Fourier Transforms in Computer Science

Daniel Štefankovič

Submitted to the Department of Computer Science
in partial fulfillment of the requirements for the

Masters Degree

at the

University of Chicago

October 2000

Abstract

We survey proofs of five Theorems which have applications in the Theory of Computing. The common theme of the proofs is the use of various variants of harmonic analysis. The proofs of following theorems are included:

- Theorem of Linial, Mansour and Nisan on the concentration of the Fourier coefficients of AC^0 functions. [LMN93] (harmonic analysis over the finite group $\mathbb{Z}_2^n$)
- Theorem of Kahn, Kalai and Linial on the influence of variables on Boolean functions [KKL88] (harmonic analysis over the finite group $\mathbb{Z}_2^n$).
- The analysis of Margulis' expander graph by Gabber and Galil [GG79] (harmonic analysis over the compact group $\mathbb{T}^2$).
- Transference Theorem of Banaszczyk [Ban93] (harmonic analysis over the locally compact group $\mathbb{R}^n$).
- Theorem of Thérien on the column sums in matrices (mod m) [Thé94] (generalized harmonic analysis over the finite group $\mathbb{Z}_{p-1}^n$ with respect to the finite field $\mathbb{F}_p$).

Acknowledgement: I would like to thank my advisor László Babai for his support and suggestions and Divakar Viswanath for helpful discussions.

Contents

Introduction

Fourier analysis originated in the works of Euler and Fourier, who were working on problems in mathematical physics. The subject had a large impact on the development of mathematics. Dirichlet, in his work dealing with the convergence of Fourier series, defined the notion of function as we know it today. Riemann introduced his notion of integral in his work on trigonometric series. Cantor's study of the so-called sets of uniqueness led him to the development of the theory of sets. Today harmonic analysis is used in every branch of mathematics including group theory (where it was started by Frobenius), probability theory, combinatorics, differential equations, and number theory.

Methods of harmonic analysis also found their way to Computer Science. One of the most important applications in Computer Science is the Fast Fourier Transform (FFT) algorithm of Cooley and Tukey [CT65] and its application to the fast multiplication of numbers by Schönhage and Strassen [SS71]. Chung, Diaconis and Graham used the convolution $\leftrightarrow$ multiplication property of the Fourier transform to analyze random walks on graphs [CDG87]. Linial, Mansour and Nisan [LMN93] investigated the properties of the Fourier coefficients of functions computed by the AC^0 circuits and obtained result about the learnability of AC^0 functions. The KM learning algorithm of Kushilevitz and Mansour [KM93] is based on estimating the Fourier coefficients of the function learned. For more applications of harmonic analysis in learning theory see [Jac95, BFJ$^+$94]. Kahn, Kalai and Linial [KKL88] used the Fourier transform and Beckner's Lemmas [Bec75] to show that every balanced Boolean function has a variable with large influence. Thérien [Thé94] applied generalized harmonic analysis to study of circuits with MOD_m gates. Gabber and Galil successfully analyzed a modified construction of Margulis [Mar73] using harmonic analysis on the two dimensional torus $\mathbb{T}^2$. Banaszczyk showed a transference theorem in lattices using Fourier transforms of Gaussian-like measures in $\mathbb{R}^n$. J. Naor and M. Naor [NN93] used the Fourier transform to design polynomial size sample spaces of ε-biased $(\log n)$-wise independent random variables.

Chapter 1

Harmonic Analysis over Finite Abelian Groups

1.1 Introduction

Let A be a measure space with a non-negative measure μ and the corresponding Lebesgue integral. For complex valued functions $f, g \in \mathbb{C}^A$ we define

- **pointwise multiplication** $(fg)(x) = f(x)g(x),\ x \in A$,
- **inner product** $\langle f, g\rangle = \int f(y)\overline{g(y)}\, d\mu(y)$, where $\overline{x}$ is the complex conjugate of $x \in \mathbb{C}$,
- ***p*-norm** $||f||_p = \left(\int |f(y)|^p\, d\mu(y)\right)^{1/p}$,
- **convolution** $(f * g)(x) = \int f(y)g(x-y)\, d\mu(y)$.

Let G be a finite abelian group written additively. Let $n = |G|$. **Measure on *G*:** We endow G with the measure $\mu(g) = 1/n,\ g \in G$. We define the inner product, the p-norm and the convolution on the space $\mathbb{C}^G$ as above.

By $\mathbb{S}^1$ we denote the multiplicative group of complex numbers of modulus 1.

Definition 1.1.1 A **character** χ of G is a homomorphism $G \to \mathbb{S}^1$. The **unit character 1** is the character which assigns 1 to every $a \in G$.

Since G is finite a function $\chi : G \to \mathbb{C}$ is a character iff it satisfies

$$\chi(a+b) = \chi(a)\chi(b),\ a, b \in G.$$

Note that

- $\chi(a)^n = \chi(n \cdot a) = \chi(0) = 1$.
- $\chi(-a) = \overline{\chi(a)}$.

- If χ and ψ are characters then $\chi\psi$ and $\overline{\chi}$ are characters.

Lemma 1.1.2 *Characters are an orthonormal set of functions.*

Proof :
Note that for any character χ and any $a \in G$

$$\chi(a)\sum_{b\in G}\chi(b) = \sum_{b\in G}\chi(a+b) = \sum_{b\in G}\chi(b).$$

If $\chi \neq \mathbf{1}$ then there is an a such that $\chi(a) \neq 1$ and hence $\sum_{b\in G}\chi(b) = 0$. If characters ψ, χ are different then $\psi\overline{\chi} \neq \mathbf{1}$ and hence

$$\langle \psi, \chi \rangle = \frac{1}{n}\sum_{b\in G}(\psi\overline{\chi})(b) = 0.$$

Clearly $\langle \psi, \psi \rangle = 1$. ■

Characters with the pointwise multiplication form a group $\widehat{G}$, called the **dual group** of G. The unit character is the unit of $\widehat{G}$. Every finite abelian group is a direct sum of cyclic groups. Fix

$$G = \mathbb{Z}_{n_1} \oplus \cdots \oplus \mathbb{Z}_{n_k}.$$

Let $\omega_i = \exp(2\pi i/n_i)$, a primitive n_i-th root of unity. For $b = (b_1, \ldots, b_k) \in G$ let

$$\chi_b(x) = \prod_{i=1}^{k}\omega_i^{b_i x_i}.$$

Note that χ_b is a character for any $b \in G$. The χ_b are all distinct. They are all characters because the dimension of $\mathbb{C}^G$ is n and the characters are orthonormal. Thus the dual group of G is

$$\widehat{G} = \Big\{\chi_b \mid b = (b_1, \ldots, b_k) \in G, \Big\}. \tag{1.1}$$

Theorem 1.1.3 *Characters form an orthonormal basis of $\mathbb{C}^G$.*

From Theorem 1.1.3 it follows that every function f in $\mathbb{C}^G$ can be expressed as a linear combination of characters. The coefficient of χ_b is denoted by $\widehat{f}(\chi_b)$ and called the **Fourier coefficient**. We have

$$f = \sum_{b\in G}\widehat{f}(\chi_b)\chi_b. \tag{1.2}$$

The function $\widehat{f} : \widehat{G} \to \mathbb{C}$ is called the **Fourier transform** of f. **Measure on $\widehat{G}$**: We endow the group $\widehat{G}$ with the measure $\nu(g) = 1,\ g \in \widehat{G}$. The corresponding inner product of $f, g \in \mathbb{C}^{\widehat{G}}$ is denoted $\widehat{\langle} f, g \widehat{\rangle}$, the p-norm is denoted $\widehat{\|} f \widehat{\|}_p$ and the convolution of $f, g \in \mathbb{C}^{\widehat{G}}$ is denoted $f \mathbin{\hat{*}} g$.

