)$ to get a G-cylinder, $\rho \in G\text{-Restr}_{\delta}(d,e,\nu,\ell)$ and a circuit C^{Θ}_{ℓ} such that for any $x \in \mathrm{vol}(\rho)$,

$$[\![\Theta(x)]\!] = C^{\Theta}_{\ell}(x).$$

Apply the G-lemma to $\{C^{\Theta}_{\ell}\}^{\infty}_{\ell_0}$ and get δ', $0 < \delta' \leq \delta$, and a finer G-cylinder, $\rho' \subseteq \rho$, $\rho' \in G\text{-Restr}_{\delta'}(d,e,\nu,\ell)$ such that $\Theta(x)$ is oblivious in $\mathrm{vol}(\rho')$. **INDEXPARITY** is *not* oblivious in $\mathrm{vol}(\rho')$. So Θ does not define **INDEXPARITY**. ■

Chapter 8

PARITY is not $\mathcal{SB}_1^{\mathbf{N}}(\mathbf{trunc}, \mathbf{flip})$ definable

In this chapter the main result is

Theorem. **PARITY** *is not ring sharply bounded definable.*

To the cut-and-paste language we add the *flip*. This new term now enables subtraction of terms.

Notation and convention

For items without local definition consult their *most recent* definition:

8a. To the cut-and-paste language we add the term $\text{flip}(x, y)$. The interpretation of the latter is

$$\text{flip}(x, y) = \begin{cases} 2^{|x|} - 1 - y, & \text{if } |x| \geq |y|; \\ 0, & \text{otherwise.} \end{cases}$$

The new sharply bounded formulae are now called *ring sharply bounded.*

8b. Integer subtraction is definable: let x and y be positive integers. Then

$$y \dot{-} x = \text{flip}(y, \text{flip}(y, y) + x).$$

8c. We extend the definitions of $e(t), d(t)$ to include the new function symbol $\text{flip}(x, y)$: $d(\text{flip}(r, s)) = \max\{d(r), d(s)\}$ and $e(\text{flip}(r, s)) = \max\{e(r), e(s)\}$.

8d. Let m be a positive integer and $t(x, \vec{z})$ a ring term. We then say that t *is in ring cluster form at* m iff

$\forall\, d \geq d(t), \epsilon\colon 0 < \epsilon \leq 1/(2 \cdot d + 2)$;

$\exists\, \mu, \nu_0, \delta\colon 0 < \delta \leq \epsilon$;

$\forall\, \nu \geq 2 \cdot \nu_0, e \geq e(t) \exists\, \ell_0 \forall\, \ell \geq \ell_0$;

$\forall\, \rho \in \mathrm{Restr}_\epsilon(d, e, \nu, \ell) \exists\, \rho' \in \mathrm{Restr}_\delta(d, e, \nu, \ell), \rho' \subseteq \rho$;

$\forall\, \vec{c} < \ell^m$;

$\exists\, \vec{a}^{(\vec{c},+)}, \vec{a}^{(\vec{c},-)}$ sequences of $\leq \mu$ bits;

$\exists\, \vec{f}^{(\vec{c},+)}, \vec{f}^{(\vec{c},-)} \in \mathrm{BOOL}_\mu^{\nu_0}$;

$\forall\, j < \nu_0$;

$\exists\, \alpha^+_{\vec{c},j}, \alpha^-_{\vec{c},j} \in \mathbf{Z}$, $\psi^+_{\vec{c},j}, \psi^-_{\vec{c},j}$ positive integers;

$\exists\, A_j^{\vec{c},+}, A_j^{\vec{c},-} \subseteq \mathcal{H}(d(t)\colon A_{d,e,\nu,\ell})$;

$\forall\, b \in A_j^{\vec{c},+}, b' \in A_j^{\vec{c},-}\ \exists\, \varphi^+_{\vec{c},j,b}, \varphi^-_{\vec{c},j,b'}$ positive integers;

such that:

1. $|\psi^+_{\vec{c},j} \cdot \varphi^+_{\vec{c},j,b}|, |\psi^-_{\vec{c},j} \cdot \varphi^-_{\vec{c},j,b'}| \leq \lceil \log_2 \ell \rceil^{e(t)}$;
2. $\forall\, x \in \rho',\ t(x, \vec{c}) = \Gamma_+ - \Gamma_- \geq 0$, where

$$\Gamma_+ = \sum_{j < \nu_0} 2^{\alpha^+_{\vec{c},j}} \cdot \psi^+_{\vec{c},j} \cdot f_j^{\vec{c},+}(x[\vec{a}^{(\vec{c},+)}]) \cdot \sum_{a \in A_j^{\vec{c},+}} \varphi^+_{\vec{c},j,a} \cdot 2^a \cdot \wedge x[\mathrm{supp}(a)]$$

and

$$\Gamma_- = \sum_{j < \nu_0} 2^{\alpha^-_{\vec{c},j}} \cdot \psi^-_{\vec{c},j} \cdot f_j^{\vec{c},-}(x[\vec{a}^{(\vec{c},-)}]) \cdot \sum_{a \in A_j^{\vec{c},-}} \varphi^-_{\vec{c},j,a} \cdot 2^a \cdot \wedge x[\mathrm{supp}(a)].$$

Remark. Quantifier tracing is as in the cluster shift argument and the blueprint for two terms in ring cluster form at m is as in the cluster + sublemma, with the understanding that both procedures are applied to the definition of ring cluster form at m.

The lemmata

The strategy is the same as in chapters 6 and 7: we prove a bit-finite lemma, a cluster lemma, and then the corresponding circuit lemma will follow. The definition of bit-finiteness at m, m a positive integer, is the same as in chapter 6. We thus have,

Ring Cluster-To-Finite Lemma 8.1. *Let m be a positive integer and $t(x, \vec{z})$ a ring term. If $t(x, \vec{z})$ is in ring cluster form then $t(x, \vec{z})$ has the bit-finite property at m.*

Proof. Since $t(x, \vec{z})$ is in ring cluster form we trace the quantifiers in the definition of ring cluster form at m to the end. Recall the definitions of cluster, as they appear cluster-to-finite of chapter 6. Here the clusters (when evaluated) can be negative or zero, while all other relationships hold, as in the cluster-to-finite lemma of chapter 6. Also preclusters may have up to $2 \cdot \nu_0$ entries. For fixed $\vec{c} < \ell^m$ and $y < ||t||_{\ell,m}$ we let:

1. $\vec{g}(y)$ be the (unique) precluster of $t(x, \vec{c})$ the range of which (cf. cluster-to-finite lemma) contains y;
2. $\vec{g}^+(y)$ be the precluster to the left of $\vec{g}(y)$;
3. $\alpha = \min_{(a+\alpha_{t,\vec{c},j}) \in \vec{g}^+(y)} \{a + \alpha_{t,j}\}$, and
4. $\beta = \min_{(a+\alpha_{t,\vec{c},j}) \in \vec{g}(y)} \{a + \alpha_{t,j}\}$.

