Complexity Theoretic Lower Bounds on Cryptographic Functions

Dissertation
zur Erlangung des Grades eines
Doktor der Mathematik
der Ruhr-Universität Bochum
am Fachbereich Mathematik

vorgelegt von
Eike Kiltz

unter der Betreuung von
Hans Ulrich Simon

Bochum
November 2003

Tag der Disputation:	4.2.2004
Dekan:	Prof. Dr. Lothar Gerritzen
Gutachter:	Prof. Dr. Hans Ulrich Simon
	Prof. Dr. Hans Dobbertin

Acknowledgments

First of all it is a pleasure to thank my supervisor Hans Ulrich Simon for the great support and the nice atmosphere in his research group. He has been a reliable source of encouragement and adv ice through all my years at the Ruhr-Universität Bochum. Thanks also to all the nice people I met at the “Lehrstuhl für Mathematik und Informatik” and the “Lehrstuhl für Informationssicherheit und Kryptologie” in Bochum, the group at “basic research in computer science BRICS” in Århus, and at various conferences and workshops. I greatly benefited from my discussions (personal or by email) with Andreas Bomke, Dietrich Braess, Hans Dobbertin, Jürgen Forster, Gregor Leander, Tanja Lange, Igor Shparlinski, and Arne Winterhof.

Thanks to the European Union who funded my three month stay at the university of Århus with a Marie-Curie fellowship. I'm very grateful to the BRICS group in Århus for their nice hospitality while my stays in spring and summer 2003.

Finally, thanks to my family and my friends for supporting me in everything I decided to do!

Contents

Chapter 1

Cryptography and Complexity
Theory:
A match made in heaven.

SHAFI GOLDWASSER

Introduction

One of the fundamental questions in theoretical computer science is to fully understand why some functions are easy to compute, while others require considerable large computing resources. Unfortunately, the present state of the art falls far away from giving an answer to this question – there are very few functions for which we know exact complexity measures (i.e., matching upper and lower bounds).

The study of complexity of algorithms began far before the invention of the computer. For many centuries the notion of an algorithm has been used by mathematicians in the idea of a constructive proof. The first known reference to the study of an algorithm is Euclid's work on the study of the greatest common divisor (GCD). This was done in approximately 300 B.C. Of course, Euclid's motivation was to reduce the amount of pencil-and-paper work in order to compute the GCD of two numbers. Interestingly, this algorithm has survived to the present day: approximately 2300 years later Euclid's algorithm and its variants still are the most efficient ones and in fact they are widely implemented in many computer programs. Knuth [29] calls Euclid's GCD algorithm the "granddaddy of all algorithms". The question whether this algorithm and its variants are the best possible stays as an important open problem. For a summary we refer the reader to [29].

The area of complexity theory is concerned with the fundamental problem: what makes some problems *difficult* to compute and what makes them *easy*. In pursuing this question, many important concepts have been introduced in theoretical computer science; we mention the concept of NP-complete problems, and circuit complexity. In general, many persons like to view the study of algorithms as providing upper bounds (in form of an efficient algorithm) and complexity theory as producing lower bounds (proof that certain functions actually are very difficult to compute). Of course, there is a lot of overlapping between these two areas such that distinction is not always clear.

This dissertation looks at problems in complexity theory; specifically, we are interested in finding explicit Boolean functions which are provably hard to compute in some reasonable model of computation. Although functions requiring exponential resources are known to be plentiful by elementary counting arguments [55], the problem of actually proving such lower bounds for explicit functions, say functions computable in NP, is much more difficult. For a function in NP, the best known lower bound on circuit size are only linear [2], the best known lower bounds on circuit depth are only logarithmic.

The two types of restrictions usually considered in complexity theory are monotone circuits and bounded depth circuits. In the first part of this thesis (Chapter 5) we concentrate on the depth restriction: we give (exponential) lower bounds on threshold circuits for some concrete Boolean functions.

One natural motivation to study threshold circuits comes from the fact that they are very closely related to neural nets (the basic element of a neural net is very similar to a threshold gate). As the term "neural net" implies, it was originally aimed towards modeling networks of real neurons in the human brain. However, the models are extremely simplified when seen from a neurophysiological point of view, though they are believed to be still of value to obtain insight into the principles of "biological computing" (see [19]).

To be more precise, we give exponential lower bounds on threshold circuits of depth 2 with bounded weights on the bottom layers. We look at very fundamental functions such as Boolean predicates of the discrete multiplication function, the discrete division function, the discrete exponentiation function, and the Diffie-Hellman function. The reasons for examining such functions are manifold.
First, the problems examined are of very fundamental structure, and sometimes they are used repeatedly in more complex calculations. Some of the considered functions are widely used in encryption schemes. In some cases, e.g. when embedded in hardware, every single gate may be expensive to build and an efficient implementation of such basic functions becomes very important. Our lower bounds give an answer to the question whether the implementation may be done with less hardware meaning *lower cost* and also *more efficiency*. Thus, it becomes very important to give more insight in understanding the *exact complexity* of such fundamental problems. Some of the examples for which we provide lower bounds are known to be computable with one more threshold layer with polynomially many gates. So our given bounds are very close to the best possible. Secondly, the analysis of easy functions adds important new tools to the techniques available to provide lower bounds for more complex functions. Third, some of the presented functions are actually used as the decryption function in various cryptographic protocols. Giving exponential lower bounds on the number of threshold circuits can also be viewed as a weak impossibility result to compute the function (to brake the cryptographic protocol) when the computational model is restricted (in our case it is restricted to threshold gates of depth 2). Thus presenting lower bounds on (very restricted) computational de-

vices may put a little more confidence in some widely believed assumptions that such functions are computationally hard. For example, one of such assumptions is the *CDH assumption* which states that the *Diffie-Hellman function* is hard to compute. Our results support this assumption in the sense that a Boolean predicate of the Diffie-Hellman functions is hard to compute when we (severely) restrict the model of computation. To put it in different words, we show that a certain kind of attack (namely an attack based on threshold circuits of depth 2) on the Diffie-Hellman function is doomed to fail since the number of necessary gates for this attack is too large. We want to stress that impossibility results of this kind should not be overestimated since the restricted model (based on threshold circuits of depth 2) seems to be far too weak to contain any practical cryptographic attack. However, since impossibility results (even in restricted models) are very rare, we think that they still are of some value.

In the second part of this thesis (Chapter 6) we look at interpolation results for (cryptographic) functions defined over a finite field. We give lower bounds on the degree and on the number of non-zero coefficients of such interpolation polynomials. Functions of interest will be functions related to the Diffie-Hellman function. Motivation of examining these function is given by the observation that whenever there is a simple (e.g. low degree) polynomial that coincides with such a functions on a sufficient large number of its inputs, then there exists an algorithm that computes the Diffie-Hellman function (and thus contradicting the DDH assumption).

Though the results seem to be similar in the two parts of this thesis, the methods are quite different. In the first part we pursue the strategy of giving lower bounds with algebraic and number theoretic methods. Specifically, we use the concept of matrices and matrix norms together with bounds on exponential sums. In the second part, the methods are more of combinatorial nature and results are obtained by counting arguments.

It is assumed that the reader is familiar with the basic notation and terminology used in complexity theory (such as the classes N and NP). Definitions of this terms can be looked up in widely used textbooks of theoretical computer science [29, 46].

Chapter 2

In mathematics you don't understand things. You just get used to them.

Johann von Neumann

Prerequisites from Algebra

In the first two chapters we recall basic definitions and facts, which, hopefully, will help the reader in understanding the remainder of the dissertation.

As usual $\mathbb{F}_q$ denotes a finite field of q elements, $\mathbb{Z}$ denotes the ring of integer numbers, $\mathbb{R}$ denotes the field of real numbers and $\mathbb{C}$ denotes the field of complex numbers.

For $N \geq 2$ we denote by $\mathbb{Z}_N$ the residue ring modulo N which we identify with the set $\{0, \ldots, N-1\}$. $\mathbb{Z}_N^*$ is the subset from $\mathbb{Z}_N$ consisting of all invertible elements. These are exactly the elements $x \in \mathbb{Z}_N$ with $\gcd(x, N) = 1$. The number of elements of $\mathbb{Z}_N^*$ is denoted by $\varphi(N)$, the Euler's Totient function. For the special case when $P = N$ is a prime, then $\mathbb{Z}_N^* = \{1, \ldots, N-1\}$. Similar to the notation $\mathbb{Z}_N^*$, we use $\{0, \ldots, N-1\}^*$ for the elements $x \in \{0, \ldots, N-1\}$ with $\gcd(x, N) = 1$.

2.1 Matrices and Matrix Norms

Let $\mathbb{F}$ be a field and let N_1 and N_2 be two natural numbers. A *matrix* over $\mathbb{F}$ is an N_1-by-N_2 array of elements from the field $\mathbb{F}$. The set of all N_1-by-N_2 matrices over $\mathbb{F}$ is denoted by $\mathbb{F}^{N_1 \times N_2}$. As usual, we denote matrices by capital letters. The entry of a matrix $A \in \mathbb{F}^{N_1 \times N_2}$ corresponding to the indices $0 \leq x \leq N_1 - 1$ and $0 \leq y \leq N_2 - 1$ is denoted by $A_{x,y}$.[1] The product BA of two matrices $B \in \mathbb{F}^{N_3 \times N_2}$ and $A \in \mathbb{F}^{N_2 \times N_1}$ is given by the matrix $(BA) \in \mathbb{F}^{N_3 \times N_1}$ with the entries $(BA)_{z,x} = \sum_{y=0}^{N_2-1} B_{z,y} A_{y,x}$.

If $A \in \mathbb{F}^{N_1 \times N_2}$ is given by the entries $A_{i,j}$, the transpose of A, denoted by $A^\top$, is the matrix $A^\top \in \mathbb{F}^{N_2 \times N_1}$ whose entries are $A_{j,i}$. That means, rows are exchanged for columns and vice versa. Note that $(A^\top)^\top = A$. The Hermitian adjoint A^* of $A \in \mathbb{C}^{N_1 \times N_2}$ is given by $A^* = \bar{A}^\top$, where $\bar{A}$ is the component-wise conjugate of the matrix A.

[1] Usually matrix indices range from 1 to N_1 and N_2, resp. But for our needs it will be more convenient to let the range start from 0.

We adopt the convention that $\mathbb{F}^{N_1} := \mathbb{F}^{N_1 \times 1}$ represents the vector space of column vectors over $\mathbb{F}$. In particular, the product of a matrix $B \in \mathbb{F}^{N_1 \times N_2}$ and a column vector $v \in \mathbb{F}^{N_2}$ is given by the column vector $Bv \in \mathbb{F}^{N_1}$ with the entries $(Bv)_z = \sum_{y=0}^{N_2-1} B_{z,y} v_y$. For a vector $x \in \mathbb{F}^{N_1}$, we denote $x^\top$ as its transpose.

For $x, y \in \mathbb{C}^{N_1}$, we define the *standard scalar product* $\langle x, y \rangle$ as

$$\langle x, y \rangle := x^* y = \sum_{i=0}^{N_1-1} \bar{x}_i y_i,$$

where $\bar{x}_i$ denotes the complex conjugate of $x_i \in \mathbb{C}$.

The *Euclidean norm* of a vector $x \in \mathbb{C}^{N_1}$ is given by

$$||x||_2 := \sqrt{\langle x, x \rangle}.$$

The *Cauchy-Schwarz Inequality* (or *Schwarz Inequality*) states that for all $x \in \mathbb{C}^N$ and $y \in \mathbb{C}^N$, we have

$$| \langle x, y \rangle | \leq \|x\|_2 \cdot \|y\|_2, \tag{2.1}$$

with equality if and only if $x = cy$ for some scalar c.

From now on we denote by the letter $\mathbb{K}$ either the field of real numbers $\mathbb{R}$ or of complex numbers $\mathbb{C}$. The *spectral* or *operator norm* of a matrix $A \in \mathbb{K}^{N_2 \times N_1}$ is defined as follows:

$$\|A\|_2 := \max_{\substack{u \in \mathbb{K}^{N_1} \\ \|u\|_2 = 1}} \|Au\|_2 \ .$$

Here, the maximum ranges over vectors with components in $\mathbb{K}$. The "max" in the above definition (rather than "sup") is justified since $\|Mu\|_2$ is a continuous function of $u \in \mathbb{K}^{N_1}$ and the unit ball of $\mathbb{K}^{N_1}$ is a compact set.

We briefly mention the following well-known facts concerning the spectral norm (see, e.g., [22]).

Fact 2.1.1. *The spectral norm of matrix $A \in \mathbb{K}^{N_1 \times N_2}$ satisfies the following equations:*

$$\begin{aligned}
\|A\|_2 &= \max_{\substack{u \in \mathbb{K}^{N_1}, v \in \mathbb{K}^{N_2} \\ \|u\|_2 = \|v\|_2 \leq 1}} |u^* A v| = \max_{\substack{u \in \mathbb{K}^{N_1}, v \in \mathbb{K}^{N_2} \\ \|u\|_2 = \|v\|_2 = 1}} |u^* A v| && (2.2) \\
&= \max\{\sqrt{\lambda} : \lambda \textit{ is an eigenvalue of } AA^*\} && (2.3) \\
&= \sqrt{\|AA^*\|_2} \ . && (2.4)
\end{aligned}$$

Obviously, by (2.2), the spectral norm of any submatrix B from A is upper bounded by the spectral norm of A:

$$\|B\|_2 \leq \|A\|_2. \tag{2.5}$$

The *Frobenius norm* of a matrix $A \in \mathbb{C}^{N_1 \times N_2}$ is defined as follows:

$$\|A\|_F := \left(\sum_{j=0}^{N_1-1} \sum_{k=0}^{N_2-1} A_{j,k}^2 \right)^{1/2} .$$

Note that $\|A\|_F = \sqrt{N_1 N_2}$ for each $A \in \{-1,1\}^{N_1 \times N_2}$.

Fact 2.1.2. *We have the following relation between the spectral norm and the Frobenius norm for each matrix $A \in \mathbb{K}^{N_1 \times N_2}$:*

$$\frac{\|A\|_F}{\sqrt{\operatorname{rank}(A)}} \leq \|A\|_2 \leq \|A\|_F \ . \tag{2.6}$$

Let J be the all-1's $N \times N$ matrix. As an immediate consequence of Fact 2.1.2 (or from the Cauchy-Schwarz inequality (2.1)), we obtain

$$\|J\|_2 = \max_{\substack{u,v \in \mathbb{K}^N \\ \|u\|_2 = \|v\|_2 = 1}} \left| \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} u_x v_y \right| = N \ . \tag{2.7}$$

By Fact 2.1.2 we can infer for square matrices with ± 1-entries

$$N^{1/2} \leq \|A\|_2 \leq N \ , \tag{2.8}$$

where $\|A\|_2 = N$ if $\operatorname{rank}(A) = 1$. It can be shown that equality $\|A\|_2 = N$ holds if *and only if* A has rank 1. Equality $N^{1/2} = \|A\|_2$ holds if and only if A has orthogonal rows or columns. This was noted, e.g., by Krause [31].

An example of ± 1 matrices with orthogonal rows and columns are the Hadamard matrices $H_{2^n} \in \{-1,1\}^{2^n \times 2^n}$. They are defined by the recursion

$$H_1 = 1 \ , \qquad H_{2^{n+1}} = \begin{pmatrix} H_{2^n} & H_{2^n} \\ H_{2^n} & -H_{2^n} \end{pmatrix} \ .$$

For example,

$$H_4 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix} \ .$$

As mentioned above, we have $\|H_{2^n}\|_2 = 2^{n/2}$.

A related measure to the spectral norm is the discrepancy of a matrix $A \in \{-1,1\}^{N \times N}$ which is given by

$$\operatorname{disc}(A) = \frac{1}{N^2} \max_{S,T} \left| \sum_{x \in S, y \in T} A_{x,y} \right| , \tag{2.9}$$

where the maximum is taken over all rectangles $S, T \subseteq \{0, \ldots, N-1\}$. The discrepancy measures how unbalanced a submatrix of A can be.

The maximum row sum matrix norm $\|A\|_\infty$ is defined on $A \in \mathbb{C}^{N_1 \times N_2}$ as

$$\|A\|_\infty = \max_{0 \leq x \leq N_1 - 1} \sum_{y=0}^{N_2 - 1} |A_{x,y}|.$$

It is well known that the *spectral radius* of matrix $A \in \mathbb{C}^{N_1 \times N_2}$ which is given by $\rho(A) = \max\{|\lambda|, \lambda \text{ is an eigenvalue of } A\}$ is a lower bound for any matrix norm $\|\!|\cdot|\!\|$:

$$\rho(A) \leq |\!|\!|A|\!|\!|.$$

Since the maximum row sum norm is a matrix norm (see [22]), we can derive the following simple (but useful) lemma:

Lemma 2.1.3. *For matrices $A \in \mathbb{C}^{N_1 \times N_2}$,*

$$|\!|\!|A|\!|\!|_2^2 \overset{(2.4)}{=} |\!|\!|AA^*|\!|\!|_2 = \rho(AA^*) \leq |\!|\!|AA^*|\!|\!|_\infty.$$

2.2 Fourier Expansion

Let $L_n = \{f : \{0,1\}^n \to \mathbb{R}\}$ denote the set of real valued functions on $\{0,1\}^n$. Though our interest is in Boolean functions, it is temporarily convenient to consider this richer space. L_n is a vector space over $\mathbb{R}$ of dimension 2^n, and has a natural inner product: for $f, g \in L_n$, we define

$$\langle f, g \rangle = \frac{1}{2^n} \sum_{x \in \{0,1\}^n} f(x)g(x) \ .$$

For $u \in \{0,1\}^n$, we define the function $\chi_u : \{0,1\}^n \to \mathbb{R}$ as $\chi_u(x) = (-1)^{\langle u,x \rangle}$. These function χ_u are the characters of $\{0,1\}^n$. Among their many wonderful properties (see, e.g., [59]) is the fact that the characters $\{\chi_u : u \in \{0,1\}^n\}$ form an orthonormal basis for L_n. Thus, any function $f \in L_n$ can be written in terms of this basis:

$$f = \sum_{u \in \{0,1\}^n} \hat{f}_u \chi_u, \tag{2.10}$$

where

$$\hat{f}_u = \langle f, \chi_u \rangle = \frac{1}{2^n} \sum_{x \in \{0,1\}^n} f(x) \cdot (-1)^{\langle u,x \rangle}$$

are called the *Fourier coefficients* or *spectrum* of f. As we have observed above, these coefficients uniquely determine the function f. We are interested in the L_1, L_2 and L_∞ norms associated with the spectrum. Namely,

$$L_1(f) = \sum_{u \in \{0,1\}^n} |\hat{f}_u|, \ \ L_2(f) = \sum_{u \in \{0,1\}^n} (\hat{f}_u)^2, \ \ L_\infty(f) = \max_{u \in \{0,1\}^n} |\hat{f}_u| \ .$$

For the L_2 norm, we have the Parseval identity

$$L_2(f) = \frac{1}{2^n} \sum_{x \in \{0,1\}^n} f^2(x) \ . \tag{2.11}$$

2.3 Boolean Functions

A *Boolean function on n variables* is a mapping

$$f_n : \{0,1\}^n \to \{0,1\} \ .$$

For our analysis it will be convenient to consider ± 1-valued functions rather than $0,1$-valued matrices. We identify 0 with the value 1, and 1 with the value -1. We can associate a Boolean function f_n on n variables with the ± 1-valued function F:

$$F : \{0,1\}^n \to \{-1,1\},$$

where $F(x)$ is given by $F(x) = (-1)^{f(x)}$.

By (2.11) it turns out that $L_2(F)$ always equals to one for any ± 1 valued function F. In fact, this implies that

$$L_1(F) \geq 1 \geq L_\infty(F), \quad L_1(F) \cdot L_\infty(F) \geq 1 \ .$$

In general, this spectral approach is a very useful tool in the study of Boolean functions (see e.g. [51]).

Most of the times we will consider Boolean functions on $2n$ variables

$$f_n : \{0,1\}^n \times \{0,1\}^n \to \{0,1\} \ .$$

They can be identified in a natural way by a $2^n \times 2^n$ matrix. Again it will be convenient to consider ± 1-valued matrices rather than $0,1$-valued matrices. We make the following definition:

Definition 2.3.1. *Let $2^{n-1} < N \leq 2^n$. We say a matrix $A \in \{-1,1\}^{N\times N}$* represents a Boolean function $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ *on $\{0,\ldots,N-1\} \times \{0,\ldots,N-1\}$, if*

$$A_{x,y} = (-1)^{f(x,y)}$$

for all $x, y \in \{0,\ldots,N-1\}$.

Let matrix A represent a Boolean function $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ on $\{0,\ldots,2^n-1\} \times \{0,\ldots,2^n-1\}$. By (2.10), we can write the entry $A_{x,y}$ of matrix A by its sum of the Fourier coefficients

$$A_{x,y} = \sum_{u\in\{0,1\}^n} \sum_{v\in\{0,1\}^n} \hat{A}_{u,v} (-1)^{\langle u,x\rangle + \langle v,y\rangle},$$

where the Fourier coefficients of A are given by

$$\hat{A}_{u,v} := \frac{1}{2^{2n}} \sum_{x\in\{0,1\}^n} \sum_{y\in\{0,1\}^n} A_{x,y} (-1)^{\langle u,x\rangle + \langle v,y\rangle}.$$

The L_1 and L_∞ norm of matrix A are then given by

$$L_1(A) = \sum_{u,v\in\{0,1\}^n} |\hat{A}(u,v)|, \quad L_\infty(A) = \max_{u,v\in\{0,1\}^n} |\hat{A}(u,v)| \ .$$

Note that the definition of the L_1 and L_∞ norm makes only sense for matrices whose dimension is a power of 2.

2.4 Relations between Norms, Discrepancy, and Fourier Spectra

We will use quadratic matrices $A \in \{-1,1\}^{N\times N}$ as our fundamental analyzing tool to derive complexity theoretic bounds. In this section we will collect some properties about these matrices.

It was noted by Forster [11], that as a simple consequence of Hoeffding's Inequality (see [20]) and the Gerschgorin Circle theorem (see, e.g., [22]) the spectral norm of a random ± 1 matrix is in probability bounded away from N:

Lemma 2.4.1. *Let the entries of the matrix $A \in \{-1,+1\}^{N\times N}$ be independently and identically distributed with* $\Pr(A_{ij}=1)=\frac{1}{2}$. *Then with probability* $1-N^2/\exp(N^{1-4\varepsilon})$, $\|A\|_2 \le \sqrt{2}N^{1-\varepsilon}$.

Rewriting the definition of the discrepancy (2.9) in the form

$$\operatorname{disc}(A) = \frac{1}{N^2}\max_{x,y\in\{0,1\}^N}|x^\top Ay| = \frac{1}{N}\max_{x,y\in\{0,1/\sqrt{N}\}^N}|x^\top Ay|, \tag{2.12}$$

we see that the discrepancy of a ± 1-valued matrix can be viewed as a sort of "discrete spectral norm", scaled by the factor $1/N$ (observe that for every vector $x \in \{0,1/\sqrt{N}\}^N$, $\|x\|_2 \le 1$). By Fact 2.1.2 it follows immediately that $\operatorname{disc}(A) \le \frac{\|A\|_2}{N}$. Surprisingly, Shaltiel [53] could show that low discrepancy also implies a low spectral norm, though in a weak sense:

$$\Omega\left(\left(\frac{\|A\|_2}{N}\right)^3\right) \le \operatorname{disc}(A) \le \frac{\|A\|_2}{N} \ . \tag{2.13}$$

In fact, Shaltiel's result was motivated by the following result from Nisan and Wigderson [44] giving a relation[2] between the discrepancy and the rank of a matrix $A \in \{-1,1\}^{N\times N}$:

$$\frac{1}{\operatorname{disc}(A)} = O(\operatorname{rank}(A)^{3/2}) \ . \tag{2.14}$$

Note that this result can also be obtained by combining Fact 2.1.2 with bound (2.13).

Let $A \in \{-1,1\}^{2^n\times 2^n}$ be a matrix. Since every Fourier coefficient $\hat{A}_{u,v}$ of A can be written in the form

$$\begin{aligned}\hat{A}_{u,v} &= \frac{1}{2^{2n}}\sum_{x,y\in\{0,1\}^n}(-1)^{\langle x,u\rangle}A_{x,y}(-1)^{\langle y,v\rangle}\\ &= \frac{1}{2^n}\sum_{x,y\in\{0,1\}^n}\frac{(-1)^{\langle x,u\rangle}}{2^{n/2}}A_{x,y}\frac{(-1)^{\langle y,v\rangle}}{2^{n/2}}\\ &= \frac{1}{2^n}|\alpha^\top A\beta|,\end{aligned} \tag{2.15}$$

[2] The original result holds only for $0,1$ matrices; but considering ± 1 matrices rather than $0,1$ matrices obviously changes the rank by at most 1, and does not change the discrepancy.