Note that $\chi_a\chi_b = \chi_{a+b}$ and hence

$$b \to \chi_b \tag{1.3}$$

is an isomorphism between G and $\widehat{G}$. Hence we can view $\widehat{f}$ as a function in $\mathbb{C}^G$ and write $\widehat{f}(b)$ instead of $\widehat{f}(\chi_b)$.

From the orthogonality of characters it follows that

$$\widehat{f}(b) = \langle f, \chi_b \rangle. \tag{1.4}$$

Expanding (1.4) and using $\chi_a(b) = \chi_b(a)$ we obtain

$$\widehat{f} = \frac{1}{n}\sum_{a\in G} f(a)\overline{\chi_a}. \tag{1.5}$$

Theorem 1.1.4 *The Fourier transform satisfies*

- *linearity* $\widehat{f+g} = \widehat{f} + \widehat{g},\ \widehat{af} = a\widehat{f},\ f,g \in \mathbb{C}^G,\ a \in \mathbb{C},$
- $\widehat{fg} = \widehat{f} \hat{*} \widehat{g},\ \widehat{f*g} = \widehat{f}\widehat{g},\ f,g \in \mathbb{C}^G.$

Proof :
The linearity follows from the linearity of inner product. To prove $\widehat{fg} = \widehat{f} \hat{*} \widehat{g}$ note that

$$fg = \Big(\sum_{a\in G}\widehat{f}(a)\chi_a\Big)\Big(\sum_{b\in G}\widehat{g}(b)\chi_b\Big) = \sum_{a,b\in G}\chi_{a+b}\widehat{f}(a)\widehat{g}(b) = \sum_{c\in G}\chi_c\sum_{a\in G}\widehat{f}(a)\widehat{g}(c-a) = \sum_{c\in G}(\widehat{f}\hat{*}\widehat{g})(c)\chi_c$$

and hence $\widehat{fg} = \widehat{f} \hat{*} \widehat{g}$. The proof of $\widehat{f*g} = \widehat{f}\widehat{g}$ uses (1.5) instead of (1.2). ∎

An important property of the Fourier transform is the following formula.

Theorem 1.1.5 (Plancherel formula) *For any* $f, g \in \mathbb{C}^G$

$$\langle f, g\rangle = \hat{\langle}\widehat{f}, \widehat{g}\hat{\rangle}$$

Proof :
Using the orthogonality of characters

$$\langle f, g\rangle = \langle\sum_{a\in G}\widehat{f}(a)\chi_a, \sum_{b\in G}\widehat{g}(b)\chi_b\rangle = \sum_{a,b\in G}\widehat{f}(a)\overline{\widehat{g}(b)}\langle\chi_a,\chi_b\rangle = \sum_{a\in G}\widehat{f}(a)\overline{\widehat{g}(a)} = \hat{\langle}\widehat{f},\widehat{g}\hat{\rangle}.$$

∎

Taking $f = g$ in the Plancherel formula we obtain

Theorem 1.1.6 (Parseval's equality) *For any* $f \in \mathbb{C}^G$

$$||f||_2 = \hat{||}\widehat{f}\hat{||}_2.$$

To obtain inequalities for other p-norms we will need the Riesz-Thorin interpolation theorem, Minkowski's and Hölder's inequality (see [Zyg59], vol 2, p.95, p.94, vol 1, p.19)

Theorem 1.1.7 (The Riesz-Thorin interpolation theorem) *If T is a linear operator from a measure space A to a measure space B such that*

$$\begin{aligned} ||Tf||_{1/q_1} &\leq c_1||f||_{1/p_1} \\ ||Tf||_{1/q_2} &\leq c_2||f||_{1/p_2} \end{aligned}$$

where $0 \leq p_1, p_2, q_1, q_2 \leq 1$, then for any $t \in [0,1]$

$$||Tf||_{1/q} \leq c||f||_{1/p}$$

where $p = tp_1 + (1-t)p_2$, $q = tq_1 + (1-t)q_2$ and $c = c_1^t c_2^{1-t}$.

Theorem 1.1.8 (Hölder's inequality) *Let A be a measure space. Let $1/p + 1/q = 1,\ p, q \geq 1$. For any $f, g \in \mathbb{C}^A$*

$$||fg||_1 \leq ||f||_p||g||_q.$$

Theorem 1.1.9 (Minkowski's inequality) *Let A be a measure space. Let $p \geq 1$. For any $f, g \in \mathbb{C}^A$*

$$||f+g||_p \leq ||f||_p + ||g||_p.$$

An important generalization of Parseval's equality is the

Theorem 1.1.10 (Hausdorff-Young inequality) *Let $1/p + 1/q = 1,\ p, q \geq 1$ and $f \in \mathbb{C}^G$. Then*

$$\begin{aligned} ||f||_p &\leq \hat{|}|\widehat{f}\hat{|}|_q \quad \text{for } 2 \leq p \leq \infty \\ ||f||_p &\geq \hat{|}|\widehat{f}\hat{|}|_q \quad \text{for } 1 \leq p \leq 2 \end{aligned}$$

Proof :
Note that

$$|\widehat{f}(x)| = \left| \frac{1}{n} \sum_{a \in G} f(a)\chi_x(a) \right| \leq \frac{1}{n} \sum_{a \in G} |f(a)| = ||f||_1$$

and hence $\hat{|}|\widehat{f}\hat{|}|_\infty \leq ||f||_1$. From Parseval's equality we know $\hat{|}|\widehat{f}\hat{|}|_2 = ||f||_2$. An application of the Riesz-Thorin interpolation Theorem yields the Lemma for $2 \leq p \leq \infty$. For $1 \leq p \leq 2$ note that for $1/p + 1/q = 1$

$$||f||_2^2 = ||f||_p||f||_q.$$

■

The following inequalities for norms in $\mathbb{C}^G$ and $\mathbb{C}^{\widehat{G}}$ will be useful

Lemma 1.1.11 *Let $p \le r$. For any $f \in \mathbb{C}^G$ and $g \in \mathbb{C}^{\widehat{G}}$*

$$\begin{aligned} ||f||_p &\le ||f||_r, \\ \hat{||}g\hat{||}_p &\ge \hat{||}g\hat{||}_r. \end{aligned}$$

Proof :
Note that $||f||_p = ||\,|f|^p\,||_1^{1/p}$ and $||f||_r = ||\,|f|^p\,||_{r/p}^{1/p}$. Hence we can w.l.o.g. assume $p = 1$ and $r > 1$. Let q be the conjugate exponent to r i.e. $1/r + 1/q = 1$. By Hölder's inequality

$$||f||_p \le ||f||_r ||\mathbf{1}||_q = ||f||_r.$$

For the second inequality using the same argument we can assume $p = 1$ and $r > 1$. Also w.l.o.g. $||g||_1 = 1$. Hence $|g(x)| \le 1$ for all $x \in \widehat{G}$. Therefore $|g(x)|^r \le |g(x)|$ and $\hat{||}g\hat{||}_r \le 1$. ■

Theorem 1.1.12 (Young's convolution inequality) *Let $1 \le p, q, r \le \infty$, $1/r = 1/p + 1/q - 1$. For any $f, g \in \mathbb{C}^G$*

$$||f * g||_r \le ||f||_p\, ||g||_q.$$

Proof :
We have $||f * g||_1 \le ||f||_1 ||g||_1$ and from Hölder inequality $||f * g||_\infty \le ||f||_\infty ||g||_1$. Hence using Riesz-Thorin interpolation Theorem

$$||f * g||_p \le ||f||_p ||g||_1. \tag{1.6}$$

From Hölder inequality

$$||f * g||_\infty \le ||f||_p ||g||_q, \quad \text{for } 1/p + 1/q = 1. \tag{1.7}$$

Now using Riesz-Thorin interpolation Theorem on (1.6) and (1.7) we obtain the Lemma. ■

Given a subgroup $H \le G$, the characters for which $\chi|_H \equiv \mathbf{1}_H$ form a subgroup of $\widehat{G}$. Its corresponding (via (1.3)) subgroup in G is denoted $H^\perp$. Note that the characters which are 1 on H are in one-to-one correspondence with the characters of G/H (let $\chi(aH) = \chi(a)$). Thus we have $H^\perp \cong \widehat{G/H}$.