Recall that given a precluster $\vec{g}$ of t, the cluster based on $\vec{g}$ is

$$t_{\vec{g}}(x) = \sum_{(a+\alpha_{t,\vec{c},j},j) \in \vec{g}} 2^{a+\alpha_{t,\vec{c},j}} \cdot \psi_{t,\vec{c},j} \cdot f_j^{t,\vec{c}}(x[\vec{a}^{(t,\vec{c})}]) \cdot \varphi_{t,\vec{c},j,a} \cdot \wedge x[\mathrm{supp}(a)].$$

Call pairs of the form $\mathrm{v} = \langle \tau, \xi \rangle$, $\tau, \xi \in \{0,1\}^{2 \cdot \nu_{t,0}}$ *relevant.* For a precluster $\vec{g}$ of t and a relevant pair v we let

1. $t_{\vec{g}}(\mathrm{v}) = \sum_{\substack{(a+\alpha_{t,j},j) \\ \in \vec{g}}} 2^{a+\alpha_{t,j}} \cdot \xi_{(j)} \cdot \psi_{t,j} \cdot \varphi_{t,j,a} \cdot \tau_{(j)}$;
2. $V_{\vec{g}}^- = \{\mathrm{v} \mid \mathrm{v} \text{ relevant}, t_{\vec{g}}(\mathrm{v}) < 0\}$;
3. $V_{\vec{g}}^0 = \{\mathrm{v} \mid \mathrm{v} \text{ relevant}, t_{\vec{g}}(\mathrm{v}) = 0\}$, and
4. $V_{\vec{g}}^+ = \{\mathrm{v} \mid \mathrm{v} \text{ relevant}, t_{\vec{g}}(\mathrm{v}) > 0\}$.

For x, $\vec{g}$ a precluster and v relevant we say that x *is* $(\vec{g}\colon \mathrm{v})$*-good* if

1. x is $(t\colon \xi)$-good, and
2. $1 = [\![\bigwedge_{(a+\alpha_{t,\vec{c},j}) \in \vec{g}} (\tau(j) = \wedge x[\mathrm{supp}(a)])]\!]$.

Clearly $[\![x \text{ is } (\vec{g}\colon \mathrm{v})\text{-good}]\!]$ can be calculated by a circuit of constant depth and size, independently of x. Next we calculate some truth values:

1. $[\![t_{\vec{g}}(x) > 0]\!] = [\![\exists\, \mathrm{v} \in V^{+}_{\vec{g}}(x \text{ is } (\vec{g}\colon \mathrm{v})\text{-good})]\!]$
2. $[\![t_{\vec{g}}(x) = 0]\!] = [\![\exists\, \mathrm{v} \in V^{0}_{\vec{g}}(x \text{ is } (\vec{g}\colon \mathrm{v})\text{-good})]\!]$, and
3. $[\![t_{\vec{g}}(x) < 0]\!] = [\![\exists\, \mathrm{v} \in V^{-}_{\vec{g}}(x \text{ is } (\vec{g}\colon \mathrm{v})\text{-good})]\!]$.

Again the above truth values and their negations can be calculated by circuits of constant depth and size, independently of x. Before we express $\mathbf{BIT}(t(x, \vec{c}), y)$ as a constant depth, polynomial size circuit we need to define three auxilliary properties:

1. $Q_{-1}(x, y) = \exists\, \vec{g} \prec \vec{g}(y)(\forall \vec{g}'\colon \vec{g}(y) \succ \vec{g}' \succ \vec{g})\, (t_{\vec{g}}(x) < 0 \wedge t_{\vec{g}'}(x) = 0)$;
2. $Q_{-}(\mathrm{v}, y) = Q_{-1} \longrightarrow \mathbf{BIT}(2^{\alpha} + t_{\vec{g}(y)}(\mathrm{v}) - 2^{\beta}, y)$, and
3. $Q_{+}(\mathrm{v}, y) = \neg Q_{-1} \longrightarrow \mathbf{BIT}(2^{\alpha} + t_{\vec{g}(y)}(\mathrm{v}), y)$.

It is evident that the truth values of the above properties can be calculated by circuits of constant depth and size polynomial in ℓ. Q_{-1} tests whether there is a carry from previous subtractions. Now, for any $x \in \rho'$, $y < \|t\|_{\ell,m}$, and $\vec{c} < \ell^m$ we have that the truth value of $\mathbf{BIT}(t(x, \vec{c}), y)$ can be calculated by a circuit of constant depth and size polynomial in ℓ because

$$\mathbf{BIT}(t(x, \vec{c}), y) \iff \bigvee_{\mathrm{v}\ \text{relevant}} [(x \text{ is } (\vec{g}\colon \mathrm{v})\text{-good}) \wedge Q_{-}(\mathrm{v}, y) \wedge Q_{+}(\mathrm{v}, y)].$$

Finally, identifying $\mathbf{BIT}(t(x, \vec{c}), y)$ with the circuit that calculates its truth value we consider the multicircuit

$$\mathbf{C} = \{\mathbf{BIT}(t(x, \vec{c}), y) \mid\ y < \|t\|_{\ell,m}, \vec{c} < \ell^m\}\,.$$

Applying the multicircuit lemma we get $\rho'' \subset \rho'$, μ', $\vec{a}^{(\vec{c},y)}$ with no more than μ' bits, and $f_{\vec{c},y} \in \mathrm{BOOL}_{\mu'}$ such that

$$\forall\, \vec{c} < \ell^m\, \forall\, y < \|t\|_{\ell,m}\, \forall\, x \in \rho''\, \left[\mathbf{BIT}(t(x, \vec{c}), y) = f_{\vec{c},y}(x[\vec{a}^{(\vec{c},y)}])\right].$$

□

The central lemma of the chapter which, together with the previous lemma, is the key to the proof of the ring circuit lemma.

Ring Cluster Lemma 8.2. *Let m be a positive integer. Every ring term, $t(x, \vec{z})$, is in ring cluster form at m.*

Proof. We proceed by induction on the complexity of the ring term $r(x, \vec{z})$.

Case: x, z_i, c. We proceed exactly as in the base cases sublemma of the cut-and-paste cluster lemma. In addition we require that $\Gamma_- = 0$. This is done by setting $f_0^{\vec{c},-} \equiv 0$, for $\vec{c} < \ell^m$.

Case: $|t(x, \vec{z})|, t(x, \vec{z}) \# s(x, \vec{z})$. Since both $t(x, \vec{z})$ and $s(x, \vec{z})$ are in ring cluster form at m they have the bit-finite property at m, by the ring cluster-to-finite lemma above. The rest of the proof now proceeds exactly as the proof of the cluster length and cluster smash sublemmata of the cut-and-paste cluster lemma and the *conclusions* are the same.

Case: $t(x, \vec{z}) + s(x, \vec{z})$. We have that,

$$t(x, \vec{c}) = \Gamma_+^t - \Gamma_-^t \geq 0$$

and

$$s(x, \vec{c}) = \Gamma_+^s - \Gamma_-^s \geq 0.$$

Clearly now,

$$s(x, \vec{c}) + t(x, \vec{c}) = (\Gamma_+^s + \Gamma_+^t) - (\Gamma_-^s + \Gamma_-^t).$$

Case: $t(x, \vec{z}) \cdot s(x, \vec{z})$. As above, only here we get that,

$$s(x, \vec{c}) \cdot t(x, \vec{c}) = (\Gamma_+^s \cdot \Gamma_+^t + \Gamma_-^s \cdot \Gamma_-^t) - (\Gamma_+^s \cdot \Gamma_-^t + \Gamma_-^s \cdot \Gamma_+^t).$$

The new parameters are obtained exactly as in the proof of the cluster $\times$ sublemma of the cut-and-paste cluster lemma.