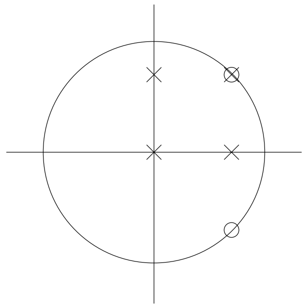

Figure 2.1: The three measures $\|A\|_2, \mathrm{disc}(A)$ and $L_\infty(A)$-norm.

with vectors $\alpha_x = \frac{(-1)^{\langle x,u\rangle}}{2^{n/2}}, \beta_y = \frac{(-1)^{\langle y,v\rangle}}{2^{n/2}}$ and $||\alpha||_2, ||\beta||_2 \leq 1$, we infer

$$L_\infty(A) \leq \frac{\|A\|_2}{2^n} \quad . \tag{2.16}$$

Similar considerations lead to

$$L_\infty(A) \leq 4\mathrm{disc}(A) \quad .$$

All three measures, the spectral norm, the discrepancy and the L_∞ norm of matrices $A \in \{-1,1\}^{2^n \times 2^n}$ can be written in the form $\max |x^\top Ay|$, where the maximum is taken over all vectors x, y from a certain subset of $\mathbb{R}^{2^n}$. The difference in the three definitions is the domain over which the maximum is taken, see Figure 2.1. We point out the difference.

For the spectral norm, the vectors are maximized over the unit ball. For the discrepancy and the L_∞ norm, the domain is discrete. By equation (2.12) the discrepancy (scaled by the factor 2^n) is maximized over all vectors from $\{0, 2^{-n/2}\}^{2^n}$ (represented by $\times$ in the figure). Similarly, by equation (2.15) the L_∞ norm (again scaled by 2^n) is maximized over a subset (e.g., a basis) of vectors from $\{2^{-n/2}, -2^{-n/2}\}^{2^n}$ (represented by $\circ$).

2.5 Exponential Sums

We will use exponential sums as a fundamental tool to analyze the spectral norm of given matrices. Nearly all results given here can be looked up in one of the textbooks [30, 60]. Since in the sequel we will not demand a lot of insight into the field of exponential sums, the reader may run quickly through this section.

For each $a \in \mathbb{Z}_N$, we define the following function $E_a^N : \mathbb{Z}_N \to \mathbb{C}$:

$$E_a^N(x) = e^{\frac{2\pi i a x}{N}} = \cos\frac{2\pi a x}{N} + i \cdot \sin\frac{2\pi a x}{N} \quad . \tag{2.17}$$

In the sequel, we drop superscript N when it is obvious from context. Clearly, the following rules are valid:

$$\begin{aligned}
|E_a(x)| &= 1 && (2.18)\\
E_a(x) &= E_x(a) && (2.19)\\
E_a(x+y) &= E_a(x)\cdot E_a(y) && (2.20)\\
E_a(x-y) &= E_a(x)\cdot E_{-a}(y) && (2.21)\\
\overline{E_a(x)} &= E_{-a}(x) && (2.22)\\
\sum_{a\in\mathbb{Z}_N} E_{-a}(y) &= \sum_{a\in\mathbb{Z}_N} E_a(y) && (2.23)
\end{aligned}$$

By simple trigonometrical transformations, we can express $|E_a(z)\pm 1|$ in terms of the cosinus and the sinus, respectively:

$$|E_a(z)+1| = |2\cos(\pi az/N)|, \quad |E_a(z)-1| = |2\sin(\pi az/N)|. \tag{2.24}$$

The following identity is well-known (see, e.g., [30]):

$$\frac{1}{N}\sum_{x\in\mathbb{Z}_N} E_a(x) = \begin{cases} 0 & \text{if } a\neq 0,\\ 1 & \text{if } a=0. \end{cases} \tag{2.25}$$

From (2.25), we can immediately infer the following identity, which holds for each binary function $B:\mathbb{Z}_N\to\{-1,1\}$:

$$B(z_0) = \frac{1}{N}\sum_{z\in\mathbb{Z}_N}\sum_{a\in\mathbb{Z}_N} B(z)E_a(z_0-z)\ . \tag{2.26}$$

It will be important to study exponential sums of the form $\sum_{x=1}^{P} E_a(x)$. If $a=0 \bmod N$, then obviously $\sum_{x=1}^{P} E_a(x) = P$. On the other hand, as it can by looked up in [30], for every integer $a\in\mathbb{Z}_N\setminus\{0\}$ we have

$$\sum_{x=1}^{P} E_a(x) = \frac{E_a(P)-1}{E_a(1)-1}E_a(1)\ . \tag{2.27}$$

As an easy consequence of this (see [60], Chapter 3, Ex. 11c) we get: For every integer $N>6$ let the functions $M(a)$ and $P(a)$ such that $P(a)\geq 0$ for values $a\in\{1,\ldots N-1\}$. Then the following holds:

$$\frac{1}{N}\sum_{a=1}^{N-1}\left|\sum_{x=M(a)}^{M(a)+P(a)} E_a(x)\right| \leq c\log N\ , \tag{2.28}$$

where $\log N=\log_2 N$ denotes the binary logarithm of N and the constant c is given by $c=\ln 2\approx 0.6931$. As another easy consequence of (2.27) we get the following:

$$\max_{a\in\mathbb{Z}_N\setminus\{0\}}\left|\frac{E_a(P)-1}{E_a(1)-1}\right| = \left|\sum_{z=1}^{P} E_a(z)\right| \leq \min\{P, N-P)\}\ . \tag{2.29}$$

We have the following Lemma for Gaussian sums [30, Theorem 3]:

Lemma 2.5.1. *For any prime N and $a \in \mathbb{Z}_N^*$, we have*

$$\left| \sum_{x \in \mathbb{Z}_N} E_a(x^2) \right| = N^{1/2} \quad .$$

We will also need the Weil bound for exponential sums [62] which we present in the following form (see also [30]):

Lemma 2.5.2. *For any prime N, $a \in \mathbb{Z}_N^*$, and any a_i with* $\gcd(a_0, a_1, \ldots, a_k, N) = 1$, *we have*

$$\left| \sum_{x \in \mathbb{Z}_N} E_a(a_0 + a_1 x + \ldots + a_k x^k) \right| \leq (k-1) N^{1/2} \quad .$$

The road to hell is paved with NAND gates.

J. GOODING

Chapter 3

Prerequisites from Complexity Theory

3.1 Asymptotics

We want to remind the reader that the notations $f(n) = O(g(n))$ or $g(n) = \Omega(f(n))$ mean that $\limsup_{n\to\infty}(f(n)/g(n)) < \infty$. Also, $f(n) = o(g(n))$ if $\lim_{n\to\infty}(f(n)/g(n)) = 0$. Furthermore, if $f(n) = O(g(n))$ and $g(n) = O(f(n))$, then $f(n) = \Theta(g(n)$ and $g(n) = \Theta(f(n))$.

We will somehow have a loose handling of asymptotic notation. If we express a result in terms of $f = O(g)$, then this always means that in fact $f = f(n)$ and $g = g(n)$ are families of functions indexed by n and hence $f(n) = O(g(n))$ holds. We will sometimes abandon this long notation and make the convention that whenever we use the notation of asymptotics (e.g., $O(f), \Omega(f)$, or $o(f)$), then $f = f(n)$ implicitly is defined as a family of functions. The parameter n (sometimes we will also use N or T) will always be clear from the context.

3.2 Circuit Complexity

We will give a quick overview on the complexity of Boolean functions. For a more detailed introduction we refer the reader to one of the textbooks [59, 61]. See also [32] for a recent survey.

A circuit is designed to compute a Boolean function $f : \{0,1\}^n \to \{0,1\}^m$. We will concentrate on the special case of $m = 1$. A circuit model needs a basis Ω of elementary one-output Boolean functions which can be realized by simple gates. In the basic circuit model a basis is a finite set (that may also depend on n).

Definition 3.2.1. *An Ω-circuit works for a fixed number n of Boolean input variables $x_1, \ldots, x_n$. It consists of a finite number b of gates $G(1), \ldots, G(b)$. Gate $G(i)$ is defined by its type $\omega_i \in \Omega$ and, if $\omega_i : \{0,1\}^{n(i)} \to \{0,1\}$, some*

$n(i)$-tuple $(P(1),\dots,P(n(i)))$ of predecessors. $P(j)$ may be some element from $\{0,1,x_1,\dots,x_n,G(1),\dots,G(i-1)\}$. By $\mathrm{res}_{G(i)}$ we denote the Boolean function computed at $G(i)$. res is defined inductively. For an input I, res_I is equal to I. If $G(i)=(\omega_i,P(1),\dots,P(n(i)))$, then

$$\mathrm{res}_{G(i)}(x)=\omega_i(\mathrm{res}_{P(1)}(x),\dots,(\mathrm{res}_{P(n(i)))}(x)).$$

Finally, the output y, where y is some input or some gate, describes what the circuit computes, namely $f:\{0,1\}^n\rightarrow\{0,1\}$.

The *fan-in of the basis* Ω is defined to be the largest number of inputs to a function of Ω. The *fan-out of an Ω-circuit* is the maximal amount that a results $\mathrm{res}_{G(i)}$ of a gate $G(i)$ is used. It was shown in [21] that as long as the fan-in of the basis Ω is bounded, general circuits with unbounded fan-out can be simulated by circuits with fan-out bounded by 2 such that size and depth grow by a small constant factor only.

We say that the circuit *realizes* or *computes* f if $f(x_1,\dots,x_n)$ equals the output of the circuit for all inputs $x\in\{0,1\}^n$. The *size* of the circuit equals the number of its gates. Obviously, every circuit can be represented by a directed acyclic graph with the sources $x_1,\dots,x_n,0,1$. The *depth* of an Ω-circuit equals the length of the longest path in its graph representation. The *circuit size* $C_\Omega(f)$ of f with respect to a chosen basis Ω is the smallest size of circuits containing instructions from Ω realizing the function f. The *circuit depth* $D_\Omega(f)$ is defined analogously.

There are different possible choices for the basis Ω. A basis is called *complete* if each Boolean function can be represented by an Ω-circuit. A typical complete basis is the basis containing (binary) AND, OR, and (unary) NOT. Only NAND is a complete basis, too. An interesting non complete basis is the *monotone basis* which consists of AND and OR. Exactly monotone functions can be computed by monotone circuits.

It is a simple fact that circuit size and depth with respect to different complete basis (with bounded fan-in) are related by a fixed multiplicative constant, only.

Small circuits in terms of depth and size imply an efficient way to compute the corresponding function. There are two main branches of research in circuit complexity. In one branch the aim is to construct small circuits for specific functions.
The other branch is to obtain lower bounds on the size and depth for various restricted circuit models. We will survey some results concerning lower bounds in circuit complexity.

It is easy to prove that there exist Boolean functions whose circuit size (for large n) is at least $2^n/n$ [55]. However, people are more interested in lower bounds for "concrete" or "explicitly defined" functions. For a complete basis and functions from NP, the best result as yet is still the lower bound of Blum [2] who presented a family of Boolean functions that can only be computed if the circuit has size at least $3n-o(n)$. Since this is the best we have so far, the state of the art is very poor: There is no proof yet of a nonlinear lower bound on the circuit size of a concrete Boolean function $f : \{0,1\}^n \rightarrow \{0,1\}$. Even for some restricted models of computation, research process may get stuck as long as the right tools for proving lower bounds are missing. For instance, there was no progress in proving exponential lower bounds on the size of *monotone* Boolean circuits until Razborov [50] discovered his celebrated lower bound on the size of *monotone circuits* that decide the Clique problem. Perhaps it is simpler to obtain lower bounds on the size if the depth is restricted? The complexity class AC^k (alternating class) consists of all Boolean functions computable in depth $O(\log^k n)$ (constant depth for $k = 0$) and polynomial size by circuits allowing negation and unbounded fan-in AND and OR. The parity function PAR of n inputs is given by $\mathrm{PAR}(x_1, \ldots, x_n) = \sum_{i=1}^{n} x_i \bmod 2$. It was a breakthrough in the eighties [14] to prove that simple functions such as the parity function PAR are not contained in AC^0.

3.3 Threshold Circuits

When considering lower bounds in unbounded fan-in circuits, the choice of the basis Ω becomes essential. Another very important (complete) basis consists of *linear threshold functions*. A linear threshold function $G_\alpha : \{0,1\}^n \rightarrow \{0,1\}$ with respect to the integer weights $\alpha_1, \ldots, \alpha_n$ and threshold α_0 is a function defined as

$$G_\alpha(x_1, \ldots, x_n) = \mathrm{sign}\Big(\sum_{i=1}^{n} \alpha_i x_i - \alpha_0\Big) \ ,$$

where the signum function $\mathrm{sign} : \mathbb{R} \rightarrow \{0,1\}$ is given by

$$\mathrm{sign}(x) = \begin{cases} 1 \ , & x \geq 0 \ , \\ 0 \ , & x < 0 \ . \end{cases}$$

Threshold gates can simulate NOT, AND, and OR. Therefore, we investigate *threshold circuits* containing only threshold gates.

For a threshold gate, the *weight* is the maximum of the absolute values of all integer weights of the gate. There is a difference when considering threshold functions of bounded and unbounded weight. One threshold gate is enough to compute the carry bit CAR of the sum of two n-bit numbers: $\mathrm{CAR}(x, y) = \mathrm{sign}(\sum_{i=1}^{n} 2^i \cdot (x_i + y_i) - 2^n)$. However, this gate has 2^{n-1} as maximal weight. Hence, we are interested in the following two complexity classes:

- LT_k: all $0,1$-valued functions computable by threshold circuits of polynomial size, depth k and unbounded integer weights.

- TC_k: all $0,1$-valued functions computable by threshold circuits of polynomial size, depth k and polynomially bounded integer weights. The class TC_k is sometimes also denoted by $\widehat{\mathrm{LT}}_k$.

Let us discuss the relation between TC_k and LT_k. By definition, $\mathrm{TC}_k \subseteq \mathrm{LT}_k$. It was proven by Goldmann, Håstad, and Razborrov [35] that $\mathrm{LT}_k \subseteq \mathrm{TC}_{k+1}$.

As we saw, $\mathrm{TC}_k \subseteq \mathrm{LT}_k$. Is this inclusion strict? At least for the case $k = 1$ and $k = 2$ we can answer this question positively: It is not hard to show that the function CAR of the introductory example above indeed cannot be realized with polynomial weights (see e.g. [59]) and thus $\mathrm{TC}_1 \subsetneq \mathrm{LT}_1$. In [18] the classes TC_2 and TC_3 were separated by the inner product function which is given by $\mathrm{IP}(x, y) = \langle x, y\rangle \bmod 2 = \sum_{i=1}^{n} x_i y_i \bmod 2$. An example (from [35]) showing that TC_2 is a proper subclass of LT_2 will be given in Section 3.4.

These separation results motivate the examination of complexity classes that are in between TC_k and LT_k. For $k \geq 2$, we define the complexity classes

- $\widetilde{\mathrm{TC}}_k$: all $0, 1$-valued functions computable by threshold circuits of polynomial size, depth k, where the top layer has polynomially bounded integer weights and other layers have unbounded integer weights.
- $\widetilde{\mathrm{LT}}_k$: all $0, 1$-valued functions computable by threshold circuits of polynomial size, depth k, where the bottom layer has polynomially bounded integer weights and the remaining gates have unbounded integer weights.

Clearly, $\mathrm{TC}_k \subseteq \widetilde{\mathrm{TC}}_k$ and $\widetilde{\mathrm{LT}}_k \subseteq \mathrm{LT}_k$. It turned out surprisingly that unbounded weights on a non-top layer cannot be used to define new functions outside TC_k. For every constant k it was proven in [35] that $\mathrm{TC}_k = \widetilde{\mathrm{TC}}_k$.

We get the following hierarchy which we zoom in for the case $k = 1, 2$, and 3:

$$\mathrm{TC}_1 \subsetneq \mathrm{LT}_1 \subsetneq \mathrm{TC}_2 \subsetneq \widetilde{\mathrm{LT}}_2 \subseteq \mathrm{LT}_2 \subseteq \mathrm{TC}_3 \subseteq \ldots \tag{3.1}$$

We want to point out that the indicated inclusions are proper. The reasons are the following:

- The classes TC_1 and LT_1 are separated by CAR [59].
- The classes LT_1 and TC_2 are separated by the parity function PAR [18].
- In [12], $\mathrm{IP} \notin \widetilde{\mathrm{LT}}_2$ was shown and hence (since $\mathrm{IP} \in \mathrm{TC}_3$ [18]) we have $\widetilde{\mathrm{LT}}_2 \subsetneq \mathrm{TC}_3$. In Section 5.7 we will present new examples of binary functions in $\mathrm{TC}_3 \setminus \widetilde{\mathrm{LT}}_2$.
- In Section 3.4 we will argue that TC_2 is a proper subclass of $\widetilde{\mathrm{LT}}_2$.

As we will explain in Section 4.1, there are good techniques for showing that a function is not contained in $\widetilde{\mathrm{LT}}_2$. On the other hand there is no explicit function from NP known yet that is not contained in LT_2. The question whether LT_2 is different from higher levels of the hierarchy (or whether it contains all of NP) is still open. So, the frontier between solved and unsolved problems concerning threshold circuits lies in the region of depth 2. Circuits of depth 3 is also the matter of much current research since TC_3 seems to be a very rich class of functions (see [59]).

3.4 The Complexity Class $\widetilde{\mathrm{LT}}_2$

In this section we want to justify the position of the complexity class $\widetilde{\mathrm{LT}}_2$ in hierarchy (3.1), i.e. we show the separation $\mathrm{TC}_2 \subsetneq \widetilde{\mathrm{LT}}_2$. As we will see this can be done with methods that already appeared in the literature. However, since the results are not explicitly stated there, we quickly discuss the arguments.

At this point we introduce the following two complexity classes:

- PT_1: all $\{0,1\}$-valued functions that can be represented as the sign of a sparse polynomial, i.e. all functions f that can be written in the form

$$f(x) = \mathrm{sign}\left(\sum_{\alpha \in S} w_\alpha x^\alpha\right), \quad S \subseteq \{0,1\}^n, \quad |S| \leq \mathrm{poly}(n),$$

  with the convention that for $\alpha \in \{0,1\}^n$, x^α denotes the ± 1-valued parity function $x^\alpha = \chi_\alpha(x) = (-1)^{\langle x, \alpha \rangle} \in \{-1, 1\}$.

- PL_∞: all $\{0,1\}$-valued functions whose Fourier expansion has L_∞-norm that is bounded from below by $1/p(n)$ for some polynomial p.

In [4] the inclusion $\mathrm{PT}_1 \subseteq \mathrm{PL}_\infty$ was shown. Furthermore it was also shown in [4] that PT_1-circuits can be simulated by LT_2-circuits. This follows by the observation that, for fixed $\alpha \in \{0,1\}^n$, the parity function $(-1)^{\langle x, \alpha \rangle}$ can be written as the sum over $2\lceil k_\alpha/2 \rceil$ linear threshold functions of unit weight, where $k_\alpha \leq n$ equals the number of one's in $\alpha \in \{0,1\}^n$ (Hamming weight of α). By this argument it is easy to see that $\widetilde{\mathrm{LT}}_2$-circuits are already sufficient to simulate PT_1-circuits as we note in the following lemma:

Lemma 3.4.1. $\mathrm{PT}_1 \subseteq \widetilde{\mathrm{LT}}_2$.

To show a separation between the two complexity classes we consider the complete quadratic function introduced in [4]:

$$\mathrm{CQ}(x_1, \ldots, x_k) = (x_1 \wedge x_2) \oplus (x_1 \wedge x_3) \oplus \cdots \oplus (x_{n-1} \wedge x_n).$$

That is, CQ consists of the parity of all the $\binom{n}{2}$ AND's between pairs of variables. In [4] it was shown that $\mathrm{CQ} \in \mathrm{TC}_2 \setminus \mathrm{PL}_\infty$. With $\mathrm{TC}_2 \subseteq \widetilde{\mathrm{LT}}_2$ and $\mathrm{PT}_1 \subseteq \mathrm{PL}_\infty$, we get the separation $\mathrm{CQ} \in \widetilde{\mathrm{LT}}_2 \setminus \mathrm{PT}_1$.

Obviously, $\mathrm{TC}_2 \subseteq \widetilde{\mathrm{LT}}_2$. Is this inclusion proper? It was shown in [35] that the function

$$p(x, y) = \mathrm{sign}\left(1 + 2\sum_{i=0}^{n-1}\sum_{j=0}^{2n-1} 2^i y_j (x_{i,2j} + x_{i,2j+1})\right)$$

is contained in $\mathrm{PT}_1 \setminus \mathrm{TC}_2$ and hence by Lemma 3.4.1 we have $\mathrm{TC}_2 \subsetneq \widetilde{\mathrm{LT}}_2$.

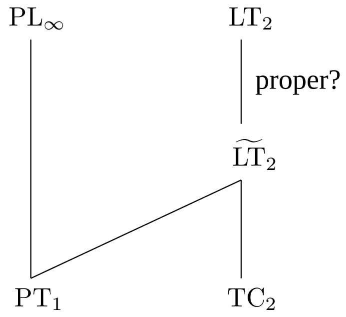

Figure 3.1: The complexity hierarchy.

It is still open problem whether $\widetilde{\mathrm{LT}}_2 \subseteq \mathrm{LT}_2$ is a proper inclusion or if $\widetilde{\mathrm{LT}}_2$-circuits are already as powerful as LT_2-circuits.

We summarize the relations in Figure 3.1. By counting arguments it was argued in [35] that $\mathrm{PL}_\infty \not\subseteq \mathrm{LT}_2$. Note that now the picture reflects all possible inclusions between the five classes shown there.

3.5 Communication Complexity

In this section we recall the concept of *communication complexity*. For the reader it will not be important to completely understand all formal aspects. A more formal treatment can be looked up in the textbook of Kushlevitz and Nisan [33].

Let there be two parties, one (Alice) knows a value x and the other (Bob) knows a value y where one party has no information about the others value. The common goal is to create a communication protocol P between Alice and Bob where at least one party at the end is able to compute a public Boolean function $f(x, y)$. Both parties have *unbounded computational power*. The communication between Alice and Bob takes place in rounds. In each round, one of the parties is active, in odd rounds Alice and in even rounds Bob. The active party chooses a message (that may depend on the known part of the input and on all prior sent and received messages) according to the communication protocol P. In the final round the active party outputs the result of the computation.

The length of such a communication protocol P is $\lfloor \log N \rfloor$, where N is the number of distinct message sequences exchanged that can occur in computations that follow the protocol. The largest possible length of such a protocol P, taken over all possible inputs x and y is called the communication complexity of P. The smallest possible value taken over all possible protocols P that correctly computes the function f for all inputs is called the *communication complexity* of the function f.

Here we consider two different types of communication complexity. For the *deterministic communication complexity*, C_f, we require the two parties to choose their messages in a deterministic way. The book of Kushlevitz and Nisan [33] gives a comprehensive overview on this topic.

Now we consider the *probabilistic communication complexity with unbounded error*, $\tilde{C}_f$. Again, the communication takes place in rounds, and the message chosen by the active party in the communication protocol may also depend on random coin tosses. In the final round the active party probabilistically chooses the result of the computation. For a *one-way probabilistic communication protocol* there are only two rounds: Alice randomly selects a message, sends it to Bob, and Bob randomly performs the binary decision of either accepting the input or rejecting it. We require the protocol P to compute for all possible inputs x and y the correct $f(x, y)$ with probability greater than half, where the probability is taken over all random coin flips made by the two parties Alice and Bob. Since the slightest edge over randomly guessing the correct value $f(x, y)$ is sufficient, this model is called the *unbounded error model*. We briefly want to note that in the unbounded error model, one-way and two-way protocols have almost the same power. In fact, any probabilistic two-way protocol of length l can be converted into a one-way protocol of length at most $l + 1$. This was shown by Paturi and Simon [47].

If I have seen farther than others,
it is because I was standing on
the shoulder of giants.

Isaac Newton

If I have not seen as far as others,
it is because giants were standing
on my shoulders.

Hal Abelson

Chapter 4

A Theorem from Complexity Theory and its Implications

4.1 A Lower Bound Method

As we have seen, the frontier between solved and unsolved problems concerning threshold circuits lies in the region of depth 2 and 3. In this section we review a method from [12] that allows to show that some concrete functions are not contained in the complexity class $\widetilde{\mathrm{LT}}_2$.

As mentioned before in Section 2.1, the spectral norm of a matrix $A \in \{-1,1\}^{N\times N}$ satisfies $\sqrt{N} \le \|A\|_2 \le N$. The extreme case $\|A\|_2 = \sqrt{N}$ corresponds to the case where A has orthogonal rows or columns. On the other extreme $\|A\|_2 = N$ corresponds to the case where $\mathrm{rank}(A) = 1$.