Theorem 1.1.13 (Poisson Summation Formula) *For a subgroup $H \le G$*

$$\frac{1}{|H|} \sum_{x \in H} f(x + a) = \sum_{y \in H^\perp} \widehat{f}(y) \chi_y(a).$$

Proof :

$$\frac{1}{|H|} \sum_{x \in H} f(x + a) = \frac{1}{|H|} \sum_{x \in H} \sum_{y \in G} \widehat{f}(y) \chi_y(x + a) = \sum_{y \in G} \widehat{f}(y) \chi_y(a) \frac{1}{|H|} \sum_{x \in H} \chi_y(x) = (*).$$

The restriction $\chi_y|_H$ is a character of H and hence by the orthogonality

$$(*) = \sum_{y \in G} \widehat{f}(y)\chi_y(a)(\chi_y|_H \equiv \mathbf{1}).$$

■

Finally we compute the Fourier coefficients of the function which has value 1 at 0 and value 0 elsewhere. For a set $A \subseteq G$ the characteristic function of A is denoted by $\mathbf{1}_A$,

$$\mathbf{1}_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{otherwise} \end{cases}$$

For a one element set $\{a\}$ we write $\mathbf{1}_a$ instead of $\mathbf{1}_{\{a\}}$.

Example 1.1.14 For $\mathbf{1}_0$ we have

$$\widehat{\mathbf{1}_0} = \frac{1}{n}\chi_x(0) = \frac{1}{n}$$

and hence

$$\mathbf{1}_0 = \frac{1}{n}\sum_{a \in G} \chi_a.$$

1.2 Representations of Boolean Functions

Let $B = \{0, 1\}$ where 0 represents false and 1 represents true. Given a Boolean function $f : B^n \to B$, we can view it as a function $f : \mathbb{Z}_2^n \to \{T, F\} \subseteq \mathbb{C}$. Usually we choose

- $F = 0,\ T = 1$; or
- $F = 1,\ T = -1$.

Let f_a be the $\{0, 1\}$-representation of f and f_b be the ± 1 representation of f. We have $f_b = 1 - 2f_a$ and hence

$$\widehat{f_b}(x) = \begin{cases} -2\widehat{f_a}(x) & \text{for } x \neq 0 \\ 1 - 2\widehat{f_a}(x) & \text{for } x = 0 \end{cases}.$$

Thus the choice of representation does not have a big influence on the Fourier transform. To simplify notation we will use the correspondence between $\mathbb{Z}_2^n$ and the subsets of $[n]$, $(a_1, \ldots, a_n) \leftrightarrow \{i; a_i = 1\}$. We define the **degree** of f as

$$\deg f = \min\{|x|; \widehat{f}(x) \neq 0\}.$$

If all $\widehat{f}(x) = 0$ we let $\deg f = 0$. Note that the degree is independent of the choice of the representation.

1.3 Random Restrictions and the Fourier Transform

Random restrictions have proven to be useful in the analysis of Boolean functions [FSS81, Hås86, LMN93]. In this section we will prove a relationship between the Fourier coefficients of a function and its random restriction. Lemma 1.3.6 will be used in section 1.4 to analyze the Fourier coefficients of the AC^0 functions.

Definition 1.3.1 Let $f : \mathbb{Z}_2^n \to \mathbb{C}$ be a function. Let $A \subseteq [n]$ be a subset of variables of f. Let α be an assignment of values to the variables in A. The **restriction** $f_{A\leftarrow\alpha}$ is the function obtained from f by assigning the values α to the variables in A.

Definition 1.3.2 Let $0 \le p \le 1$. A p-**random restriction** of $f : \mathbb{Z}_2^n \to \mathbb{C}$ is $f_{A\leftarrow\alpha}$ where

- A is random subset of $[n]$, each element is included in A independently with probability $1-p$,
- α is a random $\{0,1\}$-assignment to variables in A.

First we will look at the relation between the Fourier coefficients of a restriction of f and the Fourier coefficients of f.

Lemma 1.3.3 *Let $f : \mathbb{Z}_2^n \to \mathbb{C}$, $A \subseteq [n]$ and $x \subseteq \overline{A}$. Then*

$$\widehat{f_{A\leftarrow\alpha}}(x) = \sum_{y\subseteq A} \widehat{f}(x+y)\chi_\alpha(y).$$

Lemma 1.3.4 ([LMN93]) *Let $f : \mathbb{Z}_2^n \to \mathbb{C}$, $A \subseteq [n]$ and $x \subseteq \overline{A}$. Then we can calculate the following expected value and the second moment of $\widehat{f_{A\leftarrow\alpha}}$ where α is a random $\{0,1\}$-assignment to variables in A.*

$$E_\alpha[\widehat{f_{A\leftarrow\alpha}}(x)] = \widehat{f}(x),$$

$$E_\alpha[\widehat{f_{A\leftarrow\alpha}}(x)^2] = \sum_{y\subseteq A} \widehat{f}(x+y)^2.$$

Proof of Lemma 1.3.3:

$$\sum_{y\subseteq A} \widehat{f}(x+y)\chi_\alpha(y) = \sum_{y\subseteq A} \frac{1}{2^n} \sum_{z\subseteq[n]} f(z)\chi_z(x+y)\chi_\alpha(y) =$$

$$\frac{1}{2^{n-|A|}} \sum_{z\in[n]} f(z)\chi_z(x)\frac{1}{2^{|A|}}\sum_{y\subseteq A}\chi_y(z+\alpha) = \frac{1}{2^{n-|A|}} \sum_{z\subseteq[n];z\cap A=\alpha} f(z)\chi_z(x) = \widehat{f_{A\leftarrow\alpha}}(x)$$

■

Proof of Lemma 1.3.4:
Using Lemma 1.3.3

$$\frac{1}{2^{|A|}}\sum_{\alpha\subseteq A}\widehat{f_{A\leftarrow\alpha}}(x) = \frac{1}{2^{|A|}}\sum_{\alpha\subseteq A}\sum_{y\subseteq A}\widehat{f}(x+y)\chi_\alpha(y) = \sum_{y\subseteq A}\widehat{f}(x+y)\frac{1}{2^{|A|}}\sum_{\alpha\subseteq A}\chi_\alpha(y) = \widehat{f}(x).$$

Again using Lemma 1.3.3

$$\frac{1}{2^{|A|}}\sum_{\alpha\subseteq A}\widehat{f_{a\leftarrow\alpha}}(x)^2=\frac{1}{2^{|A|}}\sum_{\alpha\subseteq A}\sum_{y_1,y_2\subseteq A}\widehat{f}(x+y_1)\widehat{f}(x+y_2)\chi_\alpha(y_1+y_2)=$$
$$\sum_{y_1,y_2\subseteq A}\widehat{f}(x+y_1)\widehat{f}(x+y_2)\frac{1}{2^{|A|}}\sum_{\alpha\subseteq A}\chi_\alpha(y_1+y_2)=\sum_{y\subseteq A}\widehat{f}(x+y)^2.$$

■

Using the Lemma 1.3.4 we will obtain a relationship between the Fourier coefficients of a random restriction of f and the Fourier coefficients of f. For $x_i\in A$ the function $f_{A\leftarrow\alpha}$ is not a function of x_i. However we can view it as a function of x_i which does not depend on x_i. Then we have $\widehat{f_{A\leftarrow\alpha}}(y)=0$ if y contains x.