Case: $\mathrm{flip}(t(x, \vec{z}), s(x, \vec{z}))$. We use the blueprint for two terms (cf. cluster + sublemma) and the conclusion of the cluster length sublemma (which holds as remarked earlier) to get a restriction ρ'' such that for $x \in \rho''$ we have that,

$$|s(x, \vec{c})| = \vec{\chi}^{(s,\vec{c})} \bullet \vec{h}^{(s,\vec{c})}$$

and

$$|t(x,\vec{c})| = \vec{\chi}^{(t,\vec{c})} \bullet \vec{h}^{(t,\vec{c})}$$

with both $\vec{h}^{(s,\vec{c})}$ and $\vec{h}^{(t,\vec{c})}$ partitions of unity on ρ''. For $i < \nu_{s,0}$ and $j < \nu_{t,0}$, we say x *is* (i,j)*-good* iff

$$h_i^{s,\vec{c}}(x[\vec{a}^{(s,\vec{c})}]) = h_j^{t,\vec{c}}(x[\vec{a}^{(t,\vec{c})}]) = 1.$$

For (i,j)-good $x \in \rho''$ we have that,

$$|s(x,\vec{c})| = \chi_i^{s,\vec{c}} \text{ and } |t(x,\vec{c})| = \chi_j^{t,\vec{c}}.$$

Set

$$\text{okay}(x,i,j) = [\![\ \chi_i^{s,\vec{c}} \geq \chi_j^{t,\vec{c}} \text{ and } x \text{ is } (i,j)\text{-good}]\!]$$

for $i < \nu_{s,0}$, $j < \nu_{t,0}$, and $x \in \rho''$. Certainly then,

$$\begin{aligned}
\text{flip}(s(x,\vec{c}), t(x,\vec{c})) &= \sum_{\substack{i<\nu_{s,0}\\ j<\nu_{t,0}}} \text{okay}(x,i,j) \cdot \left[2^{\chi_i^{s,\vec{c}}} - 1 - t(x,\vec{c})\right] \\
&= \sum_{\substack{i<\nu_{s,0}\\ j<\nu_{t,0}}} \text{okay}(x,i,j) \cdot 2^{\chi_i^{s,\vec{c}}} - \sum_{\substack{i<\nu_{s,0}\\ j<\nu_{t,0}}} \text{okay}(x,i,j) \cdot (t(x,\vec{c}) + 1).
\end{aligned}$$

Substituting $\Gamma_+^t - \Gamma_-^t$ in the last of the above equalities and setting,

$$\Gamma'_+ = \sum_{\substack{i<\nu_{s,0}\\ j<\nu_{t,0}}} \text{okay}(x,i,j) \cdot (\Gamma_-^t + 2^{\chi_i^{s,\vec{c}}})$$

and

$$\Gamma'_- = \sum_{\substack{i<\nu_{s,0}\\ j<\nu_{t,0}}} \text{okay}(x,i,j) \cdot (\Gamma_+^t + 1)$$

we conclude that,

$$\text{flip}(s(x,\vec{c}), t(x,\vec{c})) = \Gamma'_+ - \Gamma'_-.$$

The definition of the boolean function $\text{okay}(x,i,j)$ and the fact that (i,j)-goodness is a partition of unity forces that,

$$\Gamma'_+ - \Gamma'_- \geq 0.$$

Case: $\text{trunc}(t(x,\vec{z}) : l(x,\vec{z}), r(x,\vec{z}))$.

We are going to split the proof into two cases: trunc(t: $s, 0$) and trunc(t: ∞, s). But first we need some preparation. Following the blueprint for handling two terms simultaneously (cf. cluster + sublemma) and using the conclusion of the cluster length sublemma we all the appropriate parameters and a cylinder ρ'' such that for $\vec{c} < \ell^m$ and $x \in \rho''$ we have that:

1. $s(x, \vec{c}) = \Gamma^s_+ - \Gamma^s_-$
2. $|s(x, \vec{c})| = \vec{\chi}^{(\vec{c})} \bullet \vec{h}^{(\vec{c})}(x[\vec{a}^{(\vec{c})}])$, and
3. $t(x, \vec{c}) = \Gamma^t_+ - \Gamma^t_-$.

Recall that there is m_t such that $\|t\|_{\ell,m} \leq \ell^{m_t}$ for all large ℓ. Since clusters are involved, recall all local definitions in the proof of the cluster-to-finite lemma regarding clusters, preclusters, and $\prec$. Recall the definition of small(x) and large(x) from the cluster truncate sublemma. Since the conclusions of the cluster length sublemma also hold for terms in ring cluster form at m, we freely make use of the claim proved in the cluster truncate sublemma, i.e.

Claim : *Let $\vec{g}_s$ be the first precluster in $s(x, \vec{c})$. Then,*

$$s(x, \vec{c}) = s(x, \vec{c}) \cdot \mathrm{large}(x) + \mathrm{small}(x) \cdot \cdot \sum_{(a+\alpha_{t,j}) \in \vec{g}_s} 2^{a+\alpha_{s,\vec{c},j}} \cdot \psi_{s,\vec{c},j} \cdot f_j^{s,\vec{c}}(x[\vec{a}^{(s,\vec{c})}]) \cdot \varphi_{s,\vec{c},j,a} \cdot \wedge x[\mathrm{supp}(a)].$$

The proof is the same as that in the cluster truncate sublemma. Let

$$\mathrm{v} \doteq \langle \tau_s, \xi_s, \tau_t, \xi_t \rangle ,$$

where $\tau_s, \xi_s \in \{0,1\}^{2 \cdot \nu_{s,0}}$ and $\tau_t, \xi_t \in \{0,1\}^{2 \cdot \nu_{t,0}}$. These quadruples will be called *relevant.* For relevant $\mathrm{v} = \langle \tau_s, \xi_s, \tau_t, \xi_t \rangle$, we set:

1. $s(\mathrm{v}) = \sum_{(a+\alpha_{s,\vec{c},j},j) \in \vec{g}_s} 2^{a+\alpha_{s,\vec{c},j}} \cdot \psi_{s,\vec{c},j} \cdot \xi_s(j) \cdot \varphi_{s,\vec{c},j,a} \cdot \tau_s(j)$;
2. $\vec{g}(\mathrm{v}) =$ the (unique) precluster of t the range of which contains $s(\mathrm{v})$;
3. $t_{\vec{g}(\mathrm{v})}(\mathrm{v}) = \sum_{(a+\alpha_{t,\vec{c},j},j) \in \vec{g}(\mathrm{v})} 2^{a+\alpha_{t,\vec{c},j}} \cdot \psi_{t,\vec{c},j} \cdot \xi_t(j) \cdot \varphi_{t,\vec{c},j,a} \cdot \tau_t(j)$, the value of the *critical cluster* of t relative to v;

4. $\vec{g}^{+}(\mathrm{v})$ be the $\prec$-successor precluster of $\vec{g}(\mathrm{v})$;

5. $I(\mathrm{v}) = \{a + \alpha^{+}_{t,\vec{c},j}, a + \alpha^{-}_{t,\vec{c},j} \mid (a + \alpha^{+}_{t,\vec{c},j}, j), (a + \alpha^{-}_{t,\vec{c},j}, j) \in \vec{g}(\mathrm{v})\}$;

6. $I^{+}(\mathrm{v}) = \{a + \alpha^{+}_{t,\vec{c},j}, a + \alpha^{-}_{t,\vec{c},j} \mid (a + \alpha^{+}_{t,\vec{c},j}, j), (a + \alpha^{-}_{t,\vec{c},j}, j) \in \vec{g}^{+}(\mathrm{v})\}$;