As explained below (Corollary 4.1.4), by a technique from [12], a function whose corresponding spectral norm is upper bounded by $N^{1-\varepsilon}$, cannot be contained in the complexity class $\widetilde{\mathrm{LT}}_2$. By means of this result, membership in $\widetilde{\mathrm{LT}}_2$ can be ruled out in a convenient manner.

Definition 4.1.1. *A* k-dimensional linear arrangement *representing a matrix* $A \in \{-1,+1\}^{N\times N}$ *is given by collections of vectors* u_x *and* v_y *from* $\mathbb{R}^k$, $x, y \in \mathbb{Z}_N$, *such that* $\mathrm{sign}(\langle u_x, v_y\rangle) = A_{x,y}$ *for all* $x, y \in \{0,\dots,N-1\}$.

Theorem 4.1.2 (Forster [10]). *If a matrix* $A \in \{-1,+1\}^{N\times N}$ *can be represented by a* k*-dimensional linear arrangement then*

$$k \ge N/\|A\|_2 \ .$$

In the sequel, we briefly indicate the complexity theoretic implications of Forster's result.

To simplify notation we will make no difference between the vector notation of $x = (x_1, \ldots, x_n) \in \{0,1\}^n$ and the integer notation of $x \in \{0, \ldots, 2^n - 1\}$. The connection between both notations is given by the equation $x = \sum_{i=1}^{n} x_i 2^i$.

The next conclusion is from [12] and gives nowadays the largest bounds for depth-2 threshold circuits. We present it in a slightly more general version[1].

Corollary 4.1.3. *Let (f_n) be a family of distributed Boolean functions $f_n : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$. Suppose (f_n) is computed by depth-2 threshold circuits (C_n) in which the top gate is a linear threshold gate using* unrestricted *weights and the bottom level has d_n linear threshold gates using weights whose absolute value is upper bounded by W. Circuits of this kind cannot compute (f_n) for all inputs $x, y \in \{0, \ldots, N-1\}$ ($2^{n-1} < N \leq 2^n$) unless*

$$d_n = \Omega\left(\frac{N}{nW\|A_N\|_2}\right) \quad ,$$

where $A_N \in \{-1,1\}^{N \times N}$ is the matrix representing f_n on input $\{0, \ldots, N-1\} \times \{0, \ldots, N-1\}$.

Obviously, Corollary 4.1.3 can be used to show that distributed functions whose spectral norm is less or equal $N^{1-\varepsilon}$ are not contained in the complexity class $\widetilde{\mathrm{LT}}_2$:

Corollary 4.1.4. *With the notation of Corollary 4.1.3 let the absolute value of the weights of the bottom gates be bounded by a polynomial in n and assume $\|A_N\|_2 = O(N^{1-\varepsilon})$ for a constant $0 < \varepsilon \leq 1/2$. Then the number of threshold gates d_n that is necessary to compute (f_n) is lower bounded by*

$$d_n = 2^{\varepsilon n - o(1)} \quad .$$

Hence, the family (f_n) is not contained in the complexity class $\widetilde{\mathrm{LT}}_2$.

Since for a matrix $A_n \in \{-1,1\}^{2^n \times 2^n}$ we have by (2.16) $L_\infty(A_n) \leq \|A_n\|_2 / 2^n$, we can make the following

Corollary 4.1.5. *Let (f_n) be a family of distributed Boolean functions $f_n : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$. Let A_n be the matrix representing f_n on $2^n \times 2^n$ and assume $\|A_n\|_2 \leq 2^{(1-\varepsilon)n}$ for a constant $0 < \varepsilon \leq 1/2$. Then (f_n) is not contained in the complexity class PL_∞.*

4.2 Communication Complexity

Deterministic communication complexity and the rank of the representing matrix are related by the well-know fact that $\log(\mathrm{rank}(A)) \leq C_f \leq \mathrm{rank}(A)$ (see [33], for instance). Using the relations from Nisan and Wigderson (2.14) and Shaltiel (2.13) we get the following:

[1] In the original version there were only matrices from $\{-1,1\}^{2^n \times 2^n}$ considered. In the sequel we will need to apply the result to matrices from $\{-1,1\}^{N \times N}$. However, the proof stays the same.

Corollary 4.2.1. *Let (f_N) be a family of distributed functions $f_N : \{0,\dots,N-1\} \times \{0,\dots,N-1\} \to \{0,1\}$. Then the deterministic communication complexity is at least*

$$C_{f_N} = \frac{2}{3}\log\left(\frac{1}{\operatorname{disc}(M)}\right) - O(1) = \frac{2}{3}\log\left(\frac{N}{\|M\|_2}\right) - O(1) \ ,$$

where $M = M_{f_N} \in \{-1,1\}^{N\times N}$ is the matrix representing the function f_N on $\{0,\dots,N-1\} \times \{0,\dots,N-1\}$.

Paturi and Simon [47] could prove that the minimal dimension k of a linear arrangement of a function f is closely related to its probabilistic communication complexity $\tilde{C}_f$:

$$\lceil \log k \rceil \le \tilde{C}_f \le \lceil \log k \rceil + 1 \ . \tag{4.1}$$

In [10], this observation was combined with Theorem 4.1.2 to get the following:

Corollary 4.2.2. *Let f be a distributed function $f : \{0,\dots,N-1\} \times \{0,\dots,N-1\} \to \{0,1\}$. Then the probabilistic communication complexity is at least*

$$\tilde{C}_f \ge \left\lceil \log_2 \frac{N}{\|A\|_2} \right\rceil \ ,$$

where $A \in \{-1,1\}^{N\times N}$ is the matrix representing the function f on $\{0,\dots,N-1\} \times \{0,\dots,N-1\}$.

Chapter 5

If it's provably secure,
it's probably not.

LARS R. KNUDSEN

Upper Bounds on the Spectral Norm

5.1 Introduction

Let $\mathcal{X}$ and $\mathcal{Y}$ be finite sets. In this chapter, we identify functions of the form $H : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ with matrices from $\mathcal{Z}^{\mathcal{X}\times\mathcal{Y}}$. Similarly, functions of the form $K : \mathcal{Z} \to \mathcal{K}$ are viewed as (column) vectors from $\mathcal{K}^{\mathcal{Z}}$. Note that the composition $K \circ H : \mathcal{X} \times \mathcal{Y} \to \mathcal{K}$ can then be identified with a matrix from $\mathcal{K}^{\mathcal{X}\times\mathcal{Y}}$. We are mainly interested in the following situation:

- $\mathcal{X} = \mathcal{Y} = \{0, \ldots, T-1\}$ (or $\mathcal{X} = \mathcal{Y} = \mathbb{Z}_T$) , $\mathcal{Z} = \mathbb{Z}_N$, and $H : \mathcal{X} \times \mathcal{Y} \to \mathbb{Z}_N$ is some function (in some cases we have $T = N$).
- In the role of K, we consider functions $B : \mathbb{Z}_N \to \{-1, 1\}$ and the functions $E_a : \mathbb{Z}_N \to \mathbb{C}$ from (2.17).

As noted in Lemma 2.4.1, almost all ±1-valued functions have a spectral norm upper bounded by $T^{1-\varepsilon}$. However, concrete functions with such a bound on the spectral norm generally are not easy to find. The main results of this chapter (given in Section 5.4) will be non-trivial upper bounds on the spectral norm of $B \circ H$ where B and H are as described above. As explained in Chapter 4, if the spectral norm of $B \circ H$ is upper bounded by $T^{1-\varepsilon}$ for a constant ε, we may then conclude that

- $B \circ H$ is hard to compute on some restricted computational devices.
- The deterministic and probabilistic communication complexity is linear.

Functions E_a (see Section 2.5) are used as a mathematical tool within the analysis of the computational complexity of $B \circ H$. We pursue the strategy to bound the spectral norm of $B \circ H$ in terms of the spectral norm of $E_a \circ H$ (and some other terms). The hope is that "good" upper bounds on the spectral norm of $E_a \circ H$ are known (or easy to compute) and that they can be converted into "good" upper bounds on the spectral norm of $B \circ H$.

A quick note on the motivation of using the functions E_a in our analysis. Our techniques are specialized on functions H doing computations in the residue ring $\mathbb{Z}_N$ (e.g. discrete multiplication and discrete exponentiation in $\mathbb{Z}_N$). Since the functions E_a (by definition) compute *modulo* N in their argument, they provide a natural way to model the desired "computing in $\mathbb{Z}_N$". This fact turns them the into a very useful tool for our analysis.

For some special cases of functions B and H, an upper bound on the spectral norm of $B \circ H$ can already be achieved with results known in the literature. An example is the matrix representing the least significant bit of the Diffie-Hellman function, i.e. the matrix $B_0 \circ D$ with $D(x, y) = g^{xy} \bmod N$ and $B_0(z)$ is the least significant bit (of the binary representation) of z. Here N is a prime and g is an element of maximal order $N - 1$. Shparlinski [56] could show that the discrepancy of $B_0 \circ D$ is upper bounded by $N^{-1/24+o(1)}$. Applying Shaltiel's bound (2.13) we infer that the spectral norm of $B_0 \circ D$ is bounded from above by $N^{71/72+o(1)}$.

We can improve on Shparlinski's result in several ways: By a more refined analysis we can "bypass" bound (2.13) improving the constant from 71/72 to 23/24. Furthermore, we build a whole theory around this result leading to upper bounds of the spectral norm of more general matrices of the form $B \circ H$. These bounds are valid for many (not too much *biased*) Boolean functions B and some functions H (including the Diffie-Hellman function and discrete multiplication). As already pointed out, since we deal with the algebraic object "spectral norm" our techniques have a more algebraic touch than the ones used in [56].

5.2 Considered Functions

In this section we introduce some candidates for the function H that we will consider in the sequel. To this end we present two characterizations of functions, the *multiplicative* and the *exponential structure*. Later, in Section 5.8 we will present some functions H that are actually used in cryptographic protocols and that indeed have either multiplicative or exponential structure. In fact, our applications are not only limited to cryptography. In Section 5.7 we will present some general functions H with multiplicative or exponential structure. Furthermore, we will use the techniques of this chapter to construct matrices with arbitrary spectral norm.

5.2.1 Functions with Multiplicative Structure

Let $M : \mathbb{Z}_N \times \mathbb{Z}_N \to \mathbb{Z}_N$ be the function that satisfies

$$M(x, y) = xy$$

(multiplication in $\mathbb{Z}_N$). We want to exploit the fact that some (cryptographic) functions are close relatives of M. The general concept of a "multiplicative structure" is captured in the following definition:

Definition 5.2.1. *Let $\mathcal{X}$ and $\mathcal{Y}$ be two subsets of $\{0, \ldots, N-1\}$. We say $H : \{0, \ldots, N-1\} \times \{0, \ldots, N-1\} \to \mathbb{Z}_N$ has a* multiplicative structure on subdomain $\mathcal{X} \times \mathcal{Y} \subseteq \{0, \ldots, N-1\} \times \{0, \ldots, N-1\}$ *if there exist injective functions $\sigma : \mathcal{X} \to \mathbb{Z}_N$ and $\tau : \mathcal{Y} \to \mathbb{Z}_N$ such that*

$$H(x, y) = M_{\sigma,\tau}(x, y) := \sigma(x) \cdot \tau(y)$$

for all $x \in \mathcal{X}$ and $y \in \mathcal{Y}$, where $M_{\sigma,\tau}(x, y)$ may take arbitrary values outside $\mathcal{X} \times \mathcal{Y}$.

Examples of cryptographic and non-cryptographic functions H with multiplicative structure are given in Sections 5.7 and 5.8.

5.2.2 Functions with Exponential Structure

For an element g of order T in $\mathbb{Z}_N^*$, let $G \subseteq \mathbb{Z}_N^*$ be the subgroup generated by g. For an element $x \in G$ we denote by $\mathrm{idx}_g x$ the discrete logarithm of x to the basis g, i.e. the (unique) element $a \in \mathbb{Z}_T$ that satisfies $g^a = x$. If T contains at least one large prime factor, then computing the discrete logarithm in G is assumed to be computationally hard.

The *Diffie-Hellman function* [9] $D : G \times G \to G$ with respect to g is given by the following mapping:

$$D(x, y) = g^{\mathrm{idx}_g x \cdot \mathrm{idx}_g y} \quad .$$

The security of numerous cryptographic protocols relies on the computational hardness of computing this function. The most famous of those protocols is the Diffie-Hellman key exchange protocol (see [9] or, e. g. [40, Chapters 3.7 and 12.6]) itself: Alice and Bob respectively pick random $a, b \in \mathbb{Z}_T$ and exchange $x = g^a, y = g^b$. The secret key, $g^{ab} = g^{\mathrm{idx}_g x \cdot \mathrm{idx}_g y}$, can be computed by both parties. In order to totally break the protocol, a passive eavesdropper, Eve, must compute the Diffie-Hellman function. Again, if T contains at least one large prime factor, then computing the Diffie-Hellman function in G is assumed to be computationally hard (*computational Diffie-Hellman assumption*). As an example of a cryptosystem whose security relies on the Diffie-Hellman function we can mention the El Gamal Cryptosystem [15] (see also Example 5.8.2).
Again we want to exploit the fact that some functions are close relatives to D. We define the general concept of an "exponential structure" as follows:

Definition 5.2.2. *Let $g \in \mathbb{Z}_N^*$ be an element of order T, We say $H : \{0, \ldots, T-1\} \times \{0, \ldots, T-1\} \to \mathbb{Z}_N$ has an* exponential structure on subdomain $\mathcal{X} \times \mathcal{Y} \subseteq \{0, \ldots, T-1\} \times \{0, \ldots, T-1\}$ with respect to g *if there exist injective functions $\sigma : \mathcal{X} \to \mathbb{Z}_T$ and $\tau : \mathcal{Y} \to \mathbb{Z}_T$ such that*

$$H(x, y) = D_{\sigma,\tau}(x, y) := g^{\sigma(x)\tau(y)},$$

for all $x \in \mathcal{X}$ and $y \in \mathcal{Y}$, where $D_{\sigma,\tau}(x, y)$ may take arbitrary values outside $\mathcal{X} \times \mathcal{Y}$.

Examples of functions H with exponential structure are given in Sections 5.7 and 5.8.

5.3 Boolean Predicates

In this section we will introduce some natural Boolean functions.

We will be mostly interested in the following functions $B : \mathbb{Z}_N \to \{-1, 1\}$:

- the k-th least significant bit ($0 \leq k \leq n-1$),
- predicates $B(z)$ that evaluate to 1 iff z is chosen from a (fixed) interval,
- predicates $B(z)$ that evaluate to 1 iff z is chosen from the union of distinct intervals (*strongly semilinear functions*),

Depending on the Boolean predicate B, we will introduce two terms, $EXP_B(N, d)$ and $\text{bias}_B(N)$, that will render themselves useful in our analysis. We refer to the terms $\text{bias}_B(N)$ and $EXP_B(N, d)$ as the *balance terms associated with* B. In particular, we are interested in non-trivial upper bounds on the values $EXP_B(N, d)$ and $\text{bias}_B(N)$. As we will see later, "good bounds" on $EXP_B(N, d)$ and $\text{bias}_B(N)$ will be a necessary condition for our techniques to obtain good bounds on the spectral norm of $B \circ H$.

For a divisor $d < N$ of N, define

$$EXP_B(N, d) := \frac{d}{N} \sum_{a \in \mathbb{Z}^*_{N/d}} |\langle B, E_{ad} \rangle| = \frac{d}{N} \sum_{a \in \mathbb{Z}^*_{N/d}} \left| \sum_{x \in \mathbb{Z}_N} B(x) E_{ad}(x) \right| \ . \tag{5.1}$$

We want to stress that $\langle B, E_{ad} \rangle = \sum_{x \in \mathbb{Z}_N} B(x) E_{ad}(x)$ denotes the inner product of B and E_{ad} in $\mathbb{C}$, i.e. we view the functions $B : \mathbb{Z}_N \to \{-1, 1\}$ and $E_{ad} : \mathbb{Z}_N \to \mathbb{C}$ as vectors with complex entries. By equation (2.18), $EXP_B(N, d)$ is trivially upper bounded by N. We call B *smooth* if it satisfies

$$EXP_B(N, d) = N^{o(1)} \ ,$$

for all divisors $d < N$ of N.[1]

The *bias* of a binary function $B : \mathbb{Z}_N \to \{-1, 1\}$ is given by

$$\text{bias}_B(N) := \left| \sum_{z \in \mathbb{Z}_N} B(z) \right| \in \{0, \ldots, N\} \ . \tag{5.2}$$

[1] Formally correct $B = B(N)$ is a family of Boolean functions indexed by N, where N runs through (a subset of) the natural numbers $\mathbb{N}$.

The bias measures the degree of "balance" between negative and positive entries in vector B (where value 0 reflects perfect balance). Function $B : \mathbb{Z}_N \to \{-1, 1\}$ is called *unbiased* if

$$\mathrm{bias}_B(N) = N^{o(1)} \ .$$

Functions that are at the same time unbiased and smooth will play a central role in our results.

Definition 5.3.1. *We call a binary function $B : \mathbb{Z}_N \to \{-1, 1\}$* balanced *if it is unbiased (i.e., $\mathit{bias}_B(N) = N^{o(1)}$) and smooth (i.e., $\mathit{EXP}_B(N, d) = N^{o(1)}$).*

5.4 The Bound on the Spectral Norm

We are now ready to present the main results of this chapter.

Theorem 5.4.1. *Let B be a balanced binary function. Let H be a function with a multiplicative structure on subdomain $\mathcal{X} \times \mathcal{Y}$ of $\{0, \ldots, N-1\} \times \{0, \ldots, N-1\}$. Assume that $N - |\mathcal{X}| = O(N^\alpha)$ and $N - |\mathcal{Y}| = O(N^\alpha)$ for some fixed constant $\alpha \geq 0$. Then, the spectral norm of matrix $B \circ H$ is bounded from above by*

$$\|B \circ H\|_2 = \begin{cases} N^{1/2+o(1)} & : \ \alpha = 0 \\ O(N^{(1+\alpha)/2}) & : \ \alpha \neq 0. \end{cases}$$

Theorem 5.4.2. *For integers N, let T be the order of an element $g \in \mathbb{Z}_N^*$. Let B be a balanced binary function. Let $H : \{0, \ldots, T-1\} \times \{0, \ldots, T-1\} \to \mathbb{Z}_N$ be a function with an exponential structure on subdomain $\mathcal{X} \times \mathcal{Y}$ of $\{0, \ldots, T-1\} \times \{0, \ldots, T-1\}$ with respect to g. Assume that $T - |\mathcal{X}| = O(T^{5/16} N^{5/8})$ and $T - |\mathcal{Y}| = O(T^{5/16} N^{5/8})$. Then, the spectral norm of matrix $B \circ H$ is bounded from above by*

$$\|B \circ H\|_2 = T^{21/32} N^{5/16+o(1)} \ .$$

Note that $B \circ H$ is a $T \times T$ matrix and hence the theorem gives a nontrivial bound on $\|B \circ H\|_2$ (a bound of order $T^{1-\varepsilon}$, $\varepsilon > 0$) for elements g of order $T > N^{10/11+\delta}$ for any positive constant δ.

Theorem 5.4.3. *For primes N let T be the order of an element $g \in \mathbb{Z}_N^*$. Let B be a balanced binary function. Let $H : \{0, \ldots, T-1\} \times \{0, \ldots, T-1\} \to \mathbb{Z}_N$ be a function with an exponential structure on subdomain $\mathcal{X} \times \mathcal{Y}$ of $\{0, \ldots, T-1\} \times \{0, \ldots, T-1\}$ with respect to g. Assume that $T - |\mathcal{X}| = O(T^{2/3} N^{1/4})$ and $T - |\mathcal{Y}| = O(T^{2/3} N^{1/4})$. Then, the spectral norm of matrix $B \circ H$ is bounded from above by*

$$\|B \circ H\|_2 = T^{5/6} N^{1/8+o(1)} \ .$$

This gives a nontrivial bound on $\|B \circ H\|_2$ for elements g of order $T > N^{3/4+\delta}$ for any positive constant δ.

The proof of the last three Theorems is given in Section 5.5. The next theorem (whose proof will be given in Section 5.6) shows that many natural binary functions are balanced.

Theorem 5.4.4. *The following binary functions $B : \mathbb{Z}_N \to \{-1, 1\}$ are balanced:*

- *Let B_k denote the k-th least significant bit of (the binary representation) of $z \in \mathbb{Z}_N$. Assume that k is either constant or (as a function in N) satisfies $k(N) = o(\log N)$.*[2] *Then B_k is balanced.*
- *Let B denote any unbiased and strongly semilinear function of length k (see Definition 5.6.1). Assume that k is either constant or (as a function in N) satisfies $k(N) = N^{o(1)}$. Then B is balanced.*

One may combine now any of these balanced binary functions with Theorem 5.4.1, Theorem 5.4.2, or Theorem 5.4.3 to get the result that the spectral norm of *concrete binary functions* is at most $T^{1-\varepsilon}$ for a constant $\varepsilon > 0$. As mentioned before, we may now apply the lower bound machinery from Chapter 4 to get complexity theoretic lower bounds. We will dwell on that later in Section 5.10.

5.5 The Proof

This section is devoted to the proofs of the Theorems 5.4.1, 5.4.2, and 5.4.3. The roadmap is as follows. First, in Subsection 5.5.1, we present a general "Splitting Lemma" that breaks the spectral norm of $B \circ H$ into two components. The first component will only depend on the balance terms associated with B. Providing upper bounds on the balance terms for concrete Boolean functions B will be subject of Section 5.6. The second component will only depend on the matrix H and spectral norms of complex *auxiliary matrices* $E_a \circ H$ will appear (thereby putting *exponential sums* into play). In Subsection 5.5.2 we will upper bound the spectral norm of these auxiliary matrices for concrete functions H. Finally, after putting the obtained results together in Subsection 5.5.3, we deal with the concept of matrices H attaining arbitrary values (error terms) outside a subdomain. As we will show in Subsection 5.5.4 this does not affect our results as far as the subdomain is large enough.

5.5.1 The Splitting Lemma

For $T \leq N$, let $H : \mathbb{Z}_T \times \mathbb{Z}_T \to \mathbb{Z}_N$ be a matrix and let $B : \mathbb{Z}_N \to \{-1, 1\}$ be a binary function. In this subsection we present the central idea of bounding the spectral norm of matrices of the form $B \circ H$. We will split it into two independent components, one only depending on the binary function B, the other only depending on the matrix H.

The main result in this subsection reads as follows:

[2] This rules out the possibility to select bits of high significance (not to speak of the most significant bit).

Lemma 5.5.1.

$$\|B \circ H\|_2 \leq bias_B(N) + \frac{1}{N} \cdot \sum_{a=1}^{N-1} |\langle B, E_{-a}\rangle| \, \|E_a \circ H\|_2 \ .$$

Proof. Choose $u, v \in \mathbb{R}^T$ such that $\|u\|_2 = \|v\|_2 = 1$ and $\|B \circ H\|_2 = |u^\top (B \circ H)v|$. We obtain

$$\begin{aligned}
|u^\top (B \circ H)v| &= \left| \sum_{x \in \mathbb{Z}_T} \sum_{y \in \mathbb{Z}_T} u_x v_y B(H(x,y)) \right| \\
&\overset{(2.26)}{=} \frac{1}{N} \left| \sum_{x \in \mathbb{Z}_T} \sum_{y \in \mathbb{Z}_T} u_x v_y \sum_{z \in \mathbb{Z}_T} \sum_{a \in \mathbb{Z}_N} B(z) E_a(H(x,y) - z) \right| \\
&\overset{(2.21)}{=} \frac{1}{N} \left| \sum_{a \in \mathbb{Z}_N} \sum_{z \in \mathbb{Z}_T} B(z) E_{-a}(z) \sum_{x \in \mathbb{Z}_T} \sum_{y \in \mathbb{Z}_T} u_x v_y E_a(H(x,y)) \right| \\
&\leq \frac{1}{N} \sum_{a \in \mathbb{Z}_N} S(a) \ ,
\end{aligned}$$

where $S(a)$ is given by

$$\begin{aligned}
S(a) &= \left| \sum_{z \in \mathbb{Z}_N} B(z) E_{-a}(z) \sum_{x \in \mathbb{Z}_T} \sum_{y \in \mathbb{Z}_T} u_x v_y E_a(H(x,y)) \right| \\
&\leq |\langle B, E_{-a}\rangle| \cdot |u^\top (E_a \circ H)v| \\
&\leq |\langle B, E_{-a}\rangle| \cdot \|E_a \circ H\|_2 \ .
\end{aligned}$$

The proof is completed by discussing the case $a = 0$ separately. Since $E_0(z) = 1$ for each $z \in \mathbb{Z}_N$ and $T \leq N$ we get

$$\begin{aligned}
\frac{1}{N} S(0) &= \frac{1}{N} \left| \sum_{z \in \mathbb{Z}_N} B(z) \sum_{x \in \mathbb{Z}_T} \sum_{y \in \mathbb{Z}_T} u_x v_y \right| \\
&\overset{(2.7)}{\leq} \frac{T}{N} \left| \sum_{z \in \mathbb{Z}_N} B(z) \right| \leq \left| \sum_{z \in \mathbb{Z}_N} B(z) \right| \\
&= \mathrm{bias}_B(N) \ .
\end{aligned}$$

□

The Splitting Lemma 5.5.1 shows that for computing an upper bound on the spectral norm $\|B \circ H\|_2$ it is sufficient to give an upper bound on $\langle B, E_{-a}\rangle$ and on the spectral norm $\|E_a \circ H\|_2$. Note that the matrix $E_a \circ H$ is a complex matrix from $\mathbb{C}^{T \times T}$.