Lemma 1.3.5 *Let $f:\mathbb{Z}_2^n\to\mathbb{C}$. We can compute the expected value and the second moment of $f_{A\leftarrow\alpha}$ where $A\leftarrow\alpha$ is a p-random restriction.*

$$E_{A,\alpha}[\widehat{f_{A\leftarrow\alpha}}(x)]=p^{|x|}\widehat{f}(x),$$
$$E_{A,\alpha}[\widehat{f_{A\leftarrow\alpha}}(x)^2]=p^{|x|}\sum_{y\subseteq\overline{x}}\widehat{f}(x+y)^2(1-p)^{|y|}.$$

Proof :
If the set of set variables A intersects x then $\widehat{f_{A\leftarrow\alpha}}(x)=0$, otherwise by Lemma 1.3.4, $\widehat{f_{A\leftarrow\alpha}}(x)=\widehat{f}(x)$. The probability that no element of x is chosen to A is $p^{|x|}$ and hence $E_{A,\alpha}[\widehat{f_{A\leftarrow\alpha}}(x)]=p^{|x|}\widehat{f}(x)$.

Let $p(A)$ denote the probability that the set A is chosen in the random restriction $A\leftarrow\alpha$. Then

$$E_{A,\alpha}[\widehat{f_{A\leftarrow\alpha}}(x)^2]=\sum_{A\subseteq[n]}p(A)(x\cap A=0)\sum_{y\subseteq A}\widehat{f}(x\cup y)^2=$$
$$\sum_{y\subseteq[n]}\widehat{f}(x\cup y)^2\sum_{A\subseteq[n]}p(A)(x\cap A=0)(y\subseteq A)=\sum_{y\subseteq\overline{x}}\widehat{f}(x\cup y)^2p^{|x|}(1-p)^{|y|}.$$

■

Finally we express the sum of high Fourier coefficients of a random restriction of f using the Fourier coefficients of f. Let $B(n,p)$ denote a random variable with the binomial probability distribution with parameters n (the number of coin tosses) and p (the probability of heads).

Lemma 1.3.6 *Let $f:\mathbb{Z}_2^n\to\mathbb{C}$. Then*

$$E_{A,\alpha}\left[\sum_{|x|>k}\widehat{f_{A\leftarrow\alpha}}(x)^2\right]=\sum_{x\subseteq[n]}\widehat{f}(x)^2P\big(B(|x|,p)>k\big)$$

where $A\leftarrow\alpha$ is a p-random restriction.

Proof :
For fixed A and a random assignment α using Lemma 1.3.4

$$E_\alpha\left[\sum_{|x|>k}\widehat{f_{A\leftarrow\alpha}}(x)^2\right]=\sum_{|x|>k,x\subseteq\overline{A}}\ \sum_{y\subseteq A}\widehat{f}(x+y)^2=\sum_{|x\cap\overline{A}|>k}\widehat{f}(x)^2.$$

If we let A be random where each element is chosen independently with probability p then

$$E_{A,\alpha}\left[\sum_{|x|>k}\widehat{f_{A\leftarrow\alpha}}(x)^2\right]=\sum_{x\subseteq[n]}\widehat{f}(x)^2P(x\cap\overline{A}>k)=\sum_{x\subseteq[n]}\widehat{f}(x)^2P\big(B(|x|,p)>k\big).$$

■

1.4 AC^0 Circuits

Let Ω be a set of Boolean functions. An Ω-Boolean circuit with n inputs is a directed acyclic graph. It has vertices of two types: input nodes $x_1,\ldots,x_n$ and gates. One of the gates is the output gate. Each gate g is labeled by a function f_g from Ω. We define the function $B^n\to B$ computed by a gate g of the circuit inductively as $f_g(a_1,\ldots,a_k)$ where $a_1,\ldots,a_k$ are the functions computed by the predecessors of g. The function computed by the circuit is the function computed by the output gate. The size M of the circuit is the number of gates, not counting the input nodes. The depth d of the circuit is the length of longest path from an input node to the output gate.

Let $\{C_n\}$ be a sequence of circuits where C_n has n inputs. Let $\{f_n\}$ be the corresponding sequence of Boolean functions (f_n is the function computed by C_n). With the usual abuse of language we say that the circuit C computes the function f.

An AC^0 circuit consist of AND and OR gates of unbounded fan-in and NOT gates. It has polynomial size and constant depth. Using deMorgan's laws we can normalize the circuit so that it contains only n NOT gates connected directly to the inputs. The normalization at most doubles the size of the circuit.

It is well known that the AC^0 circuits cannot compute parity [Ajt83, FSS81, FSS84, Yao85]. They even cannot approximate parity [Ajt83]. Stronger results were shown by Håstad and Boppana (see [Hås86], p. 63).

A Fourier coefficient of a Boolean function $f:\mathbb{Z}_2^n\to\{0,1\}$ can be expressed as

$$\widehat{f}(y)=P(f(x)=\oplus_{i\in y}x_i)-P(f(x)\neq\oplus_{i\in y}x_i).$$

Thus $\widehat{f}(y)$ measures the correlation between f and the parity of variables x_i, $i\in y$. Since the functions computed by AC^0 circuits cannot approximate the parity it follows that each high Fourier coefficients of $f\in AC^0$ must be small.

Theorem of Linial, Mansour and Nisan [LMN93] shows that even the sum of squares of high Fourier coefficients must be small. The rest of this section is devoted to a proof of their result. The proof is a slightly modified version of the proof in [LMN93].

Theorem 1.4.1 *Let Boolean $f : \mathbb{Z}_2^n \to \{-1,1\}$ be computed by AC^0 circuit of depth d and size M. Then*

$$\sum_{|x|>t} \widehat{f}(x)^2 \leq 2M \cdot \exp\left(-\frac{1}{5e}(t/2)^{1/d}\right).$$

Remark 1 *The result proved in [LMN93] has slightly different bound than Theorem 1.4.1*

$$\sum_{|x|>t} \widehat{f}(x)^2 \leq 2M \cdot 2^{-(1/20)t^{1/d}}.$$

Theorem 1.4.1 has interesting applications [LMN93], for example functions computed by AC^0 circuits can be learned approximately by sampling their value at quasipolynomial ($2^{\text{polylog } n}$) randomly chosen inputs (chosen under the uniform distribution). Another corollary [LMN93] of Theorem 1.4.1 is that the average sensitivity of a function computed by an AC^0 circuit of depth d is $O((\log n)^d)$.

Let f be a Boolean function on n variables, let x be an assignment and let $f(x) = i$. There exists smallest $A \subseteq [n]$ such that $f_{A\leftarrow\alpha} \equiv i$ where $\alpha = (x \cap A)$. You can view $A \leftarrow \alpha$ as a proof that $f(x) = i$. It can happen that for every x there is small such proof.

Definition 1.4.2 Define the **non-deterministic** decision tree complexity of f

$$D_i(f) = \max_{x;f(x)=i} \min\{|A|; A \leftarrow \alpha, f_{A\leftarrow\alpha} \equiv i, x \cap A = \alpha\}.$$

Let $D_*(f) = \max\{D_0(f), D_1(f)\}$.

In [Nis91] $D_*(f)$ is called the certificate complexity. We say that f is computed by a t-CNF (resp. t-DNF) formula if f is computed by a CNF (resp. DNF) formula with clauses of size at most t. Let s_c (resp. s_d) be the smallest number such that f can be expressed as an s_c-CNF (resp. s_d-DNF). Then $D_0 = s_c$ and $D_1 = s_d$.

We say that in a DNF formula clauses accept disjoint inputs if for each input either 0 or exactly 1 clause is satisfied. We will use the following stronger version of Håstad's Switching Lemma (see [Hås86], p.65).