7. $\beta(\mathrm{v}) = \min I(\mathrm{v})$, $\gamma(\mathrm{v}) = \lceil \log_2 2 \cdot \nu_{t,0} \rceil + \lceil \log_2 \ell \rceil^{e(t)} + \max I(\mathrm{v})$,
$\alpha(\mathrm{v}) = \begin{cases} \min I^{+}(\mathrm{v}), & \text{if } \vec{g}(\mathrm{v}) \text{ is not } \prec\text{-maximum;} \\ \max\{s(\mathrm{v}), \gamma(\mathrm{v})\}, & \text{otherwise.} \end{cases}$

8. Not-last(v) and Not-first(v) are as in the cluster truncate sublemma;

9. Not-first$(\mathrm{v}, j, +) = [\![$ there is $a \in A^{+}_{t,\vec{c},j}$ such that $(a + \alpha^{+}_{t,\vec{c},j}, j) \prec \vec{g}(\mathrm{v})]\!]$, for $j < \nu_{t,0}$;

10. Not-first$(\mathrm{v}, j, -) = [\![$ there is $a \in A^{-}_{t,\vec{c},j}$ such that $(a + \alpha^{-}_{t,\vec{c},j}, j) \prec \vec{g}(\mathrm{v})]\!]$, for $j < \nu_{t,0}$;

11. Not-last$(\mathrm{v}, j, +) = [\![$ there is $a \in A^{+}_{t,\vec{c},j}$ such that $(a + \alpha^{+}_{t,\vec{c},j}, j) \succ \vec{g}(\mathrm{v})]\!]$, for $j < \nu_{t,0}$.

12. Not-last$(\mathrm{v}, j, -) = [\![$ there is $a \in A^{-}_{t,\vec{c},j}$ such that $(a + \alpha^{-}_{t,\vec{c},j}, j) \succ \vec{g}(\mathrm{v})]\!]$, for $j < \nu_{t,0}$;

$\gamma(\mathrm{v}) - 1$ is the last bit we record in the binary expansion of t within the range of $\vec{g}(\mathrm{v})$. With $x \in \rho''$, say that x *is* v*-good* if the following hold:

1. x is both $(t{:}\,\xi_t)$-good and $(s{:}\,\xi_s)$-good (cf. cluster-to-finite lemma);

2. $1 = [\![\bigwedge_{(a+\alpha_{s,j},j)\in\vec{g}_s} (\tau_s(j) = \wedge x[\mathrm{supp}(a)])]\!]$, and

3. $1 = [\![\bigwedge_{(a+\alpha_{t,j},j)\in\vec{g}(\mathrm{v})} (\tau_t(j) = \wedge x[\mathrm{supp}(a)])]\!]$.

We have to take into account the carry from possible previous subtractions. Recall the definition of $Q_{-1}(x, y)$ in the ring cluster-to-finite lemma. Let $C_{-1}(x, s(\mathrm{v}))$ be the circuit that calculates the truth value for $Q_{-1}(x, s(\mathrm{v}))$. Note that $\vec{c}$ and the cylinder ρ'' are fixed. The value of $C_{-1}(x, s(\mathrm{v}))$ depends on the fixed parameters and x. In particular this value depends on $\vec{c}$ and v. Form the multicircuit

$$\mathbf{C} = \{C_{-1}(x, s(\mathrm{v})) \mid \vec{c} < \ell^m, \mathrm{v} \text{ relevant}\}$$

and apply the multicircuit lemma to obtain a finer cylinder $\rho''' \subset \rho''$, $\mu_{\mathbf{C}}$, $\vec{a}^{(t,\vec{c},\mathrm{v})}$ for $\vec{c} < \ell^m$ and relevant v, and $\mathbf{c}_{t,\vec{c},\mathrm{v}} \in \mathrm{BOOL}_{\mu_{\mathbf{C}}}$ such that:

1. $\vec{a}^{(t,\vec{c},\mathrm{v})}$ has no more than $\mu_{\mathbf{c}}$ bits, and

2. $\forall\, \vec{c} < \ell^m\, \forall\, \mathrm{v}$, relevant $\forall\, x \in \rho'''$, v-good
$$C_{-1}(x, s(\mathrm{v})) = \mathbf{c}_{t,\vec{c},\mathrm{v}}(x[\vec{a}^{(t,\vec{c},\mathrm{v})}]).$$

Now, we can proceed. Fix $\vec{c} < \ell^m$ for the rest of the considerations. Let v be relevant. Set:

$$\mathrm{left}(x, \mathrm{v}) = \sum_{\substack{\vec{g} \text{ a precluster of } t \\ \vec{g} \succ \vec{g}(\mathrm{v})}} t_{\vec{g}}(x) \cdot 2^{-s(\mathrm{v})}$$

$$\mathrm{right}(x, \mathrm{v}) = \sum_{\substack{\vec{g} \text{ a precluster of } t \\ \vec{g} \prec \vec{g}(\mathrm{v})}} t_{\vec{g}}(x)$$

$$r_-(x, \mathrm{v}) = \mathrm{trunc}(2^{\alpha(\mathrm{v})} + t_{\vec{g}(\mathrm{v})}(\mathrm{v}) - 2^{\beta(\mathrm{v})}\colon \alpha(\mathrm{v}), s(\mathrm{v}))$$

$$r_+(x, \mathrm{v}) = \mathrm{trunc}(2^{\alpha(\mathrm{v})} + t_{\vec{g}(\mathrm{v})}(\mathrm{v})\colon \alpha(\mathrm{v}), s(\mathrm{v}))$$

$$l_-(x, \mathrm{v}) = \mathrm{trunc}(2^{\alpha(\mathrm{v})} + t_{\vec{g}(\mathrm{v})}(\mathrm{v}) - 2^{\beta(\mathrm{v})}\colon s(\mathrm{v}), \beta(\mathrm{v}))$$

$$l_+(x, \mathrm{v}) = \mathrm{trunc}(2^{\alpha(\mathrm{v})} + t_{\vec{g}(\mathrm{v})}(\mathrm{v})\colon s(\mathrm{v}), \beta(\mathrm{v})).$$

Given then a relevant v and v-good $x \in \rho'''$ we have that

$$\begin{aligned}\mathrm{trunc}(t(x,\vec{c})\colon \infty, s(x,\vec{c})) = \mathrm{small}(x) \cdot [\![x \text{ is v-good}]\!] \cdot \Big[&[\text{Not-last}(\mathrm{v}) \cdot \mathrm{left}(x, \mathrm{v})] + \\ &+ \mathbf{c}_{t,\vec{c},\mathrm{v}}(x[\vec{a}^{(t,\vec{c},\mathrm{v})}]) \cdot r_-(x, \mathrm{v}) + (1 - \mathbf{c}_{t,\vec{c},\mathrm{v}}(x[\vec{a}^{(t,\vec{c},\mathrm{v})}])) \cdot r_+(x, \mathrm{v})\Big] \qquad \textbf{(R)}\end{aligned}$$

and

$$\begin{aligned}\mathrm{trunc}(t(x,\vec{c})\colon s(x,\vec{c}), 0) = \mathrm{large}(x) \cdot t(x,\vec{c}) + \\ + [\![x \text{ is v-good}]\!] \cdot \Big[&\mathbf{c}_{t,\vec{c},\mathrm{v}}(x[\vec{a}^{(t,\vec{c},\mathrm{v})}]) \cdot l_-(x, \mathrm{v}) + \\ &+ (1 - \mathbf{c}_{t,\vec{c},\mathrm{v}}(x[\vec{a}^{(t,\vec{c},\mathrm{v})}])) \cdot l_+(x, \mathrm{v}) + [\text{Not-first}(\mathrm{v}) \cdot \mathrm{right}(x, \mathrm{v})]\Big]. \qquad \textbf{(L)}\end{aligned}$$