We would like to have a bound on $\|E_a \circ H\|_2$ for a concrete function H that holds for all possible values $a \in \mathbb{Z}_N \setminus \{0\}$. Unfortunately we are not able to do so and, as it will turn out, our bound on $\|E_a \circ H\|_2$ will also depend on the value $d = \gcd(a, N)$. To this end we cut the set $\mathbb{Z}_N$ into $\tau(N)$ different slices, where $\tau(N)$ is the number of divisors of N. The d-th slice contains all $a \in \mathbb{Z}_N$ such that $\gcd(a, N) = d$. These are precisely the elements of the form ad for $a \in \mathbb{Z}_N^*$. This leads to the following

Corollary 5.5.2. *For smooth binary functions $B : \mathbb{Z}_N \to \{-1, 1\}$, we have*

$$\begin{aligned}
\|B \circ H\|_2 &\leq bias_B(N) + \frac{1}{N} \sum_{d:d|N,d<N} \sum_{a \in \mathbb{Z}^*_{N/d}} |\langle B, E_{-ad}\rangle| \, \|E_{ad} \circ H\|_2 \\
&\leq bias_B(N) + \frac{1}{N} \sum_{d:d|N,d<N} \max_{a \in \mathbb{Z}^*_{N/d}} \|E_{ad} \circ H\|_2 \sum_{a \in \mathbb{Z}^*_{N/d}} |\langle B, E_{ad}\rangle| \\
&\leq bias_B(N) + N^{o(1)} \sum_{d:d|N,d<N} \max_{a \in \mathbb{Z}^*_{N/d}} \|E_{ad} \circ H\|_2 / d \ .
\end{aligned}$$

We briefly note the well-known fact [49] that

$$\tau(N) = 2^{(1+o(1))\frac{\ln N}{\ln \ln N}} = N^{o(1)} \ . \tag{5.3}$$

Corollary 5.5.3. *Assume that*

$$\max_{d:d|N,d<N} \frac{1}{d} \max_{a \in \mathbb{Z}^*_{N/d}} \|E_{ad} \circ H\|_2 \quad \leq \quad U_H(N) \ .$$

Then, for smooth binary functions $B : \mathbb{Z}_N \to \{-1, 1\}$, we have:

$$\begin{aligned}
\|B \circ H\|_2 &= bias_B(N) + N^{o(1)} \cdot \sum_{d:d|N,d<N} \max_{a \in \mathbb{Z}^*_{N/d}} \|E_{ad} \circ H\|_2 / d \\
&\overset{(5.3)}{\leq} bias_B(N) + N^{o(1)} \cdot U_H(N) \ .
\end{aligned}$$

5.5.2 The Spectral Norm of the Auxiliary Matrices

In this section we will give a bound on the spectral norm of the complex auxiliary matrices $E_a \circ H$. For all matrices $H : \mathbb{Z}_T \times \mathbb{Z}_T \to \mathbb{Z}_N$ we have by equation (2.6):

$$T^{1/2} \leq \|E_a \circ H\|_2 \leq T \ . \tag{5.4}$$

The entries of matrix $(AA^*)_{x,y}$ with $A = E_a \circ H$ are given by

$$\begin{aligned}
(AA^*)_{x,y} &= \sum_{z \in \mathbb{Z}_T} E_a(H(x,z)) \overline{E_a(H(y,z))} \\
&\overset{(2.22)}{=} \sum_{z \in \mathbb{Z}_T} E_a(H(x,z)) E_{-a}(H(y,z)) \\
&\overset{(2.21)}{=} \sum_{z \in \mathbb{Z}_T} E_a(H(x,z) - H(y,z)) \ .
\end{aligned}$$

Using Lemma 2.1.3 this implies the following general bound on $\|E_a \circ H\|_2$:

Lemma 5.5.4. *Let $H : \mathbb{Z}_T \times \mathbb{Z}_T \to \mathbb{Z}_N$ be a matrix. For all $0 \neq a \in \mathbb{Z}_N$, the following holds:*

$$\begin{aligned}\|E_a \circ H\|_2^2 &\leq \|(E_a \circ H)(E_a \circ H)^*\|_\infty \\ &= \max_{x \in \mathbb{Z}_T} \sum_{y \in \mathbb{Z}_T} \left| \sum_{z \in \mathbb{Z}_T} E_a(H(x,z) - H(y,z)) \right| .\end{aligned}$$

However, for concrete functions H, it sometimes seems to be hard to bound the sum over y for values $x \in \mathbb{Z}_T$ with $\gcd(x,T) \neq 1$. To circumvent this we present the following

Lemma 5.5.5. *Let $H : \mathbb{Z}_T \times \mathbb{Z}_T \to \mathbb{Z}_N$ be a matrix. For all $0 \neq a \in \mathbb{Z}_N$, the following holds:*

$$\|E_a \circ H\|_2^2 = T^{o(1)} \cdot \max_{d|T, d<T} \max_{x \in \mathbb{Z}^*_{T/d}} \sum_{y \in \mathbb{Z}^*_{T/d}} \left| \sum_{z \in \mathbb{Z}_T} E_a(H(dx,z) - H(dy,z)) \right| .$$

Proof. The set $\mathbb{Z}_T$ can be cut into $\tau(T)$ different slices, where $\tau(T)$ is the number of divisors of T. The d-th slice contains all integers $x \in \mathbb{Z}_T$ such that $\gcd(x,T) = d$. These are precisely the elements of the form xd for $x \in \mathbb{Z}^*_{T/d}$. For the matrix $A = E_a \circ H$, we define the d-slice matrix $A_d \in \mathbb{C}^{\varphi(T/d) \times T}$ as

$$A_d(x,y) = A(dx,y) \ ,$$

for all $x \in \mathbb{Z}^*_{T/d}$, $y \in \mathbb{Z}_T$. For notational simplicity we let the x-indices of the matrix A_d range over all $x \in \mathbb{Z}^*_{T/d}$.

Since the operator norm is invariant under adding all-zero columns and vectors and under permutations of columns and rows, there are matrices A'_d with $\|A_d\|_2 = \|A'_d\|_2$ and $A = \sum_{d|T} A'_d$. Hence we can bound

$$\begin{aligned}\|A\|_2 &= \|\sum_{d:d|T,d<T} A'_d\|_2 \leq \sum_{d:d|T,d<T} \|A'_d\|_2 \\ &= \sum_{d:d|T,d<T} \|A_d\|_2 \\ &\overset{(5.3)}{=} T^{o(1)} \max_{d:d|T,d<T} \|A_d\|_2 \ .\end{aligned}$$

For $x, y \in \mathbb{Z}^*_{T/d}$, the entry $(A_d A_d^*)_{x,y}$ is given by

$$\sum_{z \in \mathbb{Z}_T} E_a(H(dx,z) - H(dy,z)) \ .$$

Now the assertion follows by Lemma 2.1.3. ∎

The multiplicative case

In this subsection, we will bound $\|E_a \circ M\|_2$ for the function $M(x,y) = xy$ and the case $T = N$. We will explain later how results dealing with M (multiplication in $\mathbb{Z}_N$) can be extended to arbitrary functions with a multiplicative structure. The following result (that can be found in [22], Problem 23 of Chapter 5.6) is well-known and can be proved by means of Lemma 5.5.4.

Lemma 5.5.6. *For all $a \in \mathbb{Z}_N^*$, the following holds: $\|E_a \circ M\|_2 = N^{1/2}$.*

We present a simple generalization of this Lemma to cover the case of general $a \in \mathbb{Z}_N$.

Lemma 5.5.7. *For all $0 \neq a \in \mathbb{Z}_N$, we have $\|E_a \circ M\|_2 \leq (dN)^{1/2}$, where $d = \gcd(a, N)$.*

Proof. We write a as $d \cdot a_0$ with $\gcd(a_0, N) = 1$. For every $x, y \in \mathbb{Z}_N$,

$$\left|\sum_{z\in\mathbb{Z}_N} E_a(xz - yz)\right| = \left|\sum_{z\in\mathbb{Z}_N} E_{d\cdot a_0}(z(x-y))\right| = \left|d \sum_{z\in\mathbb{Z}_{N/d}} E_{a_0}^{N/d}(z(x-y))\right|,$$

which evaluates by (2.25) to N for $x = y \pmod{N/d}$ and to 0 otherwise. By Lemma 5.5.4 we continue with

$$\begin{aligned} \|E_a \circ M\|_2^2 &\leq \max_{x\in\mathbb{Z}_N} \sum_{y\in\mathbb{Z}_N} \left|\sum_{z\in\mathbb{Z}_N} E_a(xz - zy))\right| \\ &= dN \ . \end{aligned}$$

The last equality follows since for every fixed $x \in \mathbb{Z}_N$, $x = y \pmod{N/d}$ holds for exactly d distinct $y \in \mathbb{Z}_N$. $\square$

Hence, for all $a \in \mathbb{Z}_{N/d}^*$, $\|E_{ad} \circ M\|_2 \leq (dN)^{1/2}$ what leads to the following corollary:

Corollary 5.5.8.

$$\max_{d:d|N,d<N} \frac{1}{d} \max_{a\in\mathbb{Z}_{N/d}^*} \|E_{ad} \circ M\|_2 \leq \max_{d:d|N,d<N} (N/d)^{1/2} \leq N^{1/2} \ .$$

The Exponential Case

Let g be an element from $\mathbb{Z}_N^*$ of order T. In this subsection, we will bound $\|E_{ad} \circ D\|_2$ for the function $D : \mathbb{Z}_T \times \mathbb{Z}_T \to \mathbb{Z}_N, D(x,y) = g^{xy}$. The following two results will be useful:

Lemma 5.5.9 ([13]). *Let $\gamma \in \mathbb{Z}_N^*$ be an element of order T. Let $0 \neq a \in \mathbb{Z}_N$ be an integer with $\gcd(a, N) = d < N$. Then the bound*

$$\sum_{x\in\mathbb{Z}_T} \left|\sum_{y\in\mathbb{Z}_T} E_a(\gamma^y - \gamma^{xy})\right| = d^{1/4} T^{21/16} N^{5/8+o(1)}$$

holds.

We remark that for $d = 1$ this bound is nontrivial for $T \geq N^{10/11+\delta}$, for any positive constant δ. For prime modulus the result can be improved:

Lemma 5.5.10 ([5]). *For a prime N, let $\gamma \in \mathbb{Z}_N^*$ be an element of order T. Let $a \in \mathbb{Z}_N^*$. Then the bound*

$$\sum_{x \in \mathbb{Z}_T} \left| \sum_{y \in \mathbb{Z}_T} E_a(\gamma^y - \gamma^{xy}) \right| = O(T^{5/3} N^{1/4}) \ .$$

holds.

Again we remark that this bound is nontrivial for $T \geq N^{9/20}$.

Lemma 5.5.11. *For all $0 \neq a \in \mathbb{Z}_N$ the following holds:*

$$\|E_a \circ D\|_2 = T^{21/32} N^{5/16+o(1)} d^{1/8} \ ,$$

where $d = \gcd(a, N) < N$ and g is an element of order T. Moreover, for prime N, we have

$$\|E_a \circ D\|_2 = T^{5/6+o(1)} N^{1/8} \ .$$

Proof. By Lemma 5.5.5 we have

$$\|E_a \circ D\|_2^2 = T^{o(1)} \cdot \max_{e|T, e<T} \max_{x \in \mathbb{Z}_{T/e}^*} \sigma_x(e) \ ,$$

where the e-slice $\sigma_x(e)$ is given by

$$\sigma_x(e) = \sum_{y \in \mathbb{Z}_{T/e}^*} \left| \sum_{z \in \mathbb{Z}_T} E_a(D(ex, z) - D(ey, z)) \right| \ .$$

We now present a bound on $\sigma_x(e)$ which is independent from x and e. Fix a divisor $e < T$ of T and $x \in \mathbb{Z}_{T/e}^*$.

$$\sigma_x(e) \quad = \quad \sum_{y \in \mathbb{Z}_{T/e}^*} \left| \sum_{z \in \mathbb{Z}_T} E_a(g^{exz} - g^{eyz}) \right|$$

Since g is an element of order T, the element $\gamma = g^e$ is of order T/e. We substitute g^e by γ:

$$\sigma_x(e) \quad \leq \quad \sum_{y \in \mathbb{Z}_{T/e}^*} \left| \sum_{z \in \mathbb{Z}_T} E_a(\gamma^{xz} - \gamma^{yz}) \right|$$

For every $x \in \mathbb{Z}_{T/e}^*$, we set $\gamma_x = \gamma^x$ which is an element of order T/e. After replacing the variable y by y/x we proceed with

$$\begin{aligned}
\sigma_x(e) &\leq \sum_{y \in \mathbb{Z}_{T/e}^*} \left| \sum_{z \in \mathbb{Z}_T} E_a(\gamma_x^z - \gamma_x^{yz}) \right| \\
&= e \sum_{y \in \mathbb{Z}_{T/e}^*} \left| \sum_{z \in \mathbb{Z}_{T/e}} E_a(\gamma_x^z - \gamma_x^{yz}) \right| \\
&\leq e \sum_{y \in \mathbb{Z}_{T/e}} \left| \sum_{z \in \mathbb{Z}_{T/e}} E_a(\gamma_x^z - \gamma_x^{yz}) \right| \quad .
\end{aligned}$$

For arbitrary (non-prime) N we can bound this sum using Lemma 5.5.9:

$$\sigma_x(e) = e(T/e)^{21/16} N^{5/8+o(1)} d^{1/4} \quad .$$

Using the trivial bound $e \geq 1$ this leads for any $e|T$ to

$$\sigma_x(e) = T^{21/16} N^{5/8+o(1)} d^{1/4} \quad .$$

This shows the bound $\|E_a \circ D\|_2 = T^{21/32} N^{5/16+o(1)} d^{1/8}$.

For prime N we can use the stronger bound from Lemma 5.5.10:

$$\sigma_x(e) = O(T^{5/3} N^{1/4}) \quad .$$

□

Corollary 5.5.12.

$$\begin{aligned}
\max_{d:d|N, d<N} \frac{1}{d} \max_{a \in \mathbb{Z}_{N/d}^*} \|E_{ad} \circ D\|_2 &= \max_{d:d|N,d<N} N^{5/16} T^{21/32+o(1)} d^{-7/8} \\
&= N^{5/16} T^{21/32+o(1)} \quad .
\end{aligned}$$

Moreover, for prime N we have

$$\max_{d:d|N,d<N} \frac{1}{d} \max_{a \in \mathbb{Z}_{N/d}^*} \|E_{ad} \circ D\|_2 = N^{1/8} T^{5/6+o(1)} \quad .$$

5.5.3 Putting all together

Since the spectral norm is invariant under permutations of rows and columns we have that for all bijective functions $\tilde{\sigma}, \tilde{\tau} : \{0, \ldots, T-1\} \rightarrow \mathbb{Z}_T$, and for $M_{\tilde{\sigma},\tilde{\tau}}(x,y) = M(\tilde{\sigma}(x), \tilde{\tau}(y))$,

$$\|B \circ M_{\tilde{\sigma},\tilde{\tau}}\|_2 = \|B \circ M\|_2 \quad . \tag{5.5}$$

For the **multiplicative case** we may now apply the bound $U_M(N) \leq N^{1/2}$ (according to Corollary 5.5.8) to Corollary 5.5.3:

Corollary 5.5.13. *For smooth binary functions $B : \mathbb{Z}_N \to \{-1,1\}$ and for bijective mappings $\tilde{\sigma}, \tilde{\tau} : \{0,\ldots,N-1\} \to \mathbb{Z}_N$, the following holds:*

$$\begin{aligned} \|B \circ M_{\tilde{\sigma},\tilde{\tau}}\|_2 &\overset{(5.5)}{=} \|B \circ M\|_2 \\ &\leq bias_B(N) + N^{1/2+o(1)} \ . \end{aligned}$$

In the **Exponential case** we may apply Corollary 5.5.3 with $U_D(N) = T^{21/32}N^{5/16+o(1)}$ for general N or $U_D(N) = T^{5/6+o(1)}N^{1/8}$ for prime N, respectively (according to Corollary 5.5.12):

Corollary 5.5.14. *For smooth binary functions $B : \mathbb{Z}_N \to \{-1,1\}$ and for bijective mappings $\tilde{\sigma}, \tilde{\tau} : \{0,\ldots,T-1\} \to \mathbb{Z}_T$, the following holds:*

$$\|B \circ D_{\tilde{\sigma},\tilde{\tau}}\|_2 \leq bias_B(N) + N^{5/16+o(1)}T^{21/32} \ .$$

Moreover, if N is a prime, then

$$\|B \circ D_{\tilde{\sigma},\tilde{\tau}}\|_2 \leq bias_B(N) + N^{1/8+o(1)}T^{5/6} \ .$$

5.5.4 Multiplicative and Exponential Structure on Subdomains

A function H with a multiplicative structure coincides with $M_{\sigma,\tau}$ on a subdomain $\mathcal{X} \times \mathcal{Y}$. The "non-multiplicative behavior" outside the subdomain can however be controlled as follows:

Lemma 5.5.15. *If matrices $A, B : \{0,\ldots,T-1\} \times \{0,\ldots,T-1\} \to \{-1,1\}$ coincide on subdomain $\mathcal{X} \times \mathcal{Y}$, then*

$$|\|A\|_2 - \|B\|_2| \leq 2\sqrt{T^2 - |\mathcal{X}||\mathcal{Y}|} \ .$$

Proof. Since the exception matrix $D := A - B \in \{-2,0,2\}^{T\times T}$ is zero on subdomain $\mathcal{X} \times \mathcal{Y}$, it has at most $T^2 - |\mathcal{X}||\mathcal{Y}|$ nonzero entries of absolute value 2. By the properties of the spectral norm we have

$$\|A\|_2 = \|B + D\|_2 \leq \|B\|_2 + \|D\|_2 \ ,$$

where $\|D\|_2$ can be bounded by its Frobenius norm (2.6):

$$\|D\|_2 \leq \|D\|_F \leq 2\sqrt{T^2 - |\mathcal{X}||\mathcal{Y}|} \ .$$

The same way we can show that

$$\|B\|_2 \leq \|A\|_2 + 2\sqrt{T^2 - |\mathcal{X}||\mathcal{Y}|} \ ,$$

and the result follows. □

Now we are ready to prove Theorem 5.4.1.

Proof (of Theorem 5.4.1). Let B be a balanced binary function. Let H be a function with multiplicative structure on subdomain $\mathcal{X} \times \mathcal{Y}$ of $\{0, \ldots, N-1\} \times \{0, \ldots, N-1\}$. By Definition 5.2.1 there exist injective functions $\sigma : \mathcal{X} \to \mathbb{Z}_N$ and $\tau : \mathcal{Y} \to \mathbb{Z}_N$, such that H coincides with $M_{\sigma,\tau}$ on subdomain $\mathcal{X} \times \mathcal{Y}$. Since the mapping $\sigma : \mathcal{X} \to \mathbb{Z}_N$ is injective, it can be extended to a bijective mapping from $\{0, \ldots, N-1\}$ to $\mathbb{Z}_N$ without changing its behavior on $\mathcal{X}$. The same remark holds for τ. Let $\tilde{\sigma}, \tilde{\tau} : \{0, \ldots, N-1\} \to \mathbb{Z}_N$ be the resulting bijective functions. Then $B \circ M_{\sigma,\tau}$ coincides with $B \circ M_{\tilde{\sigma},\tilde{\tau}}$ on subdomain $\mathcal{X} \times \mathcal{Y}$. The part outside of the subdomain can be controlled by Lemma 5.5.15:

$$\|B \circ H\|_2 \leq \|B \circ M_{\tilde{\sigma},\tilde{\tau}}\|_2 + 2\sqrt{N^2 - |\mathcal{X}||\mathcal{Y}|},$$

what can be bounded by Corollary 5.5.13 and by the fact that B is balanced by

$$\begin{aligned} &\leq \text{bias}_B(N) + N^{1/2+o(1)} + 2\sqrt{N^2 - |\mathcal{X}||\mathcal{Y}|} \\ &= N^{1/2+o(1)} + 2\sqrt{N^2 - |\mathcal{X}||\mathcal{Y}|}. \end{aligned}$$

Since

$$\begin{aligned} N^2 - |\mathcal{X}||\mathcal{Y}| &= N(N - |\mathcal{X}|) + N(N - |\mathcal{Y}|) - (N - |\mathcal{X}|)(N - |\mathcal{Y}|) \\ &\leq N(N - |\mathcal{X}|) + N(N - |\mathcal{Y}|) \end{aligned}$$

and $N - |\mathcal{X}| = O(N^\alpha)$ and $N - |\mathcal{Y}| = O(N^\alpha)$ for some fixed constant $0 \leq \alpha < 1$ we get

$$\|B \circ H\|_2 = N^{1/2+o(1)} + O(N^{(1+\alpha)/2})$$

which yields the theorem. □

Theorem 5.4.2 and Theorem 5.4.3 follow analogously by combining Corollary 5.5.14 with Lemma 5.5.15.

5.6 Upper Bounds on the Balance Terms

Recall the following definitions of the terms $\text{bias}_B(N)$ and $EXP_B(N, d)$. The bias of a binary function $B : \mathbb{Z}_N \to \{-1, 1\}$ is given by

$$\text{bias}_B(N) := \left| \sum_{z \in \mathbb{Z}_N} B(z) \right| \in \{0, \ldots, N\} \ .$$

Function $B : \mathbb{Z}_N \to \{-1, 1\}$ is called *unbiased* if $\text{bias}_B(N) = N^{o(1)}$. For a divisor $d < N$ of N, $EXP_B(N, d)$ is defined by

$$EXP_B(N, d) := \frac{d}{N} \sum_{a \in \mathbb{Z}^*_{N/d}} |\langle B, E_{ad} \rangle| \, .$$

B is called smooth if $EXP_B(N, d) = N^{o(1)}$ for all divisors $d < N$. A binary function $B : \mathbb{Z}_N \to \{-1, 1\}$ is called balanced if it is unbiased and smooth.

In this section we will provide bounds on the balance terms $\text{bias}_B(N)$ and $EXP_B(N, d)$ for some Boolean functions B thereby proving Theorem 5.4.4 .

5.6.1 Strongly Semilinear Functions

Definition 5.6.1. *Function $B : \mathbb{Z}_N \to \{-1, 1\}$ is called* semilinear of length k *if there exist parameters $M_i, L_i \in \mathbb{Z}_N$ and $K_i \in \mathbb{Z}_N^*$ such that, for all $z \in \mathbb{Z}_N$, condition*

$$B(z) = 1 \Longleftrightarrow z \in \bigcup_{i=1\ldots k} J_i, \quad J_i := \{K_i z + M_i \bmod N |\ 0 \le z \le L_i\},$$

is valid and the sets J_i are pairwise disjoint. We call a vector B strongly semilinear of length k *if B and $-B$ both are semilinear of length k.*

There are many natural strongly semilinear functions, for instance each indicator function of a union of k disjoint intervals in $\mathbb{Z}_N$ is strongly semilinear of length k.