Theorem 1.4.3 (Håstad's Switching Lemma [Hås86]) *Let f be computed by a t-CNF formula and $A \leftarrow \alpha$ be a p-random restriction. With probability at least $1-(5pt)^s$, $f_{A\leftarrow\alpha}$ is computed by an s-DNF formula in which clauses accept disjoint inputs.*

If a function g is just a disjunction of s atoms, then $\widehat{g}(x) = 0$ for $|x| > s$ because g does not depend on at least one of the variables in x. A function f which is computed by an s-DNF formula in which clauses accept disjoint inputs can be viewed as a sum of functions which are disjunctions of s atoms and hence $\widehat{f}(x) = 0$ for $|x| > s$. Hence we have

Corollary 1.4.4 *Let*

- *$D_i(f) \le t$, and*
- *$A \leftarrow \alpha$ be a p-random restriction.*

Then $D_{1-i}(f_{A\leftarrow\alpha}) \le s$ and $\deg f \le s$ *with probability at least* $1-(5pt)^s$.

Using Corollary 1.4.4 we will obtain the following Lemma.

Lemma 1.4.5 *Let*

- *f be computed by AC^0 circuit of depth d and size M, and*
- *$A \leftarrow \alpha$ be a p-random restriction.*

Then for any $s \ge 1$ with probability at least $1 - M(5p^{1/d}s^{1-1/d})^s$

$$\deg(f_{A\leftarrow\alpha}) \le s.$$

Proof :
Let $f_1, \ldots, f_k$ be functions such that $D_*(f_i) \le t \le s,\ i \in [k]$. Observe that

$$D_0\Big(\bigwedge_{i\in[k]} f_i\Big) \le t \qquad\qquad D_1\Big(\bigvee_{i\in[k]} f_i\Big) \le t$$

because to prove that $\wedge f_i$ is zero (or that $\vee f_i$ is one) it is enough to prove that any of the f_i is zero (or that any of the f_i is one). If $A \leftarrow \alpha$ is a q-random restriction then by Håstad's Lemma

$$D_*\Big(\Big(\bigwedge_{i\in[k]} f_i\Big)_{A\leftarrow\alpha}\Big) \le s \qquad\qquad D_*\Big(\Big(\bigvee_{i\in[k]} f_i\Big)_{A\leftarrow\alpha}\Big) \le s \tag{1.8}$$

where each of the two events happens with probability at least $1-(5qt)^s$.

Let V_i denote the set of nodes of height i in the circuit. The functions computed by the leaves of the circuit have $D_* = 1$. By (1.8) after applying qs-random restriction, all functions computed by the nodes of height 1 will have $D_* \le s$ with probability at least $1 - |V_1|(5qs)^s$.

Now inductively, if the functions computed by the nodes of height i have $D_* \le s$ then after applying a q-random restriction, the functions computed by the nodes of height $i+1$ have $D_* \le s$ with probability at least $1 - |V_{i+1}|(5qs)^s$.

After applying a q-random restriction $d-2$ times we obtain that with probability at least $1 - (M-1)(5qs)^s$ the functions computed by the nodes of height $d-1$ have $D_* \le s$. After applying one more q-random restriction we obtain that the function computed by the circuit has $\deg \le s$ with probability at least $1 - M(5qs)^s$. Hence we obtained that for a function f computed by an $AC^0[d]$ circuit of size M and sq^d-random restriction $A \leftarrow \alpha$, $\deg f_{A\leftarrow\alpha} \le s$ with probability at least $1 - M(5qs)^s$. ■

We will use Chernoff's bound to estimate $B(n,p)$ (see [MR95], p. 235):

Theorem 1.4.6 (Chernoff's Bound) *Let $X_1, \dots, X_n$ be independent random variables. Let $X = X_1 + \cdots + X_n$ and $\mu = E[X]$. For $0 \leq \delta \leq 1$*

$$P\big(X \leq (1-\delta)\mu\big) \leq e^{-\mu\delta^2/2}.$$

Proof of Theorem 1.4.1
Let $p \in [0,1]$ and $s \geq 1$. By Lemma 1.3.6

$$\sum_{x\in[n]} \widehat{f}(x)^2 P\big(B(|x|,p) > s\big) = E_{A,\alpha}\left[\sum_{|x|>s^2} \widehat{f}_{A\leftarrow\alpha}(x)^2\right].$$

For a Boolean f, $||f||_2 = 1$ and hence by the Plancherel formula the random variable in the parenthesis has value at most 1. Moreover by Lemma 1.4.5 with probability at least $1 - M(5p^{1/d}s^{1-1/d})^s$ it has value 0. Thus

$$\sum_{x\in[n]} \widehat{f}(x)^2 P\big(B(|x|,p) > s\big) \leq M(5p^{1/d}s^{1-1/d})^s.$$

Since $P(B(t,p) > s)$ is an increasing function of t we obtain

$$\sum_{|x|>t} \widehat{f}(x)^2 \leq \frac{M(5p^{1/d}s^{1-1/d})^s}{P\big(B(t,p) > s\big)}.$$

Let $p = 2s/t$ and $s \geq 3$. By Chernoff's Bound

$$P\left(B(t,p) > \frac{tp}{2}\right) \geq 1 - e^{-tp/8} \geq 1/2.$$

For $s = \frac{1}{5e}(t/2)^{1/d}$ we obtain

$$\sum_{|x|>t} \widehat{f}(x)^2 \leq 2M \cdot \exp\left(-\frac{1}{5e}(t/2)^{1/d}\right).$$

■

1.5 The Influence of Variables

The **influence** of a variable x_i on a Boolean function f is the probability that for a random assignment $A \leftarrow \alpha$, $A = [n] \setminus \{i\}$ of values to the other variables the function $f_{A\leftarrow\alpha}$ is not constant. For the AND function every variable has a tiny influence 2^{1-n}. However if the function is more balanced e.g. if f is zero on a half of the inputs then there is a variable with a large influence $\Omega(\log n/n)$. This result is due to Kahn, Kalai and Linial [KKL88].

Theorem 1.5.1 *For a function $f : \mathbb{Z}_2^n \to \{0,1\}$ such that $P(f=1) = p \leq 1/2$*

$$\sum_{i \in [n]} I_i(f)^2 \geq c_1 p^2 \frac{(\log n)^2}{n}$$

in particular there is $i \in [n]$ such that

$$I_i(f) \geq c_2 p \frac{\log n}{n}$$

where c_1, c_2 are constants.

The theorem is tight up to the multiplicative constant, see [BL90]. It has been extended to functions $[0,1] \to \mathbb{C}$ in [BKK+92]. The rest of the section contains the proof of Theorem 1.5.1.

For any $1 \geq \varepsilon > 0$ define the linear operator $T_\varepsilon : \mathbb{C}^{\mathbb{Z}_2^n} \to \mathbb{C}^{\mathbb{Z}_2^n}$

$$T_\varepsilon(f) = \sum_{x \subseteq [n]} \widehat{f}(x) \varepsilon^{|x|} \chi_x. \tag{1.9}$$

In [Bec75] it was shown that T_ε is a norm 1 operator from $L^{1+\varepsilon^2}(\mathbb{Z}_2^n)$ to $L^2(\mathbb{Z}_2^n)$.

Lemma 1.5.2 *For any $f : \mathbb{Z}_2^n \to \mathbb{C}$*

$$||T_\varepsilon f||_2 \leq ||f||_{1+\varepsilon^2}.$$

We will prove Lemma 1.5.2 using Beckner's Lemmas in section 1.6.