For fixed relevant v we shall need to partition the critical cluster of t into $2 \cdot \nu_{t,0}$ pieces. To this end we set for $j < \nu_{t,0}$,

$$\eta(j) = \begin{cases} a + \alpha^+_{t,\vec{c},j}, & \text{if } a \in A^+_{t,\vec{c},j} \text{ and } (a + \alpha^+_{t,\vec{c},j}, j) \in \vec{g}(\mathrm{v}); \\ \text{undefined}, & \text{if } a \in A^+_{t,\vec{c},j} \Longrightarrow (a + \alpha^+_{t,\vec{c},j}, j) \notin \vec{g}(\mathrm{v}). \end{cases}$$

for $\nu_{t,0} \leq j < 2 \cdot \nu_{t,0}$,

$$\eta(j) = \begin{cases} a + \alpha^-_{t,\vec{c},(j \bmod \nu_{t,0})}, & \text{if } a \in A^-_{t,\vec{c},(j \bmod \nu_{t,0})} \text{ and} \\ & (a + \alpha^-_{t,\vec{c},(j \bmod \nu_{t,0})}, (j \bmod \nu_{t,0})) \in \vec{g}(\mathrm{v}); \\ \text{undefined}, & \text{if for } a \in A^-_{t,\vec{c},(j \bmod \nu_{t,0})} \Longrightarrow \\ & \Longrightarrow (a + \alpha^-_{t,\vec{c},(j \bmod \nu_{t,0})}, (j \bmod \nu_{t,0})) \notin \vec{g}(\mathrm{v}). \end{cases}$$

and call (j) η-*defined* if $\eta(j)$ is defined. Next, we order the $(j), j < 2 \cdot \nu_{t,0}$ as follows:

$$(j) \prec (j') \text{ if both } (j), (j') \text{ are } \eta\text{-defined and } \eta(j) < \eta(j') \text{ or else if } j < j'.$$

$\eta(j)$ and $\prec$ are well defined because there is at most one $a \in A^+_{t,\vec{c},j}, A^-_{t,\vec{c},(j \bmod \nu_{t,0})}$ with $(a+\alpha^+_{t,\vec{c},j}, j) \in \vec{g}(\mathrm{v})$ or $(a+\alpha^-_{t,\vec{c},(j \bmod \nu_{t,0})}, (j \bmod \nu_{t,0})) \in \vec{g}(\mathrm{v})$ since $\mathrm{spread}(\mathcal{H}(d\colon A_{d,e,\nu,\ell})) > 4 \cdot \nu_{t,0} \cdot \lceil \log_2 \ell \rceil^e$. For (j) η-defined we set $(j)^+$ the η-defined $\prec$-successor of (j), and if (j) is the η-defined $\prec$-maximum then $(j)^+$ is undefined but we let $\eta(j)^+ = \gamma(\mathrm{v})$. We enumerate the critical precluster $\vec{g}(\mathrm{v})$ of t as $(a + \alpha_{t,\vec{c},j}, j)$ for $j < 2 \cdot \nu_{t,0}$, i.e. we drop the distinction between '+' and '−'. We now distinguish the cases as promised for a fixed relevant v:

Subcase: R. For $j < 2 \cdot \nu_{t,0}$ we set

$$\tilde{\eta}(j) = \max\{\eta(j), s(\mathrm{v})\} \text{ for } (j) \ \eta\text{-defined},$$

$$t_{\mathrm{v},j} = \begin{cases} \mathrm{trunc}(r_+(x, \mathrm{v})\colon \eta(j)^+, \tilde{\eta}(j)), & \text{if } (j) \text{ is } \eta\text{-defined}; \\ 0, & \text{if } (j) \text{ in not } \eta\text{-defined}. \end{cases}$$

$$t'_{\mathrm{v},j} = \begin{cases} \mathrm{trunc}(r_-(x, \mathrm{v})\colon \eta(j)^+, \tilde{\eta}(j)), & \text{if } (j) \text{ is } \eta\text{-defined}; \\ 0, & \text{if } (j) \text{ in not } \eta\text{-defined}. \end{cases}$$

We now give the parameters for Γ_+ and Γ_- in the ring cluster form at m. For $j < \nu_{t,0}$, the part responsible for the correct expansion of $\mathrm{left}(x, \mathrm{v})$ in **R**:

1. $A^+_{\vec{c},\mathrm{v},j} = A^+_{t,\vec{c},j} \backslash \{a \mid (a+\alpha^+_{t,\vec{c},j}, j) \in \vec{g} \text{ with } \vec{g} \preceq \vec{g}(\mathrm{v})\}$, $A^-_{\vec{c},\mathrm{v},j} = A^-_{t,\vec{c},j} \backslash \{a \mid (a+\alpha^-_{t,\vec{c},j}, j) \in \vec{g} \text{ with } \vec{g} \preceq \vec{g}(\mathrm{v})\}$;

2. $\alpha^+_{\vec{c},\mathrm{v},j} = \alpha^+_{t,\vec{c},j} - s(\mathrm{v})$, $\alpha^-_{\vec{c},\mathrm{v},j} = \alpha^-_{t,\vec{c},j} - s(\mathrm{v})$;

3. $\varphi^+_{\vec{c},\mathrm{v},j,a} = \varphi^+_{t,\vec{c},j,a}$, $\varphi^-_{\vec{c},\mathrm{v},j,a} = \varphi^-_{t,\vec{c},j,a}$ for $a \in A_{\vec{c},\mathrm{v},j}$;

4. $\psi^+_{\vec{c},\mathrm{v},j} = \psi^+_{t,\vec{c},j}$, $\psi^-_{\vec{c},\mathrm{v},j} = \psi^-_{t,\vec{c},j}$,

5. $f^+_{\vec{c},\mathrm{v},j} = \mathrm{Not\text{-}last}(\mathrm{v}, j, +) \cdot \mathrm{small}(x) \cdot [\![x \text{ is v-good}]\!] \cdot f^+_{t,\vec{c},j}$,

6. $f^-_{\vec{c},\mathrm{v},j} = \text{Not-last}(\mathrm{v}, j, -) \cdot \text{small}(x) \cdot [\![x \text{ is v-good}]\!] \cdot f^-_{t,\vec{c},j}$.

For $\nu_{t,0} \leq j < 5 \cdot \nu_{t,0}$:

7. $A^+_{\vec{c},\mathrm{v},j} = \emptyset$, $\alpha^+_{\vec{c},\mathrm{v},j} = 0$, and $\varphi^+_{\vec{c},\mathrm{v},j,a} = 1$.

For $\nu_{t,0} \leq j < 3 \cdot \nu_{t,0}$, the part responsible for getting the correct expansion for $r_+(x, \mathrm{v})$, i.e. when there is *no* carry from previous subtractions:

8. $\psi^+_{\vec{c},\mathrm{v},j} = t_{\mathrm{v},(j \bmod 2\cdot\nu_{t,0})}$, and

9. $f^+_{\vec{c},\mathrm{v},j} = \text{small}(x) \cdot [\![x \text{ is v-good}]\!] \cdot (1 - \mathbf{c}_{t,\vec{c},\mathrm{v}})$.