We now bound $EXP_B(N, d)$ from above for each strongly semilinear function B:

Lemma 5.6.2. *Let $B : \mathbb{Z}_N \to \{-1, 1\}$ be strongly semilinear of length k. Then, for every divisor $d < N$ of N, the following holds:*

$$EXP_B(N, d) \le 2k \log(N/d) \ .$$

Proof. Since B is strongly semilinear of length k, we have the following two characterizations of B and $-B$.

$$B(z) = 1 \Leftrightarrow z \in \bigcup_{i=1\ldots k} J_i, \quad J_i = \{K_i z + M_i \bmod N |\ 0 \le z \le L_i\} \tag{5.6}$$

and

$$B(z) = -1 \Leftrightarrow z \in \bigcup_{i=1\ldots k} J_i'(i), \quad J_i' = \{K_i' z + M_i' \bmod N |\ 0 \le z \le L_i'\} \ .$$

Let B_+ be the vector obtained from B by setting $B_+(z) = 1$ if $B(z) = 1$ and $B_+(z) = 0$ if $B(z) = -1$. Let B_- be the vector obtained from B by setting $B_-(z) = 1$ if $B(z) = -1$ and $B_-(z) = 0$ if $B(z) = 1$. Use $B = B_+ - B_-$ to obtain

$$\begin{aligned} EXP_B(N, d) &= \frac{d}{N} \cdot \sum_{a \in \mathbb{Z}_{N/d}^*} |\langle B, E_{ad} \rangle| \\ &\le \frac{d}{N} \cdot \sum_{a \in \mathbb{Z}_{N/d}^*} |\langle B_+, E_{ad} \rangle| + \frac{d}{N} \cdot \sum_{a \in \mathbb{Z}_{N/d}^*} |\langle B_-, E_{ad} \rangle| \ . \end{aligned} \tag{5.7}$$

Using characterization (5.6) and the fact that $E_{ad}(x) = E_a^{N/d}(x)$, we bound the first sum by

$$
\begin{aligned}
\sum_{a\in\mathbb{Z}^*_{N/d}} |\langle B_+, E_{ad}\rangle| &= \sum_{a\in\mathbb{Z}^*_{N/d}} \left|\sum_{i=1}^{k}\sum_{z=0}^{L_i} E_a^{N/d}(K_i z + M_i)\right| \\
&\overset{(2.20)}{=} \sum_{a\in\mathbb{Z}^*_{N/d}} \left|\sum_{i=1}^{k}\sum_{z=0}^{L_i} E_a^{N/d}(M_i)E_a^{N/d}(K_i z)\right| \\
&\leq \sum_{a\in\mathbb{Z}^*_{N/d}} \sum_{i=1}^{k} \left|E_a^{N/d}(M_i)\right| \left|\sum_{z=0}^{L_i} E_a^{N/d}(K_i z)\right| \\
&\overset{(2.18)}{=} \sum_{i=1}^{k}\sum_{a\in\mathbb{Z}^*_{N/d}} \left|\sum_{z=0}^{L_i} E_a^{N/d}(K_i z)\right|.
\end{aligned}
$$

Since $K_i \in \mathbb{Z}^*_{N/d}$ (because $K_i \in \mathbb{Z}^*_N$), we can substitute a/K_i for a for every $1 \leq i \leq k$. After rearranging the sum we obtain

$$
\begin{aligned}
&\leq \sum_{i=1}^{k}\sum_{a\in\mathbb{Z}^*_{N/d}} \left|\sum_{z=0}^{L_i} E_a^{N/d}(z)\right| \\
&\leq \sum_{i=1}^{k}\sum_{a=1}^{N/d-1} \left|\sum_{z=0}^{L_i} E_a^{N/d}(z)\right| \\
&\overset{(2.28)}{\leq} k\frac{N}{d}\log(N/d).
\end{aligned}
$$

Since the second sum in (5.7) can be bounded analogously the proof is complete. □

5.6.2 The Bits

We start by introducing some notation. Let $B : \mathbb{Z}_N \to \{-1, 1\}$ be a binary function. The set of integers $x \in \mathbb{Z}_N$ where B evaluates to $+1$ is denoted as B_+:

$$B_+ = \{x \in \mathbb{Z}_N : B(x) = 1\}.$$

Similarly, we define the set $B_- := \mathbb{Z}_N \setminus B_+ = \{x \in \mathbb{Z}_N : B(x) = -1\}$ as the set of integers where B evaluates to -1.

Let $B_k(x)$ denote the k-th least significant bit of (the binary representation of) $x \in \mathbb{Z}_N$. We will now give a full description of the function B_k in terms of the sets B_+ and B_-.

We start with the case of the *least significant bit* B_0. For the predicate $B(x) = B_0(x)$ it obviously holds that $B_0(x) = 1$ iff $x = 2z+1$ for $0 \leq z \leq \lfloor N/2 \rfloor$.

Thus, the set $B_+(B_0)$ can be described by

$$B_+(B_0) = \{2z+1\}_{0 \le z \le \lfloor N/2 \rfloor}.$$

For the inner bits this is getting a little bit more sophisticated. Put $r_2 = N \bmod 4$. For the *second least significant bit* $B(x) = B_1(x)$ we get:

$$B_-(B_1) = \{4z+t\}_{0 \le t \le 1, 0 \le z \le b_{r_2}(t)}$$

where $T = \lfloor (N-1)/4 \rfloor$ and $(b_0(0), b_0(1)) = (T-1, T-1)$ for $N = 0 \bmod 4$, $(b_1(0), b_1(1)) = (T, T-1)$ for $N = 1 \bmod 4$, $(b_2(0), b_2(1)) = (T, T)$ for $N = 2 \bmod 4$ and $(b_3(0), b_3(1)) = (T, T)$ for $N = 3 \bmod 4$.

Generalizing this characterization to the *inner bits* B_k yields for any integer N with $r_k = N \bmod 2^{k+1}$:

$$B_-(B_k) = \{2^{k+1}z + t\}_{0 \le t \le 2^k - 1, 0 \le z \le b_{r_k}(t)}. \tag{5.8}$$

Clearly, $b_{r_k}(t) \in \{T-1, T\}$, where T is defined as $T = \lfloor N/2^{k+1} \rfloor$. To avoid a wrap modulo N we have to choose $b_{r_k}(t) \in \{T-1, T\}$ such that the inequation

$$N - 1 - 2^{k+1} < b_{r_k}(t)2^{k+1} + t \le N - 1$$

holds. By $T2^{k+1} = N - r_k$ we get

$$b_{r_k}(t) = \begin{cases} T-1 & : \quad t \ge r_k \\ T & : \quad \text{else.} \end{cases} \tag{5.9}$$

With similar arguments as above we get:

$$\begin{aligned} B_+(B_k) &= \{2^{k+1}z + t\}_{2^k \le t \le 2^{k+1}-1, 0 \le z \le b_{r_k}(t)} \\ &= \{2^{k+1}z + t + 2^k\}_{0 \le t \le 2^k - 1, 0 \le z \le b_{r_k}(t)} \end{aligned} \tag{5.10}$$

Let $B_k(z)$ denote the k-th least significant bit of (the binary representation of) $z \in \mathbb{Z}_N$. Obviously, if the modulus N is not a power of 2, the function B_k generally is not unbiased. This is made precise in the following lemma:

Lemma 5.6.3. *$bias_{B_k}(N) = \min\{N \bmod 2^{k+1}, 2^{k+1} - N \bmod 2^{k+1}\}$.*

Proof. If 2^{k+1} divides N, the bias is zero and the result holds true. If not, we cut the elements of $\mathbb{Z}_N$ into $t := \lfloor N/2^{k+1} \rfloor$ segments S_i of consecutive elements, each of size 2^{k+1}. Since N is not divisible by 2^{k+1}, the last segment has less elements. More precisely S_t consists of the last $r := N \bmod 2^{k+1}$ elements in $\mathbb{Z}_N$. Observing that for $1 \le i \le t-1$, $\sum_{z \in S_i} B_k(z) = 0$, we obtain

$$\mathrm{bias}_{B_k}(N) = \left| \sum_{z \in \mathbb{Z}_N} B_k(z) \right| = \left| \sum_{i=1}^{t} \sum_{z \in S_i} B_k(z) \right| = \left| \sum_{z \in S_t} B_k(z) \right| = \left| \sum_{z=0}^{r-1} B_k(z) \right|.$$

The last term evaluates to r, if $r \le 2^k$, and to $2^{k+1} - r$, otherwise. $\square$

Corollary 5.6.4. *$B_k : \mathbb{Z}_N \to \{-1,1\}$ is unbiased if $k = k(N) = o(\log N)$.*

By equations (5.8) and (5.10), we can observe the following:

Lemma 5.6.5. *For odd N, $B_k : \mathbb{Z}_N \to \{-1,1\}$ is strongly semilinear of length 2^k.*

Hence, Lemma 5.6.2 implies the bound $EXP_{B_k}(N,d) \le 2^{k+1}\log(N/d)$. However, by a more tedious computation we can achieve a stronger bound for every bit B_k:

Lemma 5.6.6. *For odd N, $EXP_{B_k}(N,d) \le 3bias_{B_k}(N)\log(N/d)$.*

Proof. By definition,

$$EXP_B(N,d) = \frac{d}{N} \cdot \sum_{a\in\mathbb{Z}^*_{N/d}} |\langle B, E_{ad}\rangle| \quad .$$

Let $b_{r_k}(t)$ and T be defined as in equation (5.9). By the characterizations of equation (5.8) and equation (5.10) we can continue with

$$\begin{aligned} &= \frac{d}{N} \sum_{a\in\mathbb{Z}^*_{N/d}} \left| \sum_{t=0}^{2^k-1} \left(\sum_{z=0}^{b_{r_k}(t+2^k)} E_{ad}(2^k + 2^{k+1}z + t) - \sum_{z=0}^{b_{r_k}(t)} E_{ad}(+2^{k+1}z + t) \right) \right| \\ &\le \frac{d}{N} \cdot \sum_{a\in\mathbb{Z}^*_{N/d}} \left| \sum_{t=0}^{2^k-1} \sum_{z=0}^{b_{r_k}(t+2^k)} E_{ad}(2^k + 2^{k+1}z + t) \right| \\ &\qquad\qquad + \frac{d}{N} \cdot \sum_{a\in\mathbb{Z}^*_{N/d}} \left| \sum_{t=0}^{2^k-1} \sum_{z=0}^{b_{r_k}(t)} E_{ad}(2^{k+1}z + t) \right| \quad . \end{aligned}$$

We will show the bound

$$\frac{d}{N} \cdot \sum_{a\in Z^*_{N/d}} \left| \sum_{t=0}^{2^k-1} \sum_{z=0}^{b_{r_k}(t)} E_{ad}(2^{k+1}z + t) \right| \le \frac{3}{2}\mathrm{bias}_{B_k}(N) \log\frac{N}{d}$$

on the second term. The bound on the first term can be achieved analogously. Put

$$L(a,d) = \sum_{t=0}^{2^k-1} \sum_{z=0}^{b_{r_k}(t)} E_{ad}(2^{k+1}z + t).$$

Separating $b_{r_k}(t) = T-1$ and $b_{r_k}(t) = T$ we get

$$\begin{aligned}
|L(a,d)| &\leq \left|\sum_{t=0}^{2^k-1}\sum_{z=0}^{T-1} E_{ad}(2^{k+1}z+t)\right| + \left|\sum_{t=0}^{\min\{2^k-1,r_k-1\}} E_{ad}(2^{k+1}T+t)\right| \\
&\overset{(2.18)}{\leq} \left|\sum_{t=1}^{2^k}\sum_{z=1}^{T} E_{ad}(2^{k+1}(z-1)+(t-1))\right| + r_k \\
&\overset{(2.20)}{=} |E_{ad}(-2^{k+1}-1)| \left|\sum_{t=1}^{2^k} E_{ad}(t)\sum_{z=1}^{T} E_{ad}(2^{k+1}z)\right| + r_k \\
&\overset{(2.18)}{=} \left|\sum_{t=1}^{2^k} E_{ad}(t)\sum_{z=1}^{T} E_{ad}(2^{k+1}z)\right| + r_k.
\end{aligned}$$

We apply (2.27) twice and rewrite this term as

$$= \left|\frac{E_{ad}(2^k)-1}{E_{ad}(1)-1}\cdot\frac{E_{ad}(2^{k+1}T)-1}{E_{ad}(2^{k+1})-1}\right| + r_k.$$

By observing that $E_{ad}(2^{k+1})-1 = (E_{ad}(2^k)+1)(E_{ad}(2^k)-1)$ and $2^{k+1}T = N-r_k$ we get

$$\begin{aligned}
&= \left|\frac{1}{E_{ad}(1)-1}\cdot\frac{E_{ad}(N-r_k)-1}{E_{ad}(2^k)+1}\right| + r_k \\
&= \left|\frac{E_{ad}(-r_k)-1}{E_{ad}(1)-1}\cdot\frac{1}{E_{ad}(2^k)+1}\right| + r_k.
\end{aligned}$$

Now we use bound (2.29) and equation (2.24):

$$\leq \quad r_k\left|\frac{1}{E_{ad}(2^k)+1}\right| + r_k = \frac{r_k}{2|\cos(2^k\pi ad/N)|} + r_k. \tag{5.11}$$

On the other hand, when we did the separation between $b_{r_k}(t) = T-1$ and $b_{r_k}(t) = T$, letting the index of the first sum go to T, instead of $T-1$ leads to:

$$\begin{aligned}
|L(a,d)| &\leq \left|\sum_{t=0}^{2^k-1}\sum_{z=0}^{T} E_{ad}(2^{k+1}z+t)\right| + \left|\sum_{t=0}^{\min\{2^k-1,2^{k+1}-r_k-1\}} E_{ad}(2^{k+1}T+t)\right| \\
&\leq \left|\sum_{t=0}^{2^k-1}\sum_{z=0}^{T} E_{ad}(2^{k+1}z+t)\right| + (2^{k+1}-r_k).
\end{aligned}$$

Following along the lines of the proof above we then get

$$\leq \quad \frac{2^{k+1}-r_k}{2|\cos(2^k\pi ad/N)|} + 2^{k+1} - r_k. \tag{5.12}$$

By Lemma 5.6.3 we know that the bias of B_k equals to $\min\{r_k, 2^{k+1} - r_k\}$. Since both bounds (5.11) and (5.12) on $|L(a,d)|$ hold independently from each other this leads to the bound

$$|L(a,d)| \leq \frac{\text{bias}_{B_k}(N)}{2|\cos(2^k \pi ad/N)|} + \text{bias}_{B_k}(N) \leq \frac{3\text{bias}_{B_k}(N)}{2|\cos(2^k \pi ad/N)|}.$$

Summing the terms $|L(a,d)|$ up over all $a \in \mathbb{Z}^*_{N/d}$ we get

$$\frac{d}{N} \cdot \sum_{a \in \mathbb{Z}^*_{N/d}} |L(a,d)| \quad \leq \quad \frac{3d\text{bias}_{B_k}(N)}{2N} \cdot \sum_{a \in \mathbb{Z}^*_{N/d}} \frac{1}{|\cos(2^k \pi ad/N)|}$$

Since N is odd, the element 2^{-1} exists in $\mathbb{Z}^*_N$ and hence also in $\mathbb{Z}^*_{N/d}$. Substituting $a = 2^{-k}b$ we get

$$\begin{aligned}
&\leq \frac{3d\text{bias}_{B_k}(N)}{2N} \cdot \sum_{b \in \mathbb{Z}^*_{N/d}} \frac{1}{|\cos(\pi bd/N)|} \\
&\leq \frac{3d\text{bias}_{B_k}(N)}{2N} 2 \int_1^{(N/d-1)/2} \frac{1}{\cos(\pi bd/N)} db \\
&\leq \frac{3d\text{bias}_{B_k}(N)}{N} \int_1^{(N/d-1)/2} \frac{1}{\cos(\pi bd/N)\sin(\pi bd/N)} \\
&= \frac{3\text{bias}_{B_k}(N)}{\pi} \left(\ln(\tan(\frac{\pi}{2} - \frac{\pi d}{2N})) - \ln(\tan(\frac{\pi d}{N})) \right) \\
&\leq \frac{3\text{bias}_{B_k}(N)}{\pi} 2 \ln(\frac{2N}{\pi d}) \\
&\leq \frac{3}{2}\text{bias}_{B_k}(N) \log \frac{N}{d}
\end{aligned}$$

□

Corollary 5.6.7. *For odd N and unbiased B_k we have $EXP_{B_k}(N,d) = N^{o(1)}$ and hence B_k is balanced.*

One may wonder how $EXP_{B_k}(N,d)$ behaves for even N. We argue that $EXP_{B_k}(N,d)$ can get close to N in some cases. We write $N = 2M$ and study $EXP_{B_k}(N,d)$ for $k = 0$ (the least significant bit) and the divisor $d = M$. Since $E_M(2z) = E_M^M(z) = 1$ and $E_M(1) = -1$, we have

$$\begin{aligned}
EXP_{B_k}(N,M) &= \frac{1}{2} \left| \sum_{z=0}^{M-1} E_M(2z) - \sum_{z=0}^{M-1} E_M(2z+1) \right| \\
&\overset{(2.20)}{=} \frac{1}{2} \left| \sum_{z=0}^{M-1} E_M(2z) - E_M(1) \sum_{z=0}^{M-1} E_M(2z) \right| \\
&= \frac{1}{2} \left| (1 - E_M(1)) \sum_{z=0}^{M-1} E_M^M(z) \right| \\
&= M = N/2.
\end{aligned}$$

5.6.3 Limitations

With similar considerations as in the last subsection, one may raise the question whether for a prime modulus every unbiased Boolean function must consequently be smooth. We indicate that in general this is not true by giving a Boolean function that is "far from being smooth", though it is unbiased. Furthermore, we motivate this observation by means of Theorem 5.4.1.

Let P be a prime. If there is an integer x such that $x^2 = y$ in $\mathbb{Z}_P$, then x is said to be a *quadratic residue* (modulo P). If not, x is said to be a *quadratic nonresidue*. For primes P, the Legendre symbol $\left(\frac{x}{P}\right)$ is a number theoretic function which is defined to be equal to ± 1 depending on whether x is a quadratic residue modulo P or not. Furthermore it is convenient to define $\left(\frac{x}{P}\right)$ as 0 if $x = 0$.

It is well known that the Legendre symbol has the multiplicative property

$$\left(\frac{xy}{P}\right) = \left(\frac{x}{P}\right)\left(\frac{y}{P}\right). \tag{5.13}$$

For $i \in \{-1, 0, 1\}$, put $\mathcal{L}_i = \{x \in \mathbb{Z}_P : \left(\frac{x}{P}\right) = i\}$. By (5.13) there is a 1-1 correspondence between $\mathcal{L}_1$ and $\mathcal{L}_{-1}$ given by the relation

$$\left(\frac{x}{P}\right) = 1 \Leftrightarrow \left(\frac{\mu x}{P}\right) = -1 \ , \tag{5.14}$$

where μ be an arbitrary element with Legendre symbol -1.
Now consider the binary function $B_{\mathcal{L}} : \mathbb{Z}_P \to \{-1, 1\}$ defined as

$$B_{\mathcal{L}}(x) = \begin{cases} \left(\frac{x}{P}\right) & : \ x \in \mathbb{Z}_P^* \\ 1 & : \ x = 0. \end{cases} \tag{5.15}$$

By (5.14) we have $|\mathcal{L}_1| = |\mathcal{L}_{-1}| = (P-1)/2$. Therefore $\text{bias}_{B_{\mathcal{L}}}(P) = 1$ and $B_{\mathcal{L}}$ is unbiased.

It is not hard to show [30, Equation (42)] that for $0 \neq a \in \mathbb{Z}_P$ the following equality holds:

$$\sum_{x=0}^{P-1} E_a(x^2) = \sum_{x=1}^{P-1} \left(\frac{x}{P}\right) E_a(x) \ . \tag{5.16}$$

By Lemma 2.5.1 we have

$$\begin{aligned} EXP_{B_{\mathcal{L}}}(P,1) &= \frac{1}{P} \cdot \sum_{a \in \mathbb{Z}_P^*} \left| \sum_{x \in \mathbb{Z}_P} B(x) E_a(x) \right| \\ &= \frac{1}{P} \cdot \sum_{a \in \mathbb{Z}_P^*} \left| 1 + \sum_{x=1}^{P-1} \left(\frac{x}{P}\right) E_a(x) \right| \\ &\overset{(5.16)}{=} \frac{1}{P} \cdot \sum_{a \in \mathbb{Z}_P^*} \left| 1 + \sum_{x=0}^{P-1} E_a(x^2) \right| \\ &\geq \frac{P-1}{P} (P^{1/2} - 1) \approx P^{1/2}, \end{aligned}$$

and hence the unbiased Boolean function $B_{\mathcal{L}}$ is *not* smooth. Note that the choice of $B_{\mathcal{L}}(0) = 1$ in (5.15) is somewhat arbitrary and does not much affect the bound on $EXP_{B_{\mathcal{L}}}(P, 1)$.

Actually, the property that makes $B_{\mathcal{L}}$ being not smooth is its multiplicativity (5.13). To see this let B be any Boolean function with multiplicative property $B(xy) = B(x) \cdot B(y)$ for all $x, y \in \mathbb{Z}_N$.[3] Since, for instance, the discrete multiplication $M(x, y) = xy$ also has a multiplicative property, $M(x, y) = xy = M(x, 1)M(1, y)$, this implies that the matrix $A := B \circ M$ is completely determined by its first row and column:

$$A(x, y) = A(x, 1) \cdot A(1, y).$$

Hence A has rank one. Or, equivalently, choose vectors $u, v \in \mathbb{R}^N$, $\|u\|_2 = \|v\|_2 = 1$ as $u_x = A(x, 1)/\sqrt{N}$ and $v_y = A(1, y)/\sqrt{N}$, resp. This gives $\|A\|_2 \geq |u^\top Av| = N$ and by means of Theorem 5.4.1, B cannot be smooth.

5.7 Applications

In Section 5.4 we proved that balanced Boolean predicates of functions with multiplicative or exponential structure have relatively small spectral norms. Now we present some natural functions having multiplicative or exponential structure (on a large subdomain) allowing us to apply our results from Section 5.4.

5.7.1 Discrete Multiplication and Division modulo arbitrary Integers

The discrete multiplication function $MUL : \{0, \ldots, N{-}1\} \times \{0, \ldots, N{-}1\} \to \mathbb{Z}_N$ is defined by

$$MUL(x, y) = xy.$$

Applying Theorem 5.4.1 (with $\mathcal{X} = \mathcal{Y} = \{0, \ldots, N{-}1\}$ and $\sigma, \tau : \{0, \ldots, N{-}1\} \to \mathbb{Z}_N$ defined as the identity) we immediately get:

Corollary 5.7.1. *For balanced Boolean functions $B : \mathbb{Z}_N \to \{-1, 1\}$,*

$$\|B \circ MUL\|_2 = N^{1/2+o(1)} \quad .$$

By the results from Chapter 4 this implies (for odd N): $B_k \circ MUL \notin \widetilde{\mathrm{LT}}_2$. On the other hand it is well known (see, e.g, [59]) that $B_k \circ MUL$ can be realized with TC_3-circuits, where B_k denotes the k-th least significant bit.[4] Thus $B_k \circ MUL \in \mathrm{TC}_3 \setminus \widetilde{\mathrm{LT}}_2$ is another binary function separating TC_3 from $\widetilde{\mathrm{LT}}_2$.

[3] Note the the predicate $B_{\mathcal{L}}$ has the multiplicative property only for $x, y \in \mathbb{Z}_N^*$. However, it is the same reason that makes $EXP_{B_{\mathcal{L}}}(P, 1)$ getting large.

[4] In fact, there exists a LT_3-circuit that *simultaneously* computes all the bits of MUL.