Proof of Theorem 1.5.1
Let $f_i(x) = f(x) - f(x+i)$. Clearly for any p, $I_i(f) = ||f_i||_p^p$. We have

$$\widehat{f_i}(x) = \frac{1}{2^n} \sum_{y \subseteq [n]} \big(f(y) - f(y+i)\big)(-1)^{|x \cap y|} = \widehat{f}(x)\left(1 - (-1)^{i \in x}\right) = \begin{cases} 2\widehat{f}(x) & \text{if } i \in x \\ 0 & \text{otherwise} \end{cases}$$

Hence

$$T_\varepsilon f_i = \sum_{x \ni i} 2\widehat{f}(x) \varepsilon^{|x|} \chi_x.$$

By Plancherel's equality and Lemma 1.5.2

$$\sum_{x \ni i} 4\widehat{f}(x)^2 \varepsilon^{2|x|} = ||T_\varepsilon f_i||_2^2 \leq ||f_i||_{1+\varepsilon^2}^2 = I_i(f)^{2/(1+\varepsilon^2)}. \tag{1.10}$$

Summing (1.10) over $i \in [n]$ we obtain

$$4 \sum_{x \subseteq [n]} |x| \widehat{f}(x)^2 \varepsilon^{2|x|} \leq \sum_{i \in [n]} I_i(f)^{2/(1+\varepsilon^2)}. \tag{1.11}$$

The following equation is a linear combination of (1.11) with $\varepsilon^2 = 1$ and $\varepsilon^2 = 1/2$. For any $a \geq 0$ we have

$$\frac{1}{a}\sum_{i\in[n]} I_i(f) + \sum_{i\in[n]} I_i(f)^{4/3} \geq 4 \sum_{x\subseteq[n]} \left(\frac{1}{a} + 2^{-|x|}\right) |x| \widehat{f}(x)^2.$$

We have

$$\min_{x\geq 1} \left(\frac{1}{a} + 2^{-x}\right) x \geq \frac{\log a}{a}$$

and hence

$$4 \sum_{x\subseteq[n]} \left(\frac{1}{a} + 2^{-|x|}\right) |x| \widehat{f}(x)^2 \geq 4\frac{\log a}{a} \sum_{x\subseteq[n];x\neq\emptyset} \widehat{f}(x)^2.$$

We have $\sum_{x\subseteq[n];x\neq\emptyset} \widehat{f}(x)^2 = p - \widehat{f}(0)^2 = p - p^2 \geq p/2$ and hence

$$\frac{1}{a}\sum_{i\in[n]} I_i(f) + \sum_{i\in[n]} I_i(f)^{4/3} \geq 2p\frac{\log a}{a}.$$

Let $w = \sum_{i\in[n]} I_i(f)^2$. From Lemma 1.6 we have

$$\sum_{i\in[n]} I_i(f) \leq (wn)^{1/2} \quad \text{and} \quad \sum_{i\in[n]} I_i(f)^{4/3} \leq w^{2/3} n^{1/3}.$$

Hence

$$\frac{1}{a}(wn)^{1/2} + w^{2/3} n^{1/3} \geq 2\frac{\log a}{a} p.$$

For $a = n^{1/6}$ we obtain

$$w^{1/2} + w^{2/3} \geq \frac{1}{3}\frac{\log n}{n^{1/2}} p.$$

Since for $w \leq 1$, $w^{1/2} \geq w^{2/3}$ we have

$$w \geq \frac{1}{36}\frac{\log^2 n}{n} p^2.$$

■

1.6 Beckner's Lemmas

This section contains the proof of Lemma 1.5.2. The following Lemmas 1.6.2 and 1.6.4 needed in the proof of Lemma 1.5.2 were proven by Beckner [Bec75]. The proofs given here are simplified versions of Beckner proofs.

Lemma 1.6.1 *For $1 \le p \le 2$ and $\alpha \in \mathbb{R}$*

$$\frac{(1+\sin 2\alpha)^{p/2} + (1-\sin 2\alpha)^{p/2}}{2} \ge \big(1+(\sin\alpha)^2(p-2)\big)^{p/2}.$$

Proof :
Note that for $1 \le p \le 2$ and $-1 \le z \le 1$

$$\frac{(1+z)^p + (1-z)^p}{2} = \sum_{k=0}^{\infty} \binom{p}{2k} z^{2k} \ge 1 + \frac{p(p-1)}{2} z^2 \ge (1+z^2(p-1))^{p/2}. \tag{1.12}$$

The first inequality in (1.12) follows from the fact that the $\binom{p}{2k}$ are positive. The second one follows from $1+xt \ge (1+x)^t,\ 0 \le x, t \le 1$. Hence

$$\frac{(1+z^2+2z)^{p/2} + (1+z^2-2z)^{p/2}}{2} \ge \big(1+z^2(p-1)\big)^{p/2}.$$

Divide both sides by $(1+z^2)^{p/2}$ and let $\frac{z^2}{1+z^2} = (\sin\alpha)^2$. ■

Lemma 1.6.2 *For any $x, y \in \mathbb{C}$ and any $1 \le p \le 2$ we have*

$$\left(\frac{|x+y\sqrt{p-1}|^2 + |x-y\sqrt{p-1}|^2}{2}\right)^{1/2} \le \left(\frac{|x+y|^p + |x-y|^p}{2}\right)^{1/p}. \tag{1.13}$$

Proof :
We can scale inequality (1.13) so that $x\overline{x} + y\overline{y} = 1$. Let $a = x\overline{y} + \overline{x}y$ and $(\sin\alpha)^2 = y\overline{y}$. Note that $a^2 \le (\sin 2\alpha)^2$. The inequality becomes

$$\big(1+(\sin\alpha)^2(p-2)\big)^{p/2} \le \left(\frac{(1+a)^{p/2} + (1-a)^{p/2}}{2}\right). \tag{1.14}$$

The right hand side of (1.14) is a decreasing function of $|a|$ and hence it is enough to prove (1.14) for $|a| = \sin 2\alpha$. Now use Lemma 1.6.1. ■

For simplicity we will consider finite dimensional function spaces. Given a linear operator $T : \mathbb{C}^A \to \mathbb{C}^C$, let $K_T(x,c) = (T\mathbf{1}_x)(c)$, $K_T : A \times C \to \mathbb{C}$ is called the **kernel** of T. From the linearity of T

$$(Tf)(c) = \sum_{x \in A} K_T(x,c) f(x).$$

For any B and $f \in \mathbb{C}^{A\times B}$ we let $Tf : \mathbb{C}^{A\times B} \to \mathbb{C}^{C\times B}$ where we apply T to $a \to f(a,b)$ for each b separately

$$(Tf)(c,b) = \sum_{x\in A} K_T(x,c) f(x,b)$$

Given two operators $T_i : \mathbb{C}^{A_i} \to \mathbb{C}^{C_i},\ i = 1, 2$, their **product** is the operator $T : \mathbb{C}^{A_1\times A_2} \to \mathbb{C}^{C_1\times C_2}$, $T = T_1T_2$ where we apply T_2 for each $a \in A_1$ separately and then T_1 for each $d \in C_2$ separately. Note that the kernel of T is

$$K_T\big((a, b), (x, y)\big) = K_{T_1}(a, x)K_{T_2}(b, y).$$

The definition of the product of operators can be easily generalized for the product of more than two operators. Note the following relationship between the product and the composition of operators

$$(T_1 \circ S_1)(T_2 \circ S_2) = (T_1T_2) \circ (S_1S_2). \tag{1.15}$$

Given $\mathbb{C}^A$ with a p-norm and $\mathbb{C}^B$ with a q-norm and $f \in \mathbb{C}^{A\times B}$ let

$$||f(x, y)||_{x:p} = y \mapsto ||f(x, y)||_p$$

We define a p-norm on $\mathbb{C}^{A\times B}$

$$||f(x, y)||_p = ||\,||f(x, y)||_{x:p}\,||_{y:p} = ||\,||f(x, y)||_{y:p}\,||_{x:p}.$$

We will need generalized Minkowski's inequality, see [HLP34] p. 148.