For $3 \cdot \nu_{t,0} \leq j < 5 \cdot \nu_{t,0}$, the part responsible for getting the correct expansion for $r_-(x, \mathrm{v})$, i.e. when there *is* carry from previous subtractions:

10. $\psi^+_{\vec{c},\mathrm{v},j} = t'_{\mathrm{v},(j \bmod 2\cdot\nu_{t,0})}$, and

11. $f^+_{\vec{c},\mathrm{v},j} = \text{small}(x) \cdot [\![x \text{ is v-good}]\!] \cdot \mathbf{c}_{t,\vec{c},\mathrm{v}}$.

Now, we concern ourselves with the case when $t_{\vec{g}(\mathrm{v})}(\mathrm{v}) < 0$, or when $t_{\vec{g}(\mathrm{v})}(\mathrm{v}) = 0$ and there is a carry from previous subtractions, or when $t_{\vec{g}(\mathrm{v})}(\mathrm{v}) < 0$ and $s(\mathrm{v})$ goes past the end of the cluster but lies before $\alpha(\mathrm{v})$. This only adds one more summand in Γ_-, namely at $j = \nu_{t,0}$. So for $j = \nu_{t,0}$ we set:

12. $A^-_{\vec{c},\mathrm{v},j} = \emptyset$, $\varphi^-_{\vec{c},\mathrm{v},j,a} = 1$, $\psi^-_{\vec{c},\mathrm{v},j} = 1$;

13. $f^-_{\vec{c},\mathrm{v},j} = \text{small}(x) \cdot [\![x \text{ is v-good}]\!] \cdot ([\![t_{\vec{g}(\mathrm{v})}(\mathrm{v}) < 0]\!] + [\![t_{\vec{g}(\mathrm{v})}(\mathrm{v}) = 0]\!] \cdot \mathbf{c}_{t,\vec{c},\mathrm{v}})$, and

14. $\alpha^-_{\vec{c},\mathrm{v},j} = \max\{s(\mathrm{v}), \gamma(\mathrm{v})\} - s(\mathrm{v})$,

while for $\nu_{t,0} < j < 5 \cdot \nu_{t,0}$ we set $\psi^-_{\vec{c},\mathrm{v},j} = 0$. For $5 \cdot \nu_{t,0} \leq j < 6 \cdot \nu_{t,0}$ (which are intended for the next subcase) we only set $f^+_{\vec{c},\mathrm{v},j} = f^-_{\vec{c},\mathrm{v},j} = 0$.

The last parameter is the sequence of bits which determine all of the above properties for $x \in \rho'''$. For the v at hand we set:

$$\vec{a}^{\,(\vec{c},\mathrm{v})} = \vec{a}^{\,(s,\vec{c})} \cup \vec{a}^{\,(t,\vec{c})} \cup \vec{a}^{\,(t,\vec{c},\mathrm{v})} \cup \bigcup_{(a+\alpha_{t,j})\in\vec{g}_s} \text{supp}(a) \cup \bigcup_{(a+\alpha_{t,j})\in\vec{g}(\mathrm{v})} \text{supp}(a).$$

$\vec{a}^{(t,\vec{c},\mathbf{v})}$ determines the carry from previous subtractions, and the last two unions determine v-goodness, provided that $x \in \rho'''$. So, the final sequence is

$$\vec{a}^{(\vec{c})} = \bigcup_{\mathbf{v} \text{ relevant}} \vec{a}^{(\vec{c},\mathbf{v})},$$

The bound on its size clearly depends only on $d_s, d_t, \nu_{t,0}, \nu_{s,0}, \mu_{\mathbf{c}}, \mu_s$, and μ_t. We set,

$$\Gamma_+^{\mathbf{v}} = \sum_{j<6\cdot\nu_{t,0}} 2^{\alpha^+_{\vec{c},\mathbf{v},j}} \cdot \psi^+_{\vec{c},\mathbf{v},j} \cdot f^+_{\vec{c},\mathbf{v},j}(x[\vec{a}^{(\vec{c})}]) \cdot \sum_{a\in A^+_{\vec{c},\mathbf{v},j}} \varphi^+_{\vec{c},\mathbf{v},j,a} \cdot \wedge x[\mathrm{supp}(a)] \cdot 2^a$$

and

$$\Gamma_-^{\mathbf{v}} = \sum_{j<6\cdot\nu_{t,0}} 2^{\alpha^-_{\vec{c},\mathbf{v},j}} \cdot \psi^-_{\vec{c},\mathbf{v},j} \cdot f^-_{\vec{c},\mathbf{v},j}(x[\vec{a}^{(\vec{c})}]) \cdot \sum_{a\in A^-_{\vec{c},\mathbf{v},j}} \varphi^-_{\vec{c},\mathbf{v},j,a} \cdot \wedge x[\mathrm{supp}(a)] \cdot 2^a.$$

It is matter of unraveling the settings to see that indeed for $x \in \rho''$

$$\mathrm{trunc}(t(x,\vec{c})\colon\infty, s(x,\vec{c})) = \sum_{\mathbf{v} \text{ relevant}} \Gamma_+^{\mathbf{v}} - \Gamma_-^{\mathbf{v}}.$$

This finishes "right" subcase.

Subcase: L. Here, we set for $j < 2\cdot\nu_{t,0}$:

$$\tilde{\eta}(j) = \min\{\eta(j)^+, s(\mathbf{v})\},$$

$$t_{\mathbf{v},j} = \begin{cases} \mathrm{trunc}(l_+(x,\mathbf{v})\colon\tilde{\eta}(j),\eta(j)), & \text{if } (j) \text{ is } \eta\text{-defined;} \\ 0, & \text{if } (j) \text{ is not } \eta\text{-defined.} \end{cases}$$

$$t'_{\mathbf{v},j} = \begin{cases} \mathrm{trunc}(l_-(x,\mathbf{v})\colon\tilde{\eta}(j),\eta(j)), & \text{if } (j) \text{ is } \eta\text{-defined;} \\ 0, & \text{if } (j) \text{ is not } \eta\text{-defined.} \end{cases}$$

Next, the parameters for the ring cluster form at m. For $j < \nu_{t,0}$, the part responsible for the correct expansion of $\mathrm{right}(x,\mathbf{v})$ in **L**:

1. $A^+_{\vec{c},\mathbf{v},j} = A^+_{t,\vec{c},j} \setminus \{a \mid (a+\alpha^+_{t,\vec{c},j}, j) \in \vec{g} \text{ with } \vec{g} \succeq \vec{g}(\mathbf{v})\}$, $A^-_{\vec{c},\mathbf{v},j} = A^-_{t,\vec{c},j} \setminus \{a \mid (a+\alpha^-_{t,\vec{c},j}, j) \in \vec{g} \text{ with } \vec{g} \succeq \vec{g}(\mathbf{v})\}$;
2. $\alpha^+_{\vec{c},\mathbf{v},j} = \alpha^+_{t,\vec{c},j}$, $\alpha^-_{\vec{c},\mathbf{v},j} = \alpha^-_{t,\vec{c},j}$;
3. $\varphi^+_{\vec{c},\mathbf{v},j,a} = \varphi^+_{t,\vec{c},j,a}$, $\varphi^-_{\vec{c},\mathbf{v},j,a} = \varphi^-_{t,\vec{c},j,a}$ for $a \in A_{\vec{c},\mathbf{v},j}$;
4. $\psi^+_{\vec{c},\mathbf{v},j} = \psi^+_{t,\vec{c},j}$, $\psi^-_{\vec{c},\mathbf{v},j} = \psi^-_{t,\vec{c},j}$,

5. $f^+_{\vec{c},\mathrm{v},j} = \text{Not-first}(\mathrm{v}, j, +) \cdot \text{small}(x) \cdot [\![x \text{ is v-good}]\!] \cdot f^+_{t,\vec{c},j}$,

6. $f^-_{\vec{c},\mathrm{v},j} = \text{Not-first}(\mathrm{v}, j, -) \cdot \text{small}(x) \cdot [\![x \text{ is v-good}]\!] \cdot f^-_{t,\vec{c},j}$.