To perform discrete division in $\mathbb{Z}_N$, we may use any function $DIV : \{0, \ldots, N-1\} \times \{0, \ldots, N-1\} \to \mathbb{Z}_N$ satisfying

$$DIV(x, y) = x/y$$

for all inputs $(x, y) \in \{0, \ldots, N-1\} \times \{0, \ldots, N-1\}^*$ (recall that by $\{0, \ldots, N-1\}^*$ we denote the subset of all elements $x \in \{0, \ldots, N-1\}$ with $\gcd(x, N) = 1$). Since discrete division is nothing else than one inversion (permutation) plus discrete multiplication, we get the following:

Corollary 5.7.2. *For integers N with $N - \varphi(N) = O(N^\alpha)$ for some fixed constant $\alpha > 0$,*[5] *for balanced Boolean functions $B : \mathbb{Z}_N \to \{-1, 1\}$, we have*

$$\|B \circ DIV\|_2 = \begin{cases} N^{1/2+o(1)} & : \text{ prime } N \\ O(N^{(1+\alpha)/2}) & : \text{ non-prime } N. \end{cases}$$

Proof. Let $\mathcal{X} = \{0, \ldots, N-1\}$ and $\mathcal{Y} = \{0, \ldots, N-1\}^*$. Define the injective mapping $\tau : \mathcal{Y} \to \mathbb{Z}_N$ as $\tau(y) := y^{-1} \bmod N$. With $\sigma : \mathcal{X} \to \mathbb{Z}_N$ given by the identity, we have

$$DIV(x, y) = xy^{-1} = x\tau(y) = M_{\sigma,\tau}(x, y)$$

for all $(x, y) \in \mathcal{X} \times \mathcal{Y}$ showing that DIV has multiplicative structure on subdomain $\mathcal{X} \times \mathcal{Y}$. Noting that $N - |\mathcal{X}| = 0$ and $N - |\mathcal{Y}| = N - \varphi(N) = O(N^\alpha)$, the result follows by Theorem 5.4.1. □

5.7.2 Discrete Exponentiation modulo Primes

For primes N, the discrete exponentiation function $POW : \{0, \ldots, N-1\} \times \{0, \ldots, N-2\} \to \mathbb{Z}_N^*$ is given by

$$POW(x, y) = x^y.$$

Corollary 5.7.3. *For primes N and for balanced binary functions $B : \mathbb{Z}_N \to \{-1, 1\}$, we have*

$$\|B \circ POW\|_2 = N^{23/24+o(1)} \quad .$$

Proof. We show that the $N \times N-1$ matrix $B \circ POW$ has a $N-1 \times N-1$ submatrix whose spectral norm is bounded by $N^{23/24+o(1)}$. By similar arguments as used in Lemma 5.5.15 (and the fact that adding all-zero columns or rows does not affect the spectral norm) we know that the corresponding two spectral norms can only differ by an additive factor $O(\sqrt{N})$ leading asymptotically to the same bound on $\|B \circ POW\|_2$.
The function $POW^* : \{0, \ldots, N-2\} \times \{0, \ldots, N-2\} \to \mathbb{Z}_N^*$ is defined by

$$POW^*(x, y) = x^y$$

[5] This is the case for integers N whose smallest prime-factor is of the order $\Omega(N^{1-\alpha})$.

for all $(x, y) \in \{0, \ldots, N-2\} \times \{0, \ldots, N-2\}$, and obviously the matrix $B \circ POW^*$ is a $N-1 \times N-1$ submatrix of $B \circ POW$. Let g be an element of order $T = N-1$. Define $\mathcal{X} = \{1, \ldots, N-2\} \subseteq \{0, \ldots, N-2\}$ and $\mathcal{Y} = \{0, \ldots, N-2\}$. We claim that POW^* has exponential structure on subdomain $\mathcal{X} \times \mathcal{Y}$ with respect to g. (Clearly the Corollary follows by Theorem 5.4.3 once the claim is settled.) The function $\mathrm{idx}_g(x)$ defines a bijective mapping from $\{1, \ldots, N-1\}$ into $\mathbb{Z}_{N-1}$. Hence, the mapping $\sigma : \mathcal{X} \to \mathbb{Z}_{N-1}$ given by $\sigma(x) := \mathrm{idx}_g(x)$ is injective. Note that now σ has the property $g^{\sigma(x)} = x$ for all $x \in \mathcal{X}$. For $\tau : \mathcal{Y} \to \mathbb{Z}_{N-1}$ defined as the identity, it follows that

$$POW^*(x, y) = x^y = g^{\sigma(x)\tau(y)} = D_{\sigma,\tau}(x, y)$$

for all $(x, y) \in \mathcal{X} \times \mathcal{Y}$ and hence POW^* has exponential structure on subdomain $\mathcal{X} \times \mathcal{Y}$. This proves the claim. □

5.7.3 General Matrices with small Spectral Norm.

As mentioned in Section 2.4, the spectral norm of a matrix $A \in \{-1, 1\}^{N \times N}$ is minimal ($\|A\|_2 = N^{1/2}$) iff the matrix has orthogonal rows or columns. For the special case $N = 2^n$, an example of such matrices are the Hadamard matrices $H_{2^n} \in \{-1, 1\}^{2^n \times 2^n}$. It is well known (see [3, 23]) that general Hadamard matrices H_N (which are defined as real matrices H_N with $H_N H_N^\top = N \cdot I_N$) can only exist for $N = 1$ or 2 or for N divisible by 4. For odd N we present another natural family of ± 1-valued matrices with (nearly) minimal spectral norm. Consider the matrix $B_0 \circ M \in \{-1, 1\}^{N \times N}$ representing the least significant bit (or any other balanced binary function) of the discrete multiplication. This observations allows us to use Theorem 5.4.1 to get

$$\|B_0 \circ M\|_2 = N^{1/2 + o(1)}.$$

This gives (for odd N) a natural construction of $\{-1, 1\}^{N \times N}$ matrices with *nearly* minimal spectral norm.
The term implied by $N^{o(1)}$ (the "nearly") can be made more precise. Consider the case of prime $N \geq 11$. An easy computation shows $EXP_{B_k}(N, 1) \leq \frac{7}{5} \log N$ for $k = 0$ (the funny constant $7/5$ comes from (2.28)). Applying directly the Splitting Lemma (Lemma 5.5.1) we get $\|B_0 \circ M\|_2 \leq 1 + \frac{7}{5} N^{1/2} \log N \leq \frac{3}{2} N^{1/2} \log N$. For example, we may now apply Corollary 4.2.2 to lower bound the probabilistic communication complexity $\tilde{C}_f \geq \lceil n/2 - \log n - 3/5 \rceil$ of the function f such that matrix $B_0 \circ M$ represents f on $\{0, \ldots, N-1\} \times \{0, \ldots, N-1\}$ ($2^{n-1} < N \leq 2^n$).

5.7.4 Matrices with arbitrary Spectral Norm

Lemma 5.7.4. *If matrix $H : \mathbb{Z}_N \times \mathbb{Z}_N \to \mathbb{Z}_N$ defines a permutation on every row or every column, then*

$$\|B \circ H\|_2 \geq bias_B(N) \enspace .$$

Proof. Say matrix H defines a permutation on every column. Then, for any $y \in \mathbb{Z}_N$, we have $\sum_{x\in\mathbb{Z}_N} B(H(x,y)) = \text{bias}_B(N)$ and therefore

$$1/N \sum_{x,y\in\mathbb{Z}_N} B(H(x,y)) = \text{bias}_B(N) \ .$$

Now we can lower bound the spectral norm of $B \circ H$ in the following way

$$\|B\circ H\|_2 \geq \left| \sum_{x,y\in\mathbb{Z}_N} \frac{1}{\sqrt{N}} B(H(x,y)) \frac{1}{\sqrt{N}} \right| = \frac{1}{N} \left| \sum_{x,y\in\mathbb{Z}_N} B(H(x,y)) \right| = \text{bias}_B(N). \tag{5.17}$$

□

By changing $N^2 - \varphi^2(N)$ entries, the matrix M with $M(x,y) = xy$ can be transformed into a matrix M' defining a permutation on every row and column. By Lemma 5.5.15 (with $|\mathcal{X}| = |\mathcal{Y}| = \varphi(N)$) we get $|\|B \circ M\|_2 - \|B \circ M'\|_2| \leq 2\sqrt{N^2 - \varphi^2(N)}$ and thus by Lemma 5.7.4

$$\|B \circ M\|_2 \geq \text{bias}_B(B) - 2\sqrt{N^2 - \varphi^2(N)} \ . \tag{5.18}$$

Let $I \subseteq \mathbb{Z}_N$ be any consecutive interval of length $(N - N^{1/2+\varepsilon})/2$ for a constant $0 < \varepsilon < 1/2$. Define $B_\varepsilon : \mathbb{Z}_N \to \{-1, 1\}$ as the Boolean function that evaluates to 1 if and only if its input x belongs to the interval I. Obviously we have $\text{bias}_{B_\varepsilon}(N) = N^{1/2+\varepsilon}$. Since B_ε is strongly semilinear of length 1 (see Definition 5.6.1), we can infer from Corollary 5.5.13 the upper bound

$$\|B_\varepsilon \circ M\|_2 \leq \text{bias}_{B_\varepsilon}(N) + N^{1/2+o(1)} = O(N^{1/2+\varepsilon}) \ .$$

On the other hand, by equation (5.18) and for primes N, we have the lower bound

$$\|B_\varepsilon \circ M\|_2 \geq \text{bias}_{B_\varepsilon}(N) - \sqrt{8N} = \Omega(N^{1/2+\varepsilon}) \ .$$

We see that upper and lower bound only differ by a multiplicative constant. To be more precise, this constant is one as N goes to infinity. This shows that the bound from Corollaries 5.5.13 and 5.5.14 cannot be substantially improved beyond the value $\text{bias}_B(N)$. Moreover, we get the following explicit construction of matrices of arbitrary spectral norm:

Corollary 5.7.5. *For primes N and for every constant $0 < \varepsilon \leq 1/2$, there exist (explicitly constructible) matrices $A : \mathbb{Z}_N \times \mathbb{Z}_N \to \{-1, 1\}$ with*

$$\|A\|_2 = \Theta(N^{1/2+\varepsilon}) \ .$$

Proof. Choose $A = B_\varepsilon \circ M$ with B_ε as defined above and $M(x,y) = xy$. □

It is easy to see that our construction carries over to the discrepancy. By equation (2.13) (for the upper bound) and (5.17) (for the lower bound) we get:

Corollary 5.7.6. *For primes N and for every constant $0 < \varepsilon \leq 1/2$, there exist (explicitly constructible) matrices $A : \mathbb{Z}_N \times \mathbb{Z}_N \to \{-1, 1\}$ with*

$$\text{disc}(A) = \Theta(N^{-1/2+\varepsilon}) \ .$$

5.8 Cryptographic Applications

As we saw in the last section our two main theorems can be applied to some basic functions with multiplicative and exponential structure. In this section we show that our techniques can also be applied to functions appearing naturally in cryptographic protocols. To simplify notation we will sometimes, when it is clear to the reader, identify the set $\mathbb{Z}_N$ with the set of integers $\{0, \ldots, N-1\}$ and vice versa.

Example 5.8.1 (Pointcheval Cryptosystem [48]). *Let N be an RSA modulus, i.e., $N = pq$ for two primes p, q of (roughly) the same bit length. Let e be an integer with $\gcd(e, \varphi(N)) = 1$. The tuple (N, e) is the public key and an integer d with $ed = 1 \bmod \varphi(N)$ is the secret key. Note that by the knowledge of either p or q, $\varphi(N) = (p-1)(q-1)$ can be computed and so can d. The following function*

$$F : \mathcal{R} \times \mathbb{Z}_N \to \mathbb{Z}_N^* \times \mathbb{Z}_N, \quad F(r, m) = ((r-1)^e, r^e \cdot m),$$

where $r \in \mathcal{R} = \{r \in \mathbb{Z}_N : r - 1 \in \mathbb{Z}_N^*\}$ is chosen at random, encodes a message m. Note that it can be evaluated by knowledge of the public key (N, e) only. Define $\mathcal{X} := \{x : x^d + 1 \in \mathbb{Z}_N^*\}$ and note that $|\mathcal{X}| = \varphi(N)$. As for decoding (given the "secret key" d), we may use any function $H : \{0, \ldots, N–1\} \times \{0, \ldots, N–1\} \to \mathbb{Z}_N$ that satisfies

$$H(x, y) = y(x^d + 1)^{-e}$$

for all $x \in \mathcal{X}$ and $y \in \mathcal{Y} = \{0, \ldots, N–1\}$. It is easy to check that this implies that $H((r-1)^e, r^e \cdot m) = m$ for all $r \in \mathcal{R}$. Outside the subdomain $\mathcal{X} \times \mathcal{Y}$ (when H is applied to invalid ciphers), we may however allow H to attain arbitrary values. The mapping $\sigma : \mathcal{X} \to \mathbb{Z}_N^*$ given by $\sigma(x) := (x^d + 1)^{-e}$ defines a bijection. With $\tau : \mathcal{Y} \to \mathbb{Z}_N$ defined as the identity, we have

$$H(x, y) = \sigma(x)y = M_{\sigma,\tau}(x, y)$$

for all $x \in \mathcal{X}$ and $y \in \mathcal{Y}$, and we may conclude that any correct decoding function H has a multiplicative structure on subdomain $\mathcal{X} \times \mathcal{Y}$. Since $|\mathcal{X}| = \varphi(N) = (p-1)(q-1)$, the part outside the subdomain has a relatively small size: $N - |\mathcal{X}| = N - \varphi(N) = p + q - 1 \approx 2N^{1/2}$.

Example 5.8.2 (El Gamal Cryptosystem [15]). *Let N be a prime modulus and let $g \in \mathbb{Z}_N^*$ be an element of order $N - 1$. The "public key" has the form $h = g^s$ for some $s \in \mathbb{Z}_{N-1}$. A message $m \in \mathbb{Z}_N$ is encrypted with "randomness" $r \in \mathbb{Z}_{N-1}$ by evaluating the function*

$$F : \mathbb{Z}_{N-1} \times \mathbb{Z}_N \to \mathbb{Z}_N^* \times \mathbb{Z}_N, \quad F(r, m) = (g^r, h^r m).$$

As for decoding (given the "secret key" s), we may use any function $H : \{0, \ldots, N-1\} \times \{0, \ldots, N-1\} \to \mathbb{Z}_N$ that satisfies

$$H(x, y) = yx^{-s}$$

for all $x \in \mathcal{X} := \{1, \ldots, N-1\}$ *and all* $y \in \mathcal{Y} := \{0, \ldots, N-1\}$. *It is easy to check that* $H(g^r, h^r m) = m$. *Since for the injective mappings* $\sigma : \mathcal{X} \to \mathbb{Z}_N$ *given by* $\sigma(x) := x^{-s}$, *and* $\tau : \mathcal{Y} \to \mathbb{Z}_N$ *given by the identity,* H *satisfies*

$$H(x, y) = yx^{-s} = y\sigma(x) = M_{\sigma,\tau}(x, y),$$

for all $x \in \mathcal{X}$ *and* $y \in \mathcal{Y}$, *we may conclude that* H *has a multiplicative structure on subdomain* $\mathcal{X} \times \mathcal{Y}$. *Note that* $N - |\mathcal{X}| = N - (N-1) = 1$.

Example 5.8.3 (RSA-Paillier Cryptosystem [6]). *It is a modification [6] of the original cryptosystem from Paillier [45]. Let* $N = pq$ *be a RSA modulus and let* e *be such that* $\gcd(e, \lambda(N^2)) = 1$, *where* $\lambda(N^2) = pq \cdot \mathrm{lcm}(p-1, q-1)$ *is the Carmichael function. Choose* d *such that* $ed = 1 \bmod \varphi(N)$. *The following function*

$$F' : \mathbb{Z}_N \times \mathbb{Z}_N^* \to \mathbb{Z}_{N^2}^*, \quad F'(m, r) = r^e(1 + mN) \bmod N^2$$

encodes a message $m \in \mathbb{Z}_N$. *In his original paper [45], Paillier proves that this function is a bijection on* $\mathbb{Z}_N \times \mathbb{Z}_N^* \to \mathbb{Z}_{N^2}^*$. *Furthermore, in [7] it was proven that* F' *defines a (bijective) trapdoor function whose one-wayness is equivalent to the one-wayness of the RSA function [52]. Every element* $z \in \mathbb{Z}_{N^2}^*$ *has a unique decomposition* $z = z_1 + z_2 N$, *with* $z_1 \in \mathbb{Z}_N^*$ *and* $z_2 \in \mathbb{Z}_N$. *We denote the induced bijective decomposition mapping by*

$$\pi : \mathbb{Z}_{N^2}^* \to \mathbb{Z}_N^* \times \mathbb{Z}_N, \quad \pi(z) = (\pi_1(z), \pi_2(z)) = (z \bmod N, z \operatorname{div} N).$$

As suggested in [6], we consider the following bijective encoding function $F : \mathbb{Z}_N \times \mathbb{Z}_N^* \to \mathbb{Z}_N^* \times \mathbb{Z}_N$ *whose one-wayness is equivalent to the one-wayness of the function* F' *(and hence also to the one-wayness of RSA; see also [7] for a formal statement):*

$$F(m, r) = (\pi_1(r^e), m \cdot \pi_2(r^e)).$$

Define $\mathcal{X} = \{1, \ldots, N-1\}$. *As for decoding (given the "secret key"* d*), we may use any function* $H : \{0, \ldots, N-1\} \times \{0, \ldots, N-1\} \to \mathbb{Z}_N$ *that satisfies*

$$H(x, y) = y \cdot (\pi_2((\pi_1(x^d))^e \bmod N^2))^{-1}$$

for all $x \in \mathcal{X}$ *and* $y \in \mathcal{Y} = \{0, \ldots, N–1\}$. *It is easy to check that this implies that* $H(F(m, r)) = m$ *for all* $r \in \mathbb{Z}_N^*$. *Since for the mappings* $\sigma : \mathcal{X} \to \mathbb{Z}_N$, $\sigma(x) := (\pi_2((\pi_1(x^d))^e \bmod N^2))^{-1}$ *(*F *is bijective, so* σ *is injective on its domain) and* $\tau : \mathcal{Y} \to \mathbb{Z}_N$ *given by the identity,* H *satisfies*

$$H(x, y) = y\sigma(x) = M_{\sigma,\tau}(x, y)$$

for all $x, y \in \mathcal{X} \times \mathcal{Y}$, *we may conclude that* H *has a multiplicative structure on subdomain* $\mathcal{X} \times \mathcal{Y}$. *Again the part outside the subdomain has relatively small size:* $N - |\mathcal{X}| = N - \varphi(N) = p + q - 1 \approx 2N^{1/2}$.

In all these examples we have a multiplicative structure (Definition 5.2.1) on different (but sufficiently large) subdomains. Combining it with Theorem 5.4.1 we get the following:

Corollary 5.8.4. *Let B be a balanced binary function.*

1. *Let $H : \{0,\ldots,N-1\} \times \{0,\ldots,N-1\} \to \mathbb{Z}_N$ be a decoding function of the Pointcheval cryptosystem as given in example 5.8.1. Then,*
$$\|B \circ H\|_2 = O(N^{3/4}) \ .$$

2. *Let $H : \{0,\ldots,N-1\} \times \{0,\ldots,N-1\} \to \mathbb{Z}_N$ be a decoding function of the El Gamal cryptosystem as given in example 5.8.2. Then,*
$$\|B \circ H\|_2 = N^{1/2+o(1)} \ .$$

3. *Let $H : \{0,\ldots,N-1\} \times \{0,\ldots,N-1\} \to \mathbb{Z}_N$ be a decoding function of the RSA-Paillier's cryptosystem as given in example 5.8.3. Then,*
$$\|B \circ H\|_2 = O(N^{3/4}) \ .$$

Note that in the last three examples the decoding functions have multiplicative structure. And this is independent of the knowledge of the secret key. For example, in the El Gamal Cryptosystem, the decoding function (given the secret key s) is $H(x,y) = yx^{-s}$. On the other hand, an attacker without knowledge of the secret key s (but knowing the public key containing $h = g^s$) has to decode a given ciphertext (x,y) by computing $H(x,y) = yx^{\mathrm{idx}_g h}$. In both cases we have the same multiplicative structure. In particular, since the discrete logarithm $\mathrm{idx}_g : \mathbb{Z}_N^* \to \mathbb{Z}_{N-1}$ defines a bijection, the concept of the multiplicative structure of a function does not capture the difference between decoding with the knowledge of the secret key (what is computational feasible) and decoding without the knowledge of the secret key (what is assumed to be computational infeasible). Therefore, Corollary 5.8.4 is valid with and without the knowledge of the secret key.

We proceed with presenting two examples of functions H with exponential structure.

For primes N, let $g \in \mathbb{Z}_N^*$ be an element of order $T = N-1$. To compute the Diffie-Hellman function (see Subsection 5.2.2) we may take any function $DH : \{0,\ldots,N-1\} \times \{0,\ldots,N-1\} \to \mathbb{Z}_N$ that satisfies
$$DH(x,y) = g^{\mathrm{idx}_g x \cdot \mathrm{idx}_g y}$$
for all $(x,y) \in \{1,\ldots,N-1\} \times \{1,\ldots,N-1\}$.

Corollary 5.8.5 (Diffie-Hellman secret modulo primes). *For primes N, for elements g of order $T = N-1$, and for balanced binary functions $B : \mathbb{Z}_N \to \{-1,1\}$, we have*
$$\|B \circ DH\|_2 = N^{23/24+o(1)} \ .$$

Proof. As already argued before, it is sufficient to show that the $N \times N$ matrix $B \circ DH$ has a $N-1 \times N-1$ submatrix with spectral norm bounded by $N^{23/24+o(1)}$.

Define $\mathcal{X} = \mathcal{Y} = \{1, \ldots, N-2\}$. Let $DH^* : \{0, \ldots, N-2\} \times \{0, \ldots, N-2\} \to \mathbb{Z}_N$ be the function given by

$$DH^*(x, y) = DH(x, y)$$

for all $(x, y) \in \{0, \ldots, N-2\} \times \{0, \ldots, N-2\}$. Note that, by definition, $B \circ DH^*$ is a $N-1 \times N-1$ submatrix of $B \circ DH$. We show that DH^* has exponential structure on subdomain $\mathcal{X} \times \mathcal{Y}$ with respect to g. The function $\mathrm{idx}_g(x)$ defines a bijective mapping from $\{1, \ldots, N-1\}$ into $\mathbb{Z}_{N-1}$. Hence, the mapping $\sigma : \mathcal{X} \to \mathbb{Z}_{N-1}$ given by $\sigma(x) := \mathrm{idx}_g(x)$ is injective. For $\tau := \sigma$, it follows that

$$DH^*(x, y) = g^{\mathrm{idx}_g x \cdot \mathrm{idx}_g y} = g^{\sigma(x)\tau(y)} = D_{\sigma,\tau}(x, y)$$

for all $(x, y) \in \mathcal{X} \times \mathcal{Y}$ and DH^* has exponential structure on subdomain $\mathcal{X} \times \mathcal{Y}$. Note that the part outside the subdomain has a relatively small size: $N - 1 - |\mathcal{X}| = N-1-|\mathcal{Y}| = 1$. Now by Theorem 5.4.2 we have $\|B \circ DH^*\|_2 = N^{23/24+o(1)}$ and the result follows. □

The next example can be viewed as "computing the Diffie-Hellman function given the secret keys." Let N be an arbitrary modulus and let g be an element of order T. To compute the Diffie-Hellman secret given x and y, we may use the function $DK : \{0, \ldots, T-1\} \times \{0, \ldots, T-1\} \to \mathbb{Z}_N$ given by

$$DK(x, y) = g^{xy}.$$

The following corollary follows directly from Theorem 5.4.2 and Theorem 5.4.3 (with $\mathcal{X} = \mathcal{Y} = \{0, \ldots, N-2\}$ and $\tau, \sigma : \{0, \ldots, N-2\} \to \mathbb{Z}_{N-1}$ defined as the identity):

Corollary 5.8.6 (Diffie-Hellman key modulo arbitrary integers). *For elements g of order $T = N-1$, and for balanced binary functions $B : \mathbb{Z}_N \to \{-1, 1\}$, we have*

$$\|B \circ DK\|_2 = T^{21/32} N^{5/16+o(1)} \quad .$$

Moreover, for primes N,

$$\|B \circ DK\|_2 = T^{5/6} N^{1/8+o(1)} \quad .$$

5.9 Extensions

In Subsection 5.5.1 we saw that a good bound on the spectral norm $\|E_a \circ H\|_2$ (plus a balanced binary function B) implies a good bound on $\|B \circ H\|_2$. Examples of H with good bounds on $\|E_a \circ H\|_2$ were given in Subsection 5.5.2. In this section we extend the results of Subsection 5.5.2 by presenting some more candidates of functions H achieving sufficiently good bounds on $\|E_a \circ H\|_2$.

Lemma 5.9.1. *Let $P : \mathbb{Z}_N \times \mathbb{Z}_N \to \mathbb{Z}_N$, $P(x, y) = x^2 + p_{11}xy + y^2 + p_{10}x + p_{01}y + p_{00}$ be a monic polynomial of second degree with $\gcd(p_{11}, N) = 1$. Then for any $0 \neq a \in \mathbb{Z}_N$ with $\gcd(N, a) = d$,*

$$\|E_{ad} \circ P\|_2 \leq (dN)^{1/2} \quad .$$

Proof. We write a as $d \cdot a_0$ with $\gcd(a_0, N) = 1$. For every fixed $x, y \in \mathbb{Z}_N$ we have by (2.20) and (2.18),

$$\begin{aligned}\left|\sum_{z\in\mathbb{Z}_N} E_a(P(x,z) - P(y,z))\right| &= \left|\sum_{z\in\mathbb{Z}_N} E_{da_0}(x^2 - y^2 + zp_{11}(x-y) + p_{10}(x-y))\right| \\ &= d\left|\sum_{z\in\mathbb{Z}_N} E_{a_0}^{N/d}(zp_{11}(x-y))\right| \quad ,\end{aligned}$$

that evaluates by (2.25) to N if $x - y = 0 \pmod{N/d}$ and to 0 otherwise. By Lemma 5.5.4 we continue with

$$\begin{aligned}\|E_a \circ M\|_2^2 &\leq \max_{x\in\mathbb{Z}_N} \sum_{y\in\mathbb{Z}_N} \left|\sum_{z\in\mathbb{Z}_N} E_a(P(x,z) - P(y,z))\right| \\ &= dN \quad .\end{aligned}$$

□

Note that Lemma 5.9.1 covers the result of Lemma 5.5.7 (setting $P(x, y) = xy$). For the case of a general polynomial of maximal degree k we get the following result which is valid for primes N.