Lemma 1.6.3 *For $1 \le p \le q$ and any $f : \mathbb{C}^{A\times B} \to \mathbb{C}$*

$$||\,||f(x, y)||_{x:p}\,||_{y:q} \le ||\,||f(x, y)||_{y:q}\,||_{x:p}.$$

Lemma 1.6.4 *Let $1 \le p \le q$. Let $T_i : \mathbb{C}^{A_i} \to \mathbb{C}^{C_i},\ i = 1, 2$ be linear operators. If for any $f_i \in \mathbb{C}^{A_i}$*

$$||T_if_i||_q \le ||f_i||_p$$

then for $T = T_1T_2$ and any $f \in \mathbb{C}^{A_1\times A_2}$

$$||Tf||_q \le ||f||_p.$$

Proof :
Since $1 \le p \le q$, we can use Lemma 1.6.3

$$||(Tf)(c, d)||_q = ||\,||T_1(T_2f)(c, d)||_{c:q}\,||_{d:q} \le ||\,||(T_2f)(a, d)||_{a:p}\,||_{d:q} \le$$
$$||\,||(T_2f)(a, d)||_{d:q}\,||_{a:p} \le ||\,||f(a, b)||_{b:p}\,||_{a:p} = ||f||_p.$$

■

Lemma 1.6.4 can be easily extended for the product of more than 2 operators. Now we can finally prove Lemma 1.5.2.

Proof of Lemma 1.5.2
Take the operator $T_{n,\varepsilon} : (\mathbb{Z}_2^n \to \mathbb{C}) \to (\mathbb{Z}_2^n \to \mathbb{C})$ defined by (1.9). Take any function $f : \mathbb{Z}_2 \to \mathbb{C}$ and

let $x = \widehat{f}(0)$, $y = \widehat{f}(1)$. Then $f(0) = x + y$, $f(1) = x - y$ and $(T_{1,\varepsilon} f)(0) = x + \varepsilon y$, $(T_{1,\varepsilon} f)(1) = x - \varepsilon y$. Taking $p = 1 + \varepsilon^2$ in Lemma 1.6.2 we obtain that $T_{1,\varepsilon} : L^{1+\varepsilon^2}(\mathbb{Z}_2) \to L^2(\mathbb{Z}_2)$ is a norm 1 operator.

Now we will show that $T_{n,\varepsilon}$ is product of n copies of $T_{1,\varepsilon}$ and hence by Lemma 1.6.4 has norm 1 as an operator from $L^{1+\varepsilon^2}(\mathbb{Z}_2^n)$ to $L^2(\mathbb{Z}_2^n)$.

Let $F_n : \mathbb{C}^{\mathbb{Z}_2^n} \to \mathbb{C}^{\widehat{\mathbb{Z}}_2^n}$ be the Fourier transform operator. Let $U_n : \mathbb{C}^{\widehat{\mathbb{Z}}_2^n} \to \mathbb{C}^{\widehat{\mathbb{Z}}_2^n}$ have kernel $K_U(x, c) = (x = c)\varepsilon^{|x|}$. Clearly F_n is the product of n copies of F_1 and U_n is the product of n copies of U_1. Now since $T_n = F_n^{-1} \circ U_n \circ F_n$ we have using (1.15) that T_n is the product of n copies of T_1. ■

Chapter 2

Harmonic Analysis over Locally Compact Abelian Groups

2.1 Introduction

We will only use harmonic analysis on $\mathbb{R}^n$, $\mathbb{Z}^n$ and $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$. It might be useful to mention a more general setting. A **locally compact abelian (LCA)** group is a locally compact Hausdorff topological space with group operations which are continuous i.e. the mappings $+ : G \times G \to G$ (where $G \times G$ has Cartesian product topology) and $- : G \to G$ are continuous.

On such a group there always exists unique (up to multiplicative factor) Borel measure μ invariant under the group operations i.e. for every measurable set E and $x \in G$, $\mu(E) = \mu(E + x)$. The measure is called the **Haar measure**. The measure of a compact set is finite and the measure of an open set is positive. The integral corresponding to the Haar measure is called the Haar integral.

A **character** of G is a continuous homomorphisms from G to the multiplicative group of complex numbers of modulus 1. The group $\widehat{G}$ of all characters of G with topology generated by

$$U_{C,\varepsilon} = \{\chi \in \widehat{G}; \forall x \in C : |\chi(x) - 1| < \varepsilon, \},\ \ C - \text{compact}, \varepsilon > 0$$

(and their translates) is a locally compact abelian group called the **dual group** of G. Pontryagin's **duality theorem** says that the dual of $\widehat{G}$ is isomorphic (as a topological group) to G. Moreover if G is discrete then $\widehat{G}$ is compact and vice versa.

The **Fourier Transform** of $f \in L^1(G)$ is defined as

$$\widehat{f}(\chi) = \int_G f(g)\overline{\chi(g)}\, dg.$$

If f is continuous and $\widehat{f} \in L^1(\widehat{G})$ then we have the **inversion formula**

$$f(x) = \int_{\widehat{G}} \widehat{f}(x)\chi(x)\, d\chi, \quad \text{for all } x \in G$$

where $d\chi$ is an appropriate normalization of the Haar measure on $\widehat{G}$.

If $f \in L^1(G) \cap L^2(G)$ then $\widehat{f} \in L^2(\widehat{G})$ and we have the **Plancherel formula**

$$||f||_2 = ||\widehat{f}||_2$$

Remark 2 If G is compact we usually normalize the Haar measure on G so that $\mu(G) = 1$. This normalization makes the characters orthonormal. The normalization of the Haar measure on $\widehat{G}$ (which is discrete) for which the inversion formula holds is $\mu(x) = 1,\ x \in G$.

Note that the Haar measures defined for a finite group G in section 1 correspond to viewing G as a compact group and $\widehat{G}$ as a discrete group.

2.2 Fourier transform over $\mathbb{T}^n$

Let $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$ with the standard Lebesgue measure and integral. Two functions $f, g : \mathbb{T}^n \to \mathbb{C}$ are equivalent $f \sim g$ if they differ in a set of points of measure 0.

Let $L^2(\mathbb{T}^n)$ be the space of measurable functions $\mathbb{T}^n \to \mathbb{C}$ such that $\int_{\mathbb{T}^n} |f(x)|^2\, d\mu < \infty$ factored by the equivalence relation $\sim$. The space $L^2(\mathbb{T}^n)$ with inner product

$$\langle f, g\rangle = \int_T f(t)\overline{g(t)}\, d\mu$$

is a Hilbert space.

Let $L^2(\mathbb{Z}^n)$ be the space of functions $f : \mathbb{Z}^n \to \mathbb{C}$ such that $\sum_{z\in\mathbb{Z}^n} |f(z)|^2 < \infty$. The space $L^2(\mathbb{Z}^n)$ with inner product

$$\langle f, g\rangle = \sum_{z\in\mathbb{Z}^n} f(z)\overline{g(z)}$$

is a Hilbert space.

Theorem 2.2.1 *Inner product and metric satisfy*

- $|(f, g)| \leq ||f||_2 \cdot ||g||_2$ *(the Cauchy-Schwarz inequality)*
- $||f + g||_2 \leq ||f||_2 + ||g||_2$ *(triangle inequality)*

Theorem 2.2.2 *The characters of* $\mathbb{T}^n$ *are*

$$\{\chi_z(x) = \exp(2\pi i z^T x); z \in \mathbb{Z}^n\}.$$

They form an orthonormal basis of $L^2(\mathbb{T}^n)$ *i.e. any function* $f \in L^2(\mathbb{T}^n)$ *can be written as*

$$f = \sum_{z\in\mathbb{Z}^n} \widehat{f}(z)\chi_z.$$

The $\widehat{f}(z) \in \mathbb{C}$ are called the Fourier coefficients, $\widehat{f}(z) = \langle f, \chi_z\rangle$.