For $\nu_{t,0} \leq j < 5 \cdot \nu_{t,0}$

7. $A^+_{\vec{c},\mathrm{v},j} = \{a\}$, $\alpha^+_{\vec{c},\mathrm{v},j} = \alpha_{t,j \bmod 2\cdot\nu_{t,0}}$, $\varphi^+_{\vec{c},\mathrm{v},j,a} = 1$, for η-defined $(j \bmod 2 \cdot \nu_{t,0})$ with $(a + \alpha_{t,j \bmod 2\cdot\nu_{t,0}}) \in \vec{g}(\mathrm{v})$;

8. $A^+_{\vec{c},\mathrm{v},j} = \emptyset$, $\alpha^+_{\vec{c},\mathrm{v},j} = 0$, for $(j \bmod 2 \cdot \nu_{t,0})$ not η-defined.

For $\nu_{t,0} \leq j < 3 \cdot \nu_{t,0}$, the part responsible for getting the correct expansion of $l_+(x, \mathrm{v})$, i.e. when there is *no* carry from previous subtractions:

9. $\psi^+_{\vec{c},\mathrm{v},j} = t_{\mathrm{v},j \bmod 2\cdot\nu_{t,0}}$, and

10. $f^+_{\vec{c},\mathrm{v},j} = (1 - \mathbf{c}_{t,\vec{c},\mathrm{v}}) \cdot \text{small}(x) \cdot [\![x \text{ is v-good}]\!]$.

For $3 \cdot \nu_{t,0} \leq j < 5 \cdot \nu_{t,0}$, the part responsible for getting the correct expansion for $l_-(x, \mathrm{v})$, i.e. when there *is* carry from previous subtractions:

11. $\psi^+_{\vec{c},\mathrm{v},j} = t'_{\mathrm{v},j \bmod 2\cdot\nu_{t,0}}$, and

12. $f^+_{\vec{c},\mathrm{v},j} = \mathbf{c}_{t,\vec{c},\mathrm{v}} \cdot \text{small}(x) \cdot [\![x \text{ is v-good}]\!]$.

For $j = 5 \cdot \nu_{t,0}$ we eliminate the extra carry from the expansion of $\text{right}(x, \mathrm{v})$ because the expansion of $l_-(x, \mathrm{v})$ has already been formed with a carry:

13. $A^-_{\vec{c},\mathrm{v},j} = \emptyset$, $\psi^-_{\vec{c},\mathrm{v},j} = 1$, $\alpha^-_{\vec{c},\mathrm{v},j} = \beta(\mathrm{v})$, $f^-_{\vec{c},\mathrm{v},j} = \mathbf{c}_{t,\vec{c},\mathrm{v}}$.

Now, we concern ourselves with the case when $t_{\vec{g}(\mathrm{v})}(\mathrm{v}) < 0$, or when $t_{\vec{g}(\mathrm{v})}(\mathrm{v}) = 0$ and there is a carry from previous subtractions, or when $t_{\vec{g}(\mathrm{v})}(\mathrm{v}) < 0$ and $s(\mathrm{v})$ goes past the end of the cluster but lies before $\alpha(\mathrm{v})$. However, as before, when $s(\mathrm{v}) > \gamma(\mathrm{v})$ we have no control over $s(\mathrm{v}) - \gamma(\mathrm{v})$. Thus, it would be a mistake to simply add $2^{s(\mathrm{v})} - 2^{\gamma(\mathrm{v})}$. We create the latter by putting appropriate summands to both Γ_+ and Γ_- at $j = 5 \cdot \nu_{t,0} + 1$. So for $j = 5 \cdot \nu_{t,0} + 1$ we set:

14. $A^+_{\vec{c},\mathrm{v},j} = A^-_{\vec{c},\mathrm{v},j} = \emptyset$, $\varphi^+_{\vec{c},\mathrm{v},j,a} = \psi^+_{\vec{c},\mathrm{v},j} = 1$, $\varphi^-_{\vec{c},\mathrm{v},j,a} = \psi^-_{\vec{c},\mathrm{v},j} = 1$;

15. $\alpha^+_{\vec{c},\mathrm{v},j} = s(\mathrm{v})$, $\alpha^-_{\vec{c},\mathrm{v},j} = \gamma(\mathrm{v})$, and

16. $f^+_{\vec{c},\mathbf{v},j} = f^-_{\vec{c},\mathbf{v},j} = ([\![t_{\vec{g}(\mathbf{v})}(\mathbf{v}) < 0]\!] + [\![t_{\vec{g}(\mathbf{v})}(\mathbf{v}) = 0]\!] \cdot \mathbf{c}_{t,\vec{c},\mathbf{v}}) \cdot [\![s(\mathbf{v}) > \gamma(\mathbf{v})]\!] \cdot \mathrm{small}(x) \cdot [\![x \text{ is v-good}]\!]$,

while for $\nu_{t,0} \leq j < 5 \cdot \nu_{t,0}$ and $j = 5 \cdot \nu_{t,0} + 1$ we set $\psi^-_{\vec{c},\mathbf{v},j} = 0$. We now give the parameters when $x \in \rho''$ is "large". So for $5 \cdot \nu_{t,0} + 1 < j \leq 6 \cdot \nu_{t,0} + 1$ we put:

17. $A^+_{\vec{c},\mathbf{v},j} = A^+_{t,\vec{c},(j \bmod \nu_{t,0})}$, $A^-_{\vec{c},\mathbf{v},j} = A^-_{t,\vec{c},(j \bmod \nu_{t,0})}$,

18. $\alpha^+_{\vec{c},\mathbf{v},j} = \alpha^+_{t,\vec{c},(j \bmod \nu_{t,0})}$, $\alpha^-_{\vec{c},\mathbf{v},j} = \alpha^-_{t,\vec{c},(j \bmod \nu_{t,0})}$,

19. $\varphi^+_{\vec{c},\mathbf{v},j,a} = \varphi^+_{t,\vec{c},(j \bmod \nu_{t,0}),a}$, $\varphi^-_{\vec{c},\mathbf{v},j,a} = \varphi^-_{t,\vec{c},(j \bmod \nu_{t,0}),a}$ for $a \in A_{\vec{c},\mathbf{v},j}$,

20. $\psi^+_{\vec{c},\mathbf{v},j} = \psi^+_{t,\vec{c},(j \bmod \nu_{t,0})}$, $\psi^-_{\vec{c},\mathbf{v},j} = \psi^-_{t,\vec{c},(j \bmod \nu_{t,0})}$,

21. $f^+_{\vec{c},\mathbf{v},j} = \mathrm{large}(x) \cdot [\![x \text{ is v-good}]\!] \cdot f^+_{t,\vec{c},(j \bmod \nu_{t,0})}$,

22. $f^-_{\vec{c},\mathbf{v},j} = \mathrm{large}(x) \cdot [\![x \text{ is v-good}]\!] \cdot f^-_{t,\vec{c},(j \bmod \nu_{t,0})}$.