Lemma 5.9.2. *Let N be a prime. Let $P : \mathbb{Z}_N \times \mathbb{Z}_N \to \mathbb{Z}_N$ be a bivariate polynomial of maximal degree k over $\mathbb{Z}_N$. For all $a \in \mathbb{Z}_N^*$, the following holds:*

$$\|E_a \circ P\|_2 = O(k^{1/2}N^{3/4}) \quad .$$

Proof. Write P as $P(x, y) = \sum_{i=0}^k \sum_{j=0}^k a_{ij}x^iy^j$. For every fixed $x, y \in \mathbb{Z}_N$, $Q(x, y, z) := P(x, z) - P(y, z)$ is a polynomial of degree at most k in the indeterminate z:

$$Q(x,y,z) = \sum_{j=0}^{k} a_j(x,y)z^j, \quad a_j(x,y) = \sum_{i=0}^{k} a_{ij}(x^i - y^i) \quad .$$

For fixed $x \in \mathbb{Z}_N$, there can be at most k different $y \in \mathbb{Z}_N$ such that $Q(x, y, z)$ equals the zero polynomial in the indeterminate z. We denote those y by the exceptional set E_x ($|E_x| \leq k$ for all $x \in \mathbb{Z}_N$). For every other $y \notin E_x$ we have $\gcd(a_1(x, y), a_2(x, y), \ldots, a_k(x, y), N) = 1$ (since there must be at least one coefficient $a_j(x, y) \neq 0$) and by Lemma 2.5.2,

$$\left|\sum_{z\in\mathbb{Z}_N} E_a(P(x,z) - P(y,z))\right| = \left|\sum_{z\in\mathbb{Z}_N} E_a(Q(x,y,z))\right|$$

can be bounded by $(k-1)N^{1/2}$ for $y \notin E_x$ and by N for $y \in E_x$. By Lemma 5.5.4 we continue with

$$\begin{aligned} \|E_a \circ M\|_2^2 &\leq \max_{x\in\mathbb{Z}_N} \sum_{y\in\mathbb{Z}_N} \left| \sum_{z\in\mathbb{Z}_N} E_a(P(x,z) - P(z,y)) \right| \\ &\leq (N-k)(k-1)N^{1/2} + kN = O(kN^{3/2}) \ . \end{aligned}$$

□

Corollary 5.9.3. *For primes N let $P : \mathbb{Z}_N \times \mathbb{Z}_N \to \mathbb{Z}_N$ be a bivariate polynomial of maximal degree $O(N^\alpha)$ for a constant $0 \leq \alpha < 1/2$. Then, for balanced binary functions B,*

$$\|B \circ P\|_2 = N^{3/4+\alpha/2+o(1)} \ .$$

5.10 Complexity Theoretic Implications

In this section we summarize the most striking complexity theoretic implications of our bounds on the spectral norm we achieved earlier this chapter.

Let $B : \mathbb{Z}_N \to \{-1, 1\}$ be a balanced binary function (see Theorem 5.4.4 for examples). Let $H : \{0, \ldots, T-1\} \times \{0, \ldots, T-1\} \to \mathbb{Z}_N$ be any function having multiplicative or exponential structure on a sufficiently large subdomain. Natural examples of such functions (such as the discrete multiplication and exponentiation, the Diffie-Hellman function and others) were given in Sections 5.7 and 5.8. Applying Theorem 5.4.1, 5.4.3, or 5.4.2 leads to an upper bound of order $T^{1-\varepsilon}$ on the spectral norm of matrix $B \circ H \in \{-1, 1\}^{T\times T}$. See Corollaries 5.7.1, 5.7.2, 5.7.3, 5.8.4, 5.8.5, 5.8.6, and 5.9.3.

Let $f_n : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ be a binary function such that the matrix $B \circ H$ represents f_n on $\{0, \ldots, T-1\} \times \{0, \ldots, T-1\}$ ($2^{n-1} < T \leq 2^n$). Then,

- f_n is not contained in the complexity class $\widetilde{\mathrm{LT}}_2$ (Corollary 4.1.3). More precisely, f_n is not computable by threshold circuits of polynomial size, depth 2, where the bottom layer has polynomially bounded integer weights and the remaining gates have unbounded integer weights.
- the deterministic and probabilistic communication complexity of f_n is linear in n (Corollaries 4.2.1 and 4.2.2).

Let $g_n : \{0,1\}^{n-1} \times \{0,1\}^{n-1} \to \{0,1\}$ be the binary function such that the matrix $B \circ H$ represents g_n on $\{0, \ldots, 2^{n-1}-1\} \times \{0, \ldots, 2^{n-1}-1\}$ ($2^{n-1} \leq T < 2^n$). $B \circ H$ is then a $2^{n-1} \times 2^{n-1}$ submatrix of the $T \times T$ matrix considered above which has by (2.5) the same upper bound on its spectral norm. Then Corollary 4.1.5 shows that g_n is not contained in the complexity class PL_∞.[6]

[6] The reason for this technical restriction (considering the $2^{n-1} \times 2^{n-1}$ submatrix of $B \circ H$) is that the L_∞ norm makes only sense for matrices whose dimension is a power of 2.

Never believe on faith,
See for yourself!
What you yourself don't learn
you don't know.

BERTOLT BRECHT,
THE MOTHER, 1932

Chapter 6

Polynomial Interpolation

6.1 Introduction

In this chapter let $\mathbb{F}_q$ denote the finite field of order q with a prime power q and let $0 \neq g \in \mathbb{F}_q$ be an element of order T. As seen in the last chapter, the security of the Diffie-Hellman key exchange for the group G generated by g depends on the intractability of the *Diffie-Hellman mapping* DH_g defined by

$$\mathrm{DH}_g : G \times G \to G, \quad (g^x, g^y) \mapsto g^{xy} \ .$$

For breaking the Diffie-Hellman cryptosystem it would be sufficient to have a low degree polynomial that coincides with the mapping DH_g on a large subset of $G \times G$. In [36] and [64] it was shown that such a polynomial doesn't exist for several types of subsets. Since

$$g^{2xy} = g^{(x+y)^2} g^{-x^2} g^{-y^2},$$

and square roots in finite fields can be efficiently calculated (see e. g. [1, Chapter 7]), the problem of computing the Diffie-Hellman mapping DH_g is efficiently reducible to the problem of computing the univariate mapping

$$\mathrm{dh}_g : G \to G, \quad g^x \mapsto g^{x^2} \ .$$

For lower bounds on the degree of interpolation polynomials of dh_g see [8, 37, 57, 58].

Obviously, the Diffie-Hellman key exchange depends also on the hardness of the discrete logarithm idx_g defined by

$$\mathrm{idx}_g : G \to \mathbb{Z}_T, g^x \mapsto x \ .$$

For results on interpolation polynomials of idx_g see [8, 38, 39, 41, 42, 43, 57, 58, 63, 65].

In this chapter we consider mappings of the form

$$P\text{−}\mathrm{dh}_g : G \to G, \quad g^x \mapsto g^{P(x)}, \tag{6.1}$$

for a nonlinear polynomial $P(X) \in \mathbb{Z}_T[X]$ of small degree, with respect to T, say,

$$2 \leq \deg(P) \leq \log(T).$$

In [24] a toolbox of cryptographic functions called *P-Diffie-Hellman functions* including these mappings was suggested. In particular, it was proven that computing $P{-}\mathrm{dh}_g$ is computationally equivalent (through a polynomial time reduction) to computing dh_g. Hence, a low degree polynomial representation of $P{-}\mathrm{dh}_g$ would solve the Diffie-Hellman problem and an investigation of $P{-}\mathrm{dh}_g$ becomes important.

Moreover, for the case that $q = p$ is a prime, we consider

$$Q{-}\mathrm{idx}_g : G \to \mathbb{F}_p, \quad g^x \mapsto Q(x), \tag{6.2}$$

for a non-constant polynomial $Q(X) \in \mathbb{F}_p[X]$, where we assume that $\deg(Q)$ is small, say,

$$1 \leq \deg(Q) \leq \log(p).$$

In the following sections we will give lower bounds on the degree and the number of non-zero coefficients of the interpolation polynomial for the mappings $P{-}\mathrm{dh}_g$ and $Q{-}\mathrm{idx}_g$. For the special case $P(x) = x^2$ and $Q(x) = x$ we are in the situation of the mappings dh_g and idx_g, respectively. In these cases our bounds exactly match the ones obtained in [8] (see also [58]). Thus, our results can be seen as a generalization of earlier results to general univariate polynomials P and Q of higher degree.

In [27] the results of this chapter were extended to the case of bivariate polynomials $P(X, Y)$. More precisely, lower bounds on degree and the number of non-zero coefficients of the interpolation polynomial of the mapping $(g^x, g^y) \mapsto g^{P(x,y)}$ were obtained.

6.2 Some Results on Polynomials

The *sparsity* $\mathrm{spr}(f)$ (or *weight*) of a polynomial $f(X) \in \mathbb{F}_q[X]$ is the number of its nonzero coefficients.

We collect some obvious facts about the sparsity of a polynomial:

$$\mathrm{spr}(g + h) \leq \mathrm{spr}(g) + \mathrm{spr}(h) \tag{6.3}$$

$$\mathrm{spr}(gh) \leq \mathrm{spr}(g) \cdot \mathrm{spr}(h) \ . \tag{6.4}$$

The next lemma gives a relation between the number of zeros and the degree and sparsity of a polynomial over a finite field. The result for the degree is well known. Although the result for the sparsity can already be found in the literature ([57, Lemma 3.3] and [58, Lemma 2.5] for $T = q-1$ and [34, Lemma 1] for the general case), for completeness we give the proof.

Lemma 6.2.1. *Let $g \in \mathbb{F}_q$ be an element of order T and $f(X) \in \mathbb{F}_q[X]$ a nonzero polynomial of degree at most $T-1$ with at least b zeros of the form g^x with $0 \le x \le T-1$. Then for the degree of $f(X)$ we have*

$$\deg(f) \ge b,$$

and for the sparsity of $f(X)$ we have

$$\operatorname{spr}(f) \ge \frac{T}{T-b}.$$

Proof. Put $w = \operatorname{spr}(f)$, let $N \le T-b$ be the number of $0 \le x \le T-1$ with $f(g^x) \ne 0$ and S the number of pairs (y,i), $0 \le y \le T-1$, $0 \le i \le w-1$ with $f\left(g^{y+i}\right) \ne 0$. Since $g^x = g^{x+T}$ we have $S = wN$. Using properties of Vandermonde matrices we can verify that for every $0 \le y \le T-1$ there exists an $0 \le i \le w-1$ with $f\left(g^{y+i}\right) \ne 0$ and thus $w(T-b) \ge wN = S \ge T$. □

The following result motivated by Newton's interpolation formula is essential for the reduction algorithms and the proofs of the interpolation results. We make the convention that the degree of the zero polynomial is -1.

Lemma 6.2.2. *Let $\mathbb{D}$ be a commutative ring with identity 1. Let $B \ge 0$ be an integer and $P(X) \in \mathbb{D}[X]$ a polynomial of degree $D \ge B$ with leading coefficient a_D. Then we have*

$$\sum_{d=0}^{D-B} \binom{D-B}{d} (-1)^{D-B-d} P(X+d) = \frac{a_D D!}{B!} X^B + T_{B-1}(X),$$

where $T_{B-1}(X)$ is a polynomial of degree at most $B-1$.

Proof. Fix $B \ge 0$. For $D = B$ the result is trivial. For $D \ge B+1$ with the convention $\binom{D-1-B}{-1} = 0$ we have

$$\begin{aligned}
S &:= \sum_{d=0}^{D-B} \binom{D-B}{d} (-1)^{D-B-d} P(X+d) \\
&= \sum_{d=0}^{D-B} \left(\binom{D-1-B}{d} + \binom{D-1-B}{d-1} \right) (-1)^{D-B-d} P(X+d) \\
&= \sum_{d=0}^{D-1-B} \binom{D-1-B}{d} (-1)^{D-1-B-d} (P(X+1+d) - P(X+d)) \\
&= \sum_{d=0}^{D-1-B} \binom{D-1-B}{d} (-1)^{D-1-B-d} Q(X+d),
\end{aligned}$$

where $Q(X) := P(X+1) - P(X)$ has degree $D-1$ and leading coefficient $a_D D$. By induction we get

$$S = \frac{a_D D (D-1)!}{B!} X^B + T_{B-1}(X),$$

where $T_{B-1}(X)$ is a polynomial of degree at most $B-1$. □

In the case that $\mathbb{D} = \mathbb{K}$ is a field of positive characteristic a more general result can be proven.

Lemma 6.2.3. *Let $\mathbb{K}$ be a field of characteristic $p > 0$ and $P(X) \in \mathbb{K}[X]$ a polynomial of degree $D < p$. Let $0 \leq B < D$ be an integer, $I = \{a_0, a_1, \ldots, a_{D-B}\}$ a set of pairwise distinct elements from $\mathbb{K}$ of cardinality $|I| = D - B + 1$, and $b \in \mathbb{K}$. Then there are coefficients $c_i \in \mathbb{K}$, $0 \leq i \leq D - B$, with $\sum_{i=0}^{D-B} c_i = 0$ and*

$$\sum_{i=0}^{D-B} c_i P(X + a_i) = bX^B + T_{B-1}(X),$$

where $T_{B-1} \in \mathbb{K}[X]$ is a polynomial of degree $B - 1$.

Proof. Put $P(X) := \sum_{n=0}^{D} w_n X^n$ with $w_D \neq 0$. For any $a_i \in I$ it obviously holds true that

$$P(X + a_i) = \sum_{n=0}^{D} w_n \sum_{k=0}^{n} \binom{n}{k} a_i^{n-k} X^k.$$

Hence, for any $c_0, c_1, \ldots, c_{D-B} \in \mathbb{K}$ we have

$$\sum_{i=0}^{D-B} c_i P(X + a_i) = \sum_{k=0}^{D} d_k X^k,$$

where

$$d_k := \sum_{i=0}^{D-B} c_i \sum_{n=k}^{D} w_n \binom{n}{k} a_i^{n-k}, \quad 0 \leq k \leq D.$$

It leaves to show show the existence of a vector

$$c = (c_0, c_1, \ldots, c_{D-B})^\top \tag{6.5}$$

with $d_k = 0$ for $B + 1 \leq k \leq D$ and $d_B = b \neq 0$. Writing in a matrix form we get

$$E \cdot c = (0, 0, \ldots, 0, b)^\top, \tag{6.6}$$

where $E = (E_{k,i})$ with entries

$$E_{k,i} = \sum_{n=k}^{D} w_n \binom{n}{k} a_i^{n-k}, \quad B \leq k \leq D,\ 0 \leq i \leq D - B.$$

Note that since $D < p$, all occuring binomial coefficients are non-zero in $\mathbb{K}$. Therefore, after some simple algebraic transformations the set of linear equations (6.6) is equivalent to

$$E' \cdot c = \left(0, \ldots, 0, b/\left(w_D \binom{D}{B}\right)\right)^\top,$$

where E' is the Vandermonde Matrix of $I = \{a_0, a_1, \ldots, a_{D-B}\}$. Since

$$\det(E') = \prod_{\substack{a_i, a_j \in I \\ i > j}} (a_i - a_j) \neq 0,$$

equation (6.6) has a solution c. Clearly, since $D - B > 0$, the condition $\sum_{i=0}^{D-B} c_i = 0$ is satisfied. □

Another simple fact about Binomial coefficients (see [17, Chapter 5.3]) is the following

Lemma 6.2.4. *For integers $0 \leq c < H$,*

$$\sum_{d=0}^{H} \binom{H}{d} (-1)^{H-d} d^c = 0 \ .$$

6.3 Functions Related to the Diffie-Hellman Mapping

In this section we present results emphasizing the importance of analyzing the interpolation polynomials of $P{-}\mathrm{dh}_g$ from (6.1). More precisely, we show that a polynomial $f(X)$ that coincides with $P{-}\mathrm{dh}_g$ on some *fixed and known* points ξ can be used as an oracle to efficiently compute dh_g.

Theorem 6.3.1. *Let $0 \neq g \in \mathbb{F}_q$ be an element of order T, let $P(X) \in \mathbb{Z}_T[X]$ a polynomial of degree D with $2 \leq D \leq T - 1$ and leading coefficient a_D. Assume there exists a polynomial $f(X) \in \mathbb{F}_q[X]$ such that*

$$f\left(g^x\right) = g^{P(x)}, \quad x \in S,$$

for a set $S \subseteq \{N+1, \ldots, N+H\}$ of cardinality $|S| = H - s$ with $1 \leq H \leq T$. Then there exists a subset $R \subseteq S$ of cardinality $|R| \geq H - D + 2 - (D-1)s$ and a deterministic algorithm $\mathcal{A}$ that computes

$$\mathcal{A}(g^x) = g^{x^2},$$

for all $x \in R$, with

$$O\left(D^2 \log^2(q) \log(T) \sqrt{\gcd(a_D, T)}\right)$$

bit operations and $D - 1$ evaluations of $f(X)$.

Proof. First we assume T to be prime and we show later the generalization to composite T. Let R be the set of $x \in \{N+1, \ldots, N+H\}$ for which $x + i \in S$ for $0 \leq i \leq D - 2$. Then obviously

$$|R| \geq H - D + 2 - (D-1)s \ .$$

Let $x \in R$ be fixed. We may now describe the algorithm $\mathcal{A}$ on input g^x. It evaluates $f(X)$ in g^{x+d}, $0 \le d \le D-2$. Let η_d denote the function value of $f(g^{x+d})$:

$$\eta_d = f\left(g^{x+d}\right) = g^{P(x+d)}, \quad 0 \le d \le D-2 \ .$$

By Lemma 6.2.2 with $B = 2$ we get

$$\zeta := \prod_{d=0}^{D-2} \eta_d^{\binom{D-2}{d}(-1)^{D-d}} = g^{\sum_{d=0}^{D-2}\binom{D-2}{d}(-1)^{D-d}P(x+d)} = g^{ex^2+c_1x+c_0}$$

with some constants c_1 and c_0 and $e := a_D D!/2$. The value ζ can be computed by $\mathcal{A}$ with $O(D^2)$ additions in $\mathbb{Z}_T$ for determining recursively all binomial coefficients modulo T, $O(D)$ powers, inversions, and multiplications in $\mathbb{F}_q$, i. e.,

$$O\left(D\log(T)\max\left(D, \log^2(q)\right)\right)$$

bit operations (cf. [1, Chapters 5 and 6]). Next the algorithm $\mathcal{A}$ eliminates the linear and constant terms by computing

$$\xi := \zeta \cdot (g^x)^{-c_1} g^{-c_0} = g^{ex^2}.$$

Finally, it determines the unique root of

$$X^e - \xi, \tag{6.7}$$

i.e. $g^{x^2} = \xi^{e^{-1}}$, where e^{-1} denotes the (unique) inverse of e modulo (the prime) T, in $O(\log(T)\log^2(q))$ bit operations (cf. [1, Theorem 7.3.1]). This concludes the proof for the case of prime T.

Now we consider the general case of composite T. If we assume $\gcd(e,T) = 1$, then the proof is still valid. Without this restriction the solution of $X^e - \xi$ in (6.7) is no longer unique. In general, such an equation has either no or exactly $d = \gcd(e,T)$ different solutions. By construction we know that $X = g^{x^2}$ is a solution of (6.7), so our goal is to identify the correct solution out of a pool of d different solutions. Depending on the prime factorization of e, in many cases this can be done efficiently.

For $m|T$, let C_m denote the (unique) subgroup of order m contained in G, where $G \subseteq \mathbb{F}_q^*$ is the subgroup of order T generated by g. Let $e = \prod_{i=1}^{l} r_i^{v_i}$ be the factorization of e. Lets us for the moment assume that we know both, the factorizations of e and T. Then we can write T as $u\prod_{i=1}^{l} r_i^{s_i}$, where the product does not divide u and $v_i \le s_i$. We have the isomorphism

$$G \cong C_{r_1^{s_1}} \times \ldots \times C_{r_l^{s_l}} \times C_u.$$

Thus we can represent any element $x \in G$ as a pair $(x_{r_1^{s_1}}, \ldots, x_{r_l^{s_l}}, x_u) \in C_{r_1^{s_1}} \times \ldots \times C_{r_l^{s_l}} \times C_u$. The transformation is given by $x \mapsto (x^{T/r_1^{s_1}}, \ldots, x^{T/r_l^{s_l}}, x^{T/u})$

and $(x_{r_1^{s_1}}, \ldots, x_{r_l^{s_l}}, x_u) \mapsto x_{r_1^{s_1}}^{\alpha_1} \cdot \ldots \cdot x_{r_1^{s_1}}^{\alpha_1} \cdot x_u^{\beta}$, where $\sum \alpha_i r_i^{s_i} + \beta u = 1$. The coefficients α_i and β can be computed by the extended Euclidean Algorithm. Computing both mappings can be achieved in $O(l \log^2(q)) = O(\log(T) \log^2(q))$ bit operations [1]. The idea is to compute the eth root of ξ by computing it in each direct factor separately. The eth root of ξ in C_u is unique and can again be efficiently computed as mentioned in the proof for prime T above. Remember that our initial goal was to compute g^{x^2}. Instead of computing the $r_i^{v_i}$th root of ξ in $C_{r_i^{s_i}}$ we compute g^{x^2} in $C_{r_i^{s_i}}$ directly from g^x. This can be done by a "baby step-giant step" algorithm [54], see also [40]. For every prime factor $r_i^{s_i}$ of e, the running time is $O(\sqrt{r_i} \log^2(T))$ bit operations. The overall running time to compute the eth root is $O(\mathcal{R}(d) \log^2(q) + \log(T) \log^2(q))$ bit operations, where $\mathcal{R}(y)$ denotes the sum of the square roots of all prime factors of y with multiplicity one, i.e. $\mathcal{R}(d) = \sum_{i=1}^{l} \sqrt{r_i}$. Using $d = \gcd(e, T) | \gcd(a_D, T) \cdot D!$, we get

$$\mathcal{R}(d) \leq \mathcal{R}(\gcd(a_D, T)) \cdot \mathcal{R}(D!) \ .$$

The term $\mathcal{R}(D!)$ can be bounded by $\sum_{i=1}^{D} \sqrt{i} \leq D^{3/2}$ and the term $\mathcal{R}(\gcd(a_D, T))$ by $\sqrt{\gcd(a_D, T)}$.

It leaves to argue how we can determine the factorizations of e and T. Actually for our algorithm it is not important to know all prime factors of e and T, we only need to know which prime factors are shared by e and T. Since $e = a_D D!$ this can be done in $O((D + \sqrt{\gcd(a_D, T)}) \log^2 q)$ bit operations by factoring $\gcd(a_D, T)$ using, for instance, a Pollard's rho algorithm (see, e.g., [16, 40]). $\square$

Note that in [24] similar results were proven though the presented results in this work are more efficient in terms of running time of the reduction algorithm and number of evaluations of the function $f(X)$. The reason for the more efficient reduction is that in our setting the points x for which the polynomial $f(g^x)$ coincides with the function $g^{P(x)}$ are *known*. In [24], the function $f(X)$ is viewed as an *oracle* that produces the correct answers $g^{P(x)}$ for a certain fraction of all inputs, randomized over internal coin tosses. So, for fixed x it is not known if $f(g^x) = g^{P(x)}$ does hold true or not.

6.3.1 Interpolation

In this section we prove lower bounds on degree and sparsity of interpolation polynomials of the mappings $P\text{-}\mathrm{dh}_g$ of the form (6.1).

Theorem 6.3.2. *Let $0 \neq g \in \mathbb{F}_q$ be an element of order T, $P(X) \in \mathbb{Z}_T[X]$ a polynomial of degree D with $2 \leq D \leq T-1$, leading coefficient a_D, and $a_D D! \neq 0 \bmod T$. Let $f(X) \in \mathbb{F}_q[X]$ such that*

$$f(g^x) = g^{P(x)}, \quad x \in S,$$

for a set $S \subseteq \{N+1, \ldots, N+H\}$ of cardinality $|S| = H - s$ with $1 \leq H \leq T$. Then we have

$$\deg(f) \geq \max\left(\frac{H - (D+1)(s+1) + 1}{2^{D-1}}, \frac{H - D(s+1) + 1 - r}{2^{D-2}}\right),$$

where r denotes the least residue of $a_D D!$ modulo T. For the sparsity of $f(X)$ we have

$$\mathrm{spr}(f) \geq \max\left(\left(\frac{T}{2(T-H+(D+1)(s+1)-1)}\right)^{2^{1-D}}, \left(\frac{T}{2(T-H+D(s+1)-1)}\right)^{2^{2-D}} \right).$$

Proof. In the proof we will pursue the following strategy: first we construct a set R of high cardinality together with a non-zero polynomial $F(X)$ having (by construction) at least $|R|$ zeros. Since $F(X)$ is non-zero and we are working in a finite field, we can infer that $\deg(F) \geq |R|$ and $\mathrm{spr}(F) \geq T/(T-|R|)$ (Lemma 6.2.1). Furthermore, we show that $\deg(F)$ and $\deg(f)$, and $\mathrm{spr}(F)$ and $\mathrm{spr}(f)$, respectively, are related, i. e., that a lower bound on $\deg(F)$ implies a lower bound on $\deg(f)$ and a lower bound on $\mathrm{spr}(F)$ implies a lower bound on $\mathrm{spr}(f)$. We describe two slightly different constructions yielding results that complement each other. Details follow.

Construction of R_1. Let R_1 be the set of $x \in \{N+1, \ldots, N+H\}$ for which $x+i \in S$ for $0 \leq i \leq D$. It is easy to see that

$$|R_1| \geq H - D - (D+1)s.$$

By Lemma 6.2.2 with $B = 0$ we have

$$\prod_{d=0}^{D} f\left(g^{x+d}\right)^{\binom{D}{d}(-1)^{D-d}} = g^{\sum_{d=0}^{D}\binom{D}{d}(-1)^{D-d}P(x+d)} = g^{a_D D!}, \quad x \in R_1.$$

Construction of $F_1(X)$. Let M_{odd} and M_{even} denote the set of integers d between 0 and D such that $D-d$ is odd or even, resp. The polynomial

$$F_1(X) = \prod_{d \in M_{\text{even}}} f\left(g^d X\right)^{\binom{D}{d}} - g^{a_D D!} \prod_{d \in M_{\text{odd}}} f\left(g^d X\right)^{\binom{D}{d}}$$

has at least $|R_1|$ zeros, namely g^x with $x \in R_1$. Now we show that $F_1(X)$ is not identical to zero.
Let $f_l X^l$ be the leading term of $f(X)$. Then by Lemma 6.2.4 (applied with $c = 0$ and $c = 1$) we get for the leading coefficient lc of $F_1(X)$:

$$\begin{aligned} lc(F_1) &= \prod_{d \in M_{\text{odd}}} (f_l g^{dl})^{\binom{D}{d}} - g^{a_D D!} \prod_{d \in M_{\text{even}}} (f_l g^{dl})^{\binom{D}{d}} \\ &= f_l^{\sum_{d \in M_{\text{odd}}} \binom{D}{d}} (g^l)^{\sum_{d \in M_{\text{odd}}} d\binom{D}{d}} (1 - g^{a_D D!}) \ . \end{aligned}$$

Since $a_D D! \neq 0 \bmod T$, the leading coefficient of $F_1(X)$ is not zero and hence $F_1(X)$ is not identical to zero. This implies the bound $\deg(F_1) \geq |R_1|$. Now we have

$$\deg(F_1) = \sum_{d \in M_{\text{odd}}} \binom{D}{d} \deg(f) = \frac{1}{2} \sum_{d=0}^{D} \binom{D}{d} \deg(f) = 2^{D-1} \deg(f),$$

and hence

$$\deg(f) \geq \frac{|R_1|}{2^{D-1}}.$$

From equations (6.3) and (6.4) we get

$$\operatorname{spr}(F_1) \leq 2\operatorname{spr}(f)^{2^{D-1}}.$$

On the other hand, Lemma 6.2.1 yields

$$\operatorname{spr}(F_1) \geq \frac{T}{T - |R_1|}.$$

Construction of R_2. Now let R_2 be the set of $x \in \{N+1, \dots, N+H\}$ for which $x + i \in S$ for $0 \leq i \leq D-1$. We see that

$$|R_2| \geq H - D + 1 - Ds.$$

By Lemma 6.2.2 with $B = 1$ we have

$$\begin{aligned} \prod_{d=0}^{D-1} f\left(g^{x+d}\right)^{\binom{D-1}{d}(-1)^{D-1-d}} &= g^{\sum_{d=0}^{D-1} \binom{D-1}{d}(-1)^{D-1-d} P(x+d)} \\ &= g^{a_D D! x + b}, \quad x \in R_2, \end{aligned}$$

for some integer b.

Construction of $F_2(X)$. The polynomial

$$F_2(X) = \prod_{\substack{d=0 \\ D-1-d \text{ even}}}^{D-1} f\left(g^d X\right)^{\binom{D-1}{d}} - g^b X^r \prod_{\substack{d=0 \\ D-1-d \text{ odd}}}^{D-1} f\left(g^d X\right)^{\binom{D-1}{d}}$$

for $r := a_D D! \bmod T$ is non-zero and it has at least $|R_2|$ zeros, namely g^x with $x \in R_2$. We have

$$\deg(F_2) = 2^{D-2} \deg(f) + r$$

and thus

$$\deg(f) \geq (|R_2| - r)/2^{D-2}.$$

Moreover, we have

$$\operatorname{spr}(F_2) \leq 2\operatorname{spr}(f)^{2^{D-2}}$$

and

$$\operatorname{spr}(F_2) \geq \frac{T}{T - |R_2|},$$

from which the lower bound on $\operatorname{spr}(f)$ follows. □

Theorem 6.3.2 is non-trivial if the set S is dense enough on some interval, $|S| > H(1 - \frac{1}{D-1})$. Note that for the special case $P(x) = x^2$ the (second) bound of Theorem 6.3.2 gives us $\deg(f) \geq H - 2s - 3$. This is exactly the bound obtained in [8] (see also [58]).

6.4 Functions Related to the Discrete Logarithm

Now we prove similar results for the functions $Q-\mathrm{idx}_g$ of the form (6.2). We focus on the case where $q = p$ is a prime. For the general case, similar results can be obtained using techniques from [43, 65].

6.4.1 Reduction

Analogously to Section 6.3.1 we present reduction results that emphasize the importance of analyzing the interpolation polynomials of $Q-\mathrm{idx}_g$ from (6.2). Since now the polynomial $Q(X)$ is acting on the finite field $\mathbb{F}_p$, things are easier than in the Diffie-Hellman case. Details follow.

Let $Q(X) \in \mathbb{F}_p[X]$. Assume there exists a polynomial $f(X) \in \mathbb{F}_p[X]$ that evaluates to $Q(x)$ for all g^x, where x is in a fixed set S. Let g^x be given for fixed $x \in S$. The goal of the reduction algorithm is to compute x. But given the value $\xi := f(g^x) = Q(x)$, we can compute all possible roots of the equation

$$Q(X) - \xi = 0.$$

This can be done with a *randomized* algorithm using an expected number of

$$O(D \log^2(D) \log(\log(D)) \log^3(p))$$

bit-operations (see e. g. [1, 16]). One of the (at most D) roots must be x. Since we also know g^x, we can test which solution is the correct one. The overall time consumption is clearly dominated by computing all roots, i. e., by $O(D \log^2(D) \log(\log(D)) \log^3(p))$ bit operations. For $1 \le D \le 4$ explicit formulas for the roots of $Q(X)$ are known which provide fast deterministic reduction algorithms.

Obviously, Lemma 6.2.2 can be used to design a *deterministic* reduction algorithm that computes x from g^x for all x in a certain subset of S (provided that S is dense enough) using

$$O\left(D \log(p) \min\left(D, \log(p)\right)\right)$$

bit operations and $D - 1$ evaluations of $f(X)$.

Now the proof of the following reduction result is obvious.

Theorem 6.4.1. *Let $0 \ne g \in \mathbb{F}_p$ be an element of order T, $Q(X) \in \mathbb{F}_p[X]$ a polynomial of degree D with $1 \le D \le p - 1$, and $f(X) \in \mathbb{F}_p[X]$ such that*

$$f(g^x) = Q(x), \quad x \in S,$$

for a set $S \subseteq \{N + 1, \ldots, N + H\}$ of cardinality $|S| = H - s$ with $1 \le H \le T$. Then there exists a randomized algorithm $\mathcal{A}_1$ that computes

$$\mathcal{A}_1(g^x) = x, \quad x \in S,$$

in an expected number of

$$O\left(D\log^2(D)\log(\log(D))\log^3(p)\right)$$

bit operations and one evaluation of $f(X)$. *Moreover, there exist a subset* $R \subseteq S$ *of cardinality* $|R| \geq H - D + 1 - Ds$ *and a deterministic algorithm* $\mathcal{A}_2$ *that computes*

$$\mathcal{A}_2(g^x) = x, \quad x \in R,$$

with

$$O\left(D\min(D, \log(p))\log(p)\right)$$

bit operations and $D-1$ *evaluations of* $f(X)$.

6.4.2 Interpolation

In this section we prove lower bounds on degree and sparsity of interpolation polynomials of the functions $Q-\mathrm{idx}_g$.

Theorem 6.4.2. *Let* $0 \neq g \in \mathbb{F}_p$ *be an element of order* T, $Q(X) \in \mathbb{F}_p[X]$ *a polynomial of degree* D *with* $1 \leq D \leq p-1$, *and* $f(X) \in \mathbb{F}_p[X]$ *such that*

$$f\left(g^x\right) = Q(x), \quad x \in S,$$

for a set $S \subseteq \{N+1, \dots, N+H\}$ *of cardinality* $|S| = H - s$ *with* $1 \leq H \leq T$. *Then we have*

$$\deg(f) \geq H - (D+1)(s+1) + 1.$$

Furthermore,

$$\mathrm{spr}(f) \geq \frac{T}{T - H + (D+1)(s+1) - 1} - 1.$$

Proof. The proof follows along the lines of the proof of Theorem 6.3.2 instead that since now the polynomial $Q(X)$ is not longer in the exponent, the construction of $F_1(X)$ must be adapted. This leads to better bounds on degree and sparsity.
Let a_D denote the leading coefficient of the polynomial $Q(X)$. Let R_1 be defined as in the proof of Theorem 6.3.2, i.e. the set of elements x from S such that $x+i \in S$ for $0 \leq i \leq D$. Again, $|R_1| \geq H - D - (D+1)s$ and using Lemma 6.2.2 with $B = 0$ it can be easily verified that the polynomial

$$F_1(X) = \sum_{d=0}^{D} f\left(g^d X\right)\binom{D}{d}(-1)^{D-d} - a_D D!$$

has at least $|R_1|$ zeros (all elements g^x with $x \in R_1$). Since

$$\begin{aligned} F_1(0) &= f(0)\sum_{d=0}^{D}\binom{D}{d}(-1)^{D-d} - a_D D! \\ &= -a_D D! \neq 0, \end{aligned}$$

the polynomial $F_1(X)$ is non-zero and we conclude the bound on the degree by noting that

$$|R_1| \le \deg(F_1) \le \deg(f).$$

By construction, $F_1(X)$ contains $\mathrm{spr}(F_1) \le \mathrm{spr}(f) + 1$ monomials. Since we are dealing with $g^x \ne 0$, we may assume $\deg(f) \le T - 1$. Applying Lemma 6.2.1 we see that

$$\mathrm{spr}(f) + 1 \ge \frac{T}{T - |R_1|},$$

and the result follows. ☐

Theorem 6.4.2 is only nontrivial if the set S is almost consecutive. more precisely, we need

$$|S| \ge \left(1 - \frac{1}{D+1}\right) H + 1.$$

Now we prove a bound for more general but sufficiently large sets S.

Theorem 6.4.3. *Let $0 \ne g \in \mathbb{F}_p$ be an element of order T, $Q(X) \in \mathbb{F}_p[X]$ a polynomial of degree D with $1 \le D \le p - 1$, and $f(X) \in \mathbb{F}_p[X]$ such that*

$$f(g^x) = Q(x), \quad x \in S,$$

for a set $S \subseteq \{0, 1, \dots, T-1\}$ of cardinality at least $D + 2$. Then we have

$$\deg(f) \ge \frac{|S|!(T-1-D)!}{2^D(|S|-1-D)!(T-1)!}.$$

Proof. Let A be the set of all vectors $a = (a_1, a_2, \dots, a_D) \in \{1, 2, \dots, T-1\}^D$ with pairwise different coordinates satisfying $a_i = y_i - x \bmod T$ for some pairwise distinct elements $x, y_1, \dots, y_D \in S$. Then we have

$$|A| \le \frac{(T-1)!}{(T-1-D)!}$$

and there exists a vector $a \in A$ such that

$$R_a := \{x \in S : a_i + x \in S \text{ or } a_i + x - T \in S \text{ for } 1 \le i \le D\}$$

has at least

$$\frac{|S|!}{(|S|-1-D)!|A|} \ge \frac{|S|!(T-1-D)!}{(|S|-1-D)!(T-1)!}$$

elements. Now we have

$$f(g^{a_i+x}) = Q(a_i + x - \varepsilon_i(x) \cdot T)$$

with $\varepsilon_i(x) \in \{0, 1\}$, $i = 1, 2, \dots, D$, and at least one of the 2^D equation systems

$$f(g^{a_i+x}) = Q(a_i + x - \varepsilon_i \cdot T)$$

has at least $|R_a|/2^D$ solutions $x \in S$. We choose the ε_i from this system. For any $0 \neq b \in \mathbb{F}_p$, Lemma 6.2.3 provides coefficients $c_0, c_1, \ldots, c_D$ with $\sum_{i=0}^{D} c_i = 0$, such that

$$\sum_{i=0}^{D} c_i f(g^{a_i+x}) = \sum_{i=0}^{D} c_i Q(a_i + x - \varepsilon_i T) = b$$

for at least $|R_a|/2^D$ different $x \in S$ and the polynomial

$$F(X) = \sum_{i=0}^{D} c_i f(g^{a_i} X) - b$$

has at least $|R_a|/2^D$ zeros. Since

$$F(0) = f(0) \sum_{i=0}^{D} c_i - b = -b \neq 0,$$

we get

$$\deg(f) \geq \deg(F) \geq \frac{|R_a|}{2^D}$$

and the bound on the degree of $f(X)$ follows. □

Theorem 6.4.3 is only nontrivial if $|S|$ is at least of the order of magnitude $T^{1-\frac{1}{D+1}}$. Note that for the special case $Q(x) = x$ the bounds of Theorems 6.4.2 and 6.4.3 give us $\deg(f) \geq H - 2s - 1$ and $\deg(f) \geq \frac{|S|(|S|-1)}{2(T-1)}$, respectively. These bounds exactly match the bounds obtained in [8] (see also [58]).

Bibliography

[1] E. Bach and J. Shallit. *Algorithmic Number Theory, Volume I: Efficient Algorithms.* MIT Press, August 1996.

[2] N. Blum. A Boolean function requiring $3n$ network size. *Theoretical Computer Science*, 28(3):337–345, February 1984.

[3] J. Brenner and L. Cummings. The hadamard maximum determinant problem. *Amer. Math. Monthly*, 79:626–630, 1972.

[4] J. Bruck. Harmonic analysis of polynomial threshold functions. *SIAM Journal on Discrete Mathematics*, 3(2):168–177, May 1990.

[5] R. Canetti, J. B. Friedlander, S. V. Konyagin, M. Lars en, D. Lieman, and I. Shparlinski. On the statistical properties of Diffie-Hellman distributions. *Israel J. Math.*, 120:23–46, 2000.

[6] D. Catalano, R. Gennaro, N. Howgrave-Graham, and P. Q. Nguyen. Paillier's cryptosystem revisited. In *Proceedings of the 8th ACM Conference on Computer and Communications Security*, pages 206–214, 2001.

[7] D. Catalano, P. Q. Nguyen, and Jaques Stern. The hardness of Hensel lifting: The case of RSA and discrete logarithm. In *Advances in Cryptology – ASIACRYPT 2002*. LNCS, Springer-Verlag, 2002.

[8] D. Coppersmith and I. Shparlinski. On polynomial approximation of the Discrete Logarithm and the Diffie-Hellman mapping. *Journal of Cryptology*, 13(3):339–360, March 2000.

[9] W. Diffie and M. Hellman. New directions in cryptography. *IEEE Transactions on information Theory*, 22(6):644–654, 1976.

[10] J. Forster. A linear lower bound on the unbounded error probabilistic communication complexity. In *Proceedings of the Sithteenth Annual Conference on Computational Complexity*, pages 100–106. IEEE Computer Society, 2001.

[11] J. Forster. *personal communication*, 2002.

[12] J. Forster, M. Krause, S. V. Lokam, R. Mubarakzjanov, N. Schmitt, and H. U. Simon. Relations between communication complexity, linear arrangements, and computational complexity. *FSTTCS: Foundations of Software Technology and Theoretical Computer Science*, 21, 2001.

[13] J. B. Friedlander, S. V. Konyagin, and I. E. Shparlinski. Some doubly exponential sums over $\mathbb{Z}_m$. *Acta Arithm.*, 105:249–370, 2002.

[14] M. L. Furst, S. B. Saxe, and M. Sipser. Parity circuits and the polynomial time hierarchy. *Math. Systems Theory*, 17:13–27, 1984.

[15] T. El Gamal. A public key cryptosystem and a signature scheme based on discrete logarithms. *Advances in Cryptology—CRYPTO '84*, pages 10–18, 1984.

[16] J. von zur Gathen and J. Gerhard. *Modern computer algebra.* Cambridge University Press, New York, 1999.

[17] R. L. Graham, D. E. Knuth, and O. Patashnik. *Concrete Mathematics.* Addison-Wesley, 1989.

[18] A. Hajnal, W. Maass, P. Pudlak, M. Szegedy, and G. Turan. Threshold circuits of bounded depth. *Journal of Computer and System Sciences*, 46:129–1154, 1993.

[19] J. Hertz, A. Krogh, and R. G. Palmer. *Introduction to the Theory of Neural Computing.* Addison-Wesley, 1991.

[20] W. Hoeffding. Probability inequalities for sums of bounded random variables. *American Statistical Journal*, pages 13–30, March 1963.

[21] H. J. Hoover, M. N. Klawe, and N. J. Pippenger. Bounding fan-out in logical networks. *Journal of the ACM*, 13:13–18, 1984.

[22] R. A. Horn and C. R. Johnson. *Matrix analysis.* Cambridge Univ. Press, 1991.

[23] J J. Hadamard. Résolution d'une question relative aux déterminants. *Bull. Sci. Math.*, 17:240–246, 1893.

[24] E. Kiltz. A tool box of cryptographic functions related to the Diffie-Hellman Function. *Progress in Cryptology – INDOCRYPT 2001*, pages 339–349, 2001.

[25] E. Kiltz. On the representation of boolean predicates of the Diffie-Hellman function. In *Proc. of 20th International Symposium on Theoretical Aspects of Computer Science STACS*, pages 223–233, 2003.

[26] E. Kiltz and H. U. Simon. Complexity theoretic aspects of some cryptographic functions. *Proceedings of the Ninth International Computing and Combinatorics Conference, COCOON 2003, LNCS 2697*, pages 294–303, 2003.

[27] E. Kiltz and A. Winterhof. Lower bounds on weight and degree of bivariate polynomials related to the Diffie-Hellman mapping. *preprint*, 2003.

[28] E. Kiltz and A. Winterhof. Polynomial interpolation of cryptographic functions related to the Diffie-Hellman problem. *Proceedings of the International Workshop on Coding and Cryptography, WCC 2003*, 2003.

[29] D. E. Knuth. *The Art of Computer Programming, Vol. 2, Seminumerical Algorithms.* Addison-Wesley, Reading, MA, 2nd edition, 1981.

[30] N. M. Korobov. *Exponential Sums and their Applications.* Kluwer Academic Publishers, 1992.

[31] M. Krause. Geometric arguments yield better bounds for threshold circuits and distributed computing. *Theoretical Computer Science*, 156(1–2):99–117, 25 March 1996.

[32] M. Krause and I. Wegener. Circuit complexity. *Boolean Functions Vol. II*, to appear.

[33] E. Kushilevitz and N. Nisan. *Communication Complexity.* Cambridge University Press, Cambridge, 1997.

[34] T. Lange and A. Winterhof. Polynomial interpolation of the elliptic curve and XTR discrete logarithm. In *Proceedings of the 8th Annual International Computing and Combinatorics Conference (COCOON'02)*, pages 137–143, 2002.

[35] A. A. Razborov M. Goldmann, J. Håstad. Majority gates versus general weighted threshold gates. *Journal of Computational complexity*, 2:277–300, 1992.

[36] E. El Mahassni and I. Shparlinski. Polynomial representations of the Diffie-Hellman mapping. *Bull. Austral. Math. Soc.*, 63:467–473, 2001.

[37] W. Meidl and A. Winterhof. A polynomial representation of the Diffie-Hellman mapping. *Appl. Algebra Engrg. Comm. Comput. 3*, pages 313–318, 2002.

[38] G. C. Meletiou. Explicit form for the discrete logarithm over the field GF(p, k). *Arch. Math. (Brno) 29*, pages 25–28, 1993.

[39] G. C. Meletiou and G. L. Mullen. A note on discrete logarithms in finite fields. *Appl. Algebra Engrg. Comm. Comput. 3*, pages 75–78, 1992.

[40] A. J. Menezes, P. C. van Oorschot, and S. A. Vanstone. *Handbook of applied cryptography.* CRC Press, Boca Raton, 1997.

[41] G. L. Mullen and D. White. A polynomial representation for logarithms in gf(q). *Acta Arith. 47*, pages 255–261, 1986.

[42] H. Niederreiter. A short proof for explicit formulas for discrete logarithms in finite fields. *Appl. Algebra Engrg. Comm. Comput. 1*, pages 55–57, 1990.

[43] H. Niederreiter and A. Winterhof. Incomplete character sums and polynomial interpolation of the discrete logarithm. *Finite Fields Appl. 8*, pages 184–192, 2002.

[44] N. Nisan and A. Wigderson. On rank vs. communication complexity. In Shafi Goldwasser, editor, *Proceedings of the 35th Annual Symposium on Foundations of Computer Science*, pages 831–836, Los Alamitos, CA, USA, November 1994. IEEE Computer Society Press.

[45] P. Paillier. Public-key cryptosystems based on composite degree residuosity classes. *Proc. of EUROCRYPT'99*, pages 223–238, 1999.

[46] C. H. Papadimitriou. *Computational Complexity*. Addison Wesley, 1994.

[47] R. Paturi and J. Simon. Probabilistic communication complexity. *Journal of Computer and System Sciences*, 33(1):106–123, aug 1986.

[48] D. Pointcheval. New public key cryptosystems based on the dependent — RSA problems. *Lecture Notes in Computer Science*, 1592:239–254, 1999.

[49] K. Prachar. *Primzahlverteilung*. Springer-Verlag, Berlin, 1957.

[50] A. A. Razborov. Lower bounds for monotone complexity of boolean functions. *Amer. Math. Soc. Transl.*, 147:75–84, 1990.

[51] A. A. Razborov. On small depth threshold circuits. *in Proceedings of the SWAT 92, Lecture Notes in Computer Science*, 621:42–52, 1992.

[52] R. L. Rivest, A. Shamir, and L. Adleman. A method for obtaining digital signatures and public-key cryptosystems. *Communications of the ACM*, 21(2):120–126, 1978.

[53] R. Shaltiel. Towards proving strong direct product theorems. In *Proceedings of the Sithteenth Annual Conference on Computational Complexity*, pages 107–119. IEEE Computer Society, 2001.

[54] D. Shanks. Class number, a theory of factorization and genera. In *in Proc. Symp. Pure Math.*, volume 20, pages 415–440. AMS, 1971.

[55] C. E. Shannon. A synthesis of two-terminal switching circuits. *Bell Systems Technical Journal*, 28:59–98, 1949.

[56] I. Shparlinski. Communication complexity and Fourier coefficients of the Diffie-Hellman key. *Proc. the 4th Latin American Theoretical Informatics Conf., LNCS 1776*, pages 259–268, 2000.

[57] I. E. Shparlinski. *Number theoretic methods in cryptography. Complexity lower bounds.* Progress in Computer Science and Applied Logic, 17. Birkhäuser Verlag, 1999.

[58] I. E. Shparlinski. *Cryptographic Application of Analytic Number Theory.* Birkhäuser Verlag, 2003.

[59] K.-Y. Siu, V. Roychowdhury, and T. Kailath. *Discrete Neural Computation - A Theoretical Foundation.* Prentice Hall PTR, 1995.

[60] I. M. Vinogradov. *Elements of number theory.* Dover Publications., 1954.

[61] I. Wegener. *The complexity of Boolean functions.* Wiley-Teubner, 1987.

[62] A. Weil. On some exponential bounds. *Proc. Nat. Acad. Sci. USA*, 35(5):204–207, 1948.

[63] A. L. Wells. A polynomial form for logarithms modulo a prime. *IEEE Trans. Inform. Theory 30*, pages 845–846, 1984.

[64] A. Winterhof. A note on the interpolation of the Diffie-Hellman mapping. *Bull. Austral. Math. Soc. 64*, pages 475–477, 2001.

[65] A. Winterhof. Polynomial interpolation of the discrete logarithm. *Des. Codes Cryptogr. 25*, pages 63–72, 2002.