The bits $\vec{a}^{(\vec{c})}$ remain as in subcase **R**. We set,

$$\Gamma^{\mathbf{v}}_+ = \sum_{j \leq 6 \cdot \nu_{t,0}} 2^{\alpha^+_{\vec{c},\mathbf{v},j}} \cdot \psi^+_{\vec{c},\mathbf{v},j} \cdot f^+_{\vec{c},\mathbf{v},j}(x[\vec{a}^{(\vec{c})}]) \cdot \sum_{a \in A^+_{\vec{c},\mathbf{v},j}} \varphi^+_{\vec{c},\mathbf{v},j,a} \cdot \wedge x[\mathrm{supp}(a)] \cdot 2^a$$

and

$$\Gamma^{\mathbf{v}}_- = \sum_{j \leq 6 \cdot \nu_{t,0}} 2^{\alpha^-_{\vec{c},\mathbf{v},j}} \cdot \psi^-_{\vec{c},\mathbf{v},j} \cdot f^-_{\vec{c},\mathbf{v},j}(x[\vec{a}^{(\vec{c})}]) \cdot \sum_{a \in A^-_{\vec{c},\mathbf{v},j}} \varphi^-_{\vec{c},\mathbf{v},j,a} \cdot \wedge x[\mathrm{supp}(a)] \cdot 2^a.$$

It is matter of unraveling the settings to see that indeed,

$$\mathrm{trunc}(t(x,\vec{c})\colon s(x,\vec{c}),0) = \sum_{\mathbf{v}\ \mathbf{relevant}} \Gamma^{\mathbf{v}}_+ - \Gamma^{\mathbf{v}}_-.$$

This finishes the "left" subcase. In both subcases it will suffice to set

$$\nu_{r,0} = \max\{(6 \cdot \nu_{t,0} + 1) \cdot 2^\kappa, \nu_{s,0}\}$$

, where $\kappa = 8 \cdot (\nu_{t,0} + \nu_{s,0})$. An upper bound on the number of bits in $\vec{a}^{(\vec{c})}$ is

$$\mu_{\mathbf{c}} + \mu_s + \mu_t + d \cdot (2 \cdot \nu_{s,0} + 2 \cdot \nu_{t,0}).$$

□

We omit the boolean ring circuit lemma because its statement and proof are exactly the same as those of the boolean cut-and-paste circuit lemma of chapter 6. However, as a matter of reference we do state the local sharp version, which again is the same as the sharp version of chapter 6.

Sharp Ring Circuit Lemma 8.3. *Let* $\Theta(x)$ *be a ring sharply bounded formula. Then,*
$\exists\, N, m_0, \nu, d, \ell_0, \delta : 0 < \delta \leq 1;$
$\forall \ell \geq \ell_0;$
$\exists \rho \in \mathrm{Restr}_\delta(d, e, \nu, \ell);$
$\exists C^{\Theta}_{\ell,\rho}$ *a circuit of size* $\leq \ell^{m_0}$ *and depth* $\leq N + 6;$
$\forall x \in \rho,$

$$[\![\Theta(x)]\!] = C^{\Theta}_{\ell,\rho}(x).$$

Proof. □

Proof of the theorem

If **PARITY** were ring sharply bounded definable then by the sharp ring cluster lemma we would be able to calculate $[\![\mathbf{PARITY}(x)]\!]$ with a circuit of constant depth and size polynomial in the length of x, which is not possible. ■

Bibliography

[Ajt83] M. Ajtai. $\sum_1^1$-formulae on finite structures. *Annals of Pure and Applied Logic*, 2(24):1–48, 1983.

[Ajt88] M. Ajtai. The complexity of the pigeonhole principle. *Proceedings of the 29th Annual IEEE Symposium on Foundations of Computer Science*, pages 346–355, 1988.

[Ben62] J.H. Bennet. *On spectra.* PhD thesis, Princeton University, 1962.

[Bus86] Samuel R. Buss. *Bounded Arithmetic.* Bibliopolis, 1986. Revision of Ph.D. dissertation, Princeton University, 1985.

[CT86] P. Clote and G. Takeuti. Exponential time and bounded arithmetic. *Springer Lecture Notes in Computer Science 221*, pages 125–143, 1986.

[DG82] C. Dimitrakopoulos and H. Gaifman. Fragments of arithmetic and the mrdp theorem. *Logic and Algorithmic, Monographie No. 30 de L'Enseignment Matèmatique*, pages 187–206, 1982.

[DP82] C. Dimitrakopoulos and J.B. Paris. Truth definitions for Δ_0 formulae. *Logic and Algorithmic, Monographie No. 30 de L'Enseignment Matèmatique*, pages 317–329, 1982.

[Fer88] Fernando Jorge Innocéncio Ferreira. *Polynomial Time Computable Arithmetic and Conservative Extensions.* PhD thesis, The Pennsylvania State University, 1988.

[FSS84] M. Furst, J.B. Saxe, and M. Sipser. Parity, Circuits, and the Polynomial-Time Hierarchy. *Mathematical Systems Theory*, 17:13–27, 1984.

[Hås86] Johan Håstad. *Computational Limitations for Small Depth Circuits.* PhD thesis, MIT, 1986.

[Imm87] N. Immerman. Expressibility and parallel complexity. 1987. revised August 1988.

[KPT] J. Krajíček, P. Pudlák, and G. Takeuti. Bounded arithmetic and the polynomial hierarchy. *Annals of Pure and Applied Logic.* to appear.

[Par71] R. Parikh. Existence and feasibility in arithmetic. *J. Symb. Logic*, 36(3):494–508, 1971.

[PW87] J.B. Paris and A.J. Wilkie. On the scheme of induction for bounded arithmetic formulas. *Annals of Pure and Applied Logic*, 35(3):205–303, 1987.

[PWW88] J.B Paris, A.J. Wilkie, and A.R. Woods. A note on the provability of the Δ_0-php principle and the existence of infinitely many primes. *J. Symb. Logic*, 53(4):1235–1244, 1988.

[Raz87] A. Razborov. Lower bounds for the size of circuits of bounded depth with basis $\{\wedge, \oplus\}$. *Math. notes of the Academy , of Sciences of the USSR*, 41(4):333–338, 1987.

[Ruz81] W.L. Ruzzo. On uniform circuit complexity. *J. Comput. System Sci.*, 22:365–383, 1981.

[Smo87] R. Smolensky. Algebraic methods in the theory of lower bounds for Boolean circuit complexity. *Proc. 16th ACM Symp. on Theory of Computing*, 1987.

[Sto77] L. Stockmeyer. The polynomial time hierarchy. *Theoretical Computer Science*, 3, 1977.

[Tak90] G. Takeuti. Sharply bounded arithmetic and the function $a \dot{-} 1$. *Contemporary Mathematics*, 106, 1990.

[Weg87] I. Wegener. *The Complexity of Boolean Functions.* Wiley-Teubner series in Computer Science, Stuttgart, 1987.

[Woo81] A. Woods. *Some Problems in Logic and Number Theory and their Connections.* PhD thesis, Manchester University, 1981.

[Yao85] A.C. Yao. Separating the polynomial time hierarchy by oracles: Part I. *Proceedings of the 26th Annual IEEE Symposium on Foundation of Computer Science*, pages 1–10, 1985.