Theorem 2.2.3 (Riesz-Fischer) *The map* $L^2(\mathbb{T}^n) \xrightarrow{\wedge} L^2(\mathbb{Z}^n)$ *is linear, bijective map.*

Moreover we have

- $\langle f, g\rangle = \langle \widehat{f}, \widehat{g}\rangle$ (Parseval's identity)
- $||f||_2 = ||\widehat{f}||_2$ (the Plancherel formula)
- $||\widehat{f}||_\infty \leq ||f||_1$

and hence the Fourier transform is an isometry of $L^2(\mathbb{T}^n)$ and $L^2(\mathbb{Z}_n)$.

A nice application of the Plancherel formula is computing the sum

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

Consider $f(x) = x$ on $\mathbb{T}^1$. Then $||f||_2^2 = 1/3$, $\widehat{f}(0) = 1/2$ and $\widehat{f}(n) = i/(2\pi n)$ for $n \neq 0$. Now use the Plancherel formula.

2.3 Expander Graphs Construction

The first expander graph construction is due to Margulis [Mar73]. We are going to show an expander graph construction due to Gabber and Galil [GG79]. The construction can be analyzed without harmonic analysis, using linear algebra (eigenvalues) [JM87]. Expanders with much better expansion have been constructed in [LPS88, Mar88].

Definition 2.3.1 A bipartite graph $G = (A \cup B, E)$ is (n, d, α) **expander** if $|A| = |B| = n$, $|E| \leq nd$ and for every $X \subseteq A$

$$N(X) \geq \left(1 + \alpha \frac{|\overline{X}|}{n}\right) |X|.$$

Explicit expanders of bounded degree have a great number of applications in the Theory of Computing. They are used in the $\log n$-depth sorting network of Ajtai, Komlós and Szemerédi [AKS83], in extraction of random bits from weak random sources [AKS87, CW89, IZ89], or in explicit construction of fault tolerant networks [AC88].

Definition 2.3.2 Let G_m be graph with $A, B = \mathbb{Z}_m^2$ and $(x, y) \in A$ connected to $\{\sigma_i(x, y); i \in \{0, ..., 4\}\}$ where

$$\begin{aligned}
&\sigma_0(x, y) = (x, y) \\
&\sigma_1(x, y) = (x + y, y), \quad \sigma_2(x, y) = (x + y + 1, y) \\
&\sigma_3(x, y) = (x, x + y), \quad \sigma_4(x, y) = (x, x + y + 1)
\end{aligned}$$

Theorem 2.3.3 *G_m is an $(m^2, 5, (2 - \sqrt{3})/4)$ expander.*

Proof :
We need to show for any $X \subseteq \mathbb{Z}_m^2$

$$|\sigma_1(X) \cup \cdots \cup \sigma_4(X) - X| \geq \alpha \frac{|X| \cdot |\overline{X}|}{m^2}.$$

There is a natural mapping between the subsets of $\mathbb{Z}_m^2$ and the subsets of the 2-dimensional torus with sides of length m (i.e. $\mathbb{R}^2/m\mathbb{Z}^2$), where (i,j) corresponds to $\Box(i,j) = \langle i, i+1) \times \langle j, j+1)$.

Let τ_1 be a linear transformation on $\mathbb{R}^2/m\mathbb{Z}^2$, $\tau_1(x,y) = (x+y, y)$. Under τ_1 the $\Box(i,j)$ is transformed to a parallelogram one half of which lies in $\Box(i+j,j)$ and the other in $\Box(i+j+1,j)$. Hence the nonempty squares in $\tau_1(X') - X'$ correspond to the elements in $(\sigma_1(X) \cup \sigma_2(X)) - X$ where $X' \subseteq \mathbb{R}^2/\mathbb{Z}_m^2$ corresponds to $X \subseteq \mathbb{Z}_m^2$. Thus

$$|\sigma_1(X) \cup \sigma_2(X) - X| \geq \mu(\tau_1(X') - X').$$

Similarly for $\tau_2(x,y) = (x, x+y)$, $|\sigma_3(X) \cup \sigma_4(X) - X| \geq \mu(\tau_2(X') - X')$. Hence

$$|\sigma_1(X) \cup \cdots \cup \sigma_4(X) - X| \geq \max\{|\sigma_1(X) \cup \sigma_2(X) - X|, |\sigma_3(X) \cup \sigma_4(X) - X|\} \geq \frac{1}{2}\big(\mu(\tau_1(X') - X') + \mu(\tau_2(X') - X')\big)$$

and to show the expansion property of G_m it is enough to show

$$\mu(\tau_1(X') - X') + \mu(\tau_2(X') - X') \geq 2\alpha\mu(X')\mu(\overline{X'})/m^2. \tag{2.1}$$

We will show that (2.1) is true not only for X' arising from the mapping between $\mathbb{Z}_m^2$ and $\mathbb{R}^2/m\mathbb{Z}^2$ but for all measurable X'. Scaling down to standard torus $\mathbb{S}^2 = \mathbb{R}^2/\mathbb{Z}^2$ inequality (2.1) becomes

$$\mu(\tau_1(X) - X) + \mu(\tau_2(X) - X)) \geq 2\alpha\mu(X)(1 - \mu(X)). \tag{2.2}$$

From now on we are working on the standard torus $\mathbb{S}^2 = \mathbb{R}^2/\mathbb{Z}^2$. Note that the τ_i are measure preserving and hence $\mu(\tau_i(X) - X) = \mu(X) - \mu(\tau_i(X) \cap X)$. Now (2.2) becomes

$$\mu(\tau_1(X) \cap X) + \mu(\tau_2(X) \cap X) \leq 2\mu(X)(1 - \alpha\mu(\overline{X})). \tag{2.3}$$

Now we rephrase (2.3) in terms of functions

$$\langle \mathbf{1}_X, (T_1+T_2)\mathbf{1}_X\rangle = \langle \mathbf{1}_X, T_1\mathbf{1}_X\rangle + \langle \mathbf{1}_X, T_2\mathbf{1}_X\rangle \leq 2\langle \mathbf{1}_X, \mathbf{1}_X\rangle(1 - \alpha + \alpha\langle \mathbf{1}_X, \mathbf{1}_X\rangle) \tag{2.4}$$

where $T_i f = f \circ \tau_i^{-1}$ is a linear operator. Now it looks that we need to estimate the Rayleigh quotient of the operator $T_1 + T_2$ on some class of functions. Theorem 2.4.4 shows that on the space of L^2 functions for which $\int f\, d^2x = 0$ the operator $T_1 + T_2$ has Rayleigh quotient $\beta = \sqrt{2} + 1/2$. Let us subtract a constant function from $\mathbf{1}_X$ so that we can apply Lemma 2.4.4

$$\langle \mathbf{1}_X - \mu(X), (T_1+T_2)(\mathbf{1}_X - \mu(X))\rangle \leq \beta\langle \mathbf{1}_X - \mu(X), \mathbf{1}_X - \mu(X)\rangle. \tag{2.5}$$

Using $\langle \mu(X), \mu(X)\rangle = \mu(X)^2$, $\langle \mu(X), \mathbf{1}_X\rangle = \mu(X)^2$ and $\langle \mathbf{1}_X, \mathbf{1}_X\rangle = \mu(X)$ on (2.5) we obtain

$$\langle \mathbf{1}_X, (T_1 + T_2)\mathbf{1}_X\rangle \leq 2\mu(X)(\beta/2 + (1-\beta/2)\mu(X))$$

which proves (2.4) for $\alpha = 1 - \beta/2 = (3-\sqrt{2})/4 < (4-2\sqrt{3})/4$.

To fix the proof note that if for every measurable X,

$$\mu(\tau_1^2(X) - X) + \mu(\tau_2^2(X) - X)) \geq 4\alpha\mu(X)\mu(\overline{X}) \tag{2.6}$$

then also (2.2). The proof of (2.6) follows exactly the same steps as the proof of (2.2), using an estimate on the Rayleigh quotient of $T' = T_1 \circ T_1 + T_2 \circ T_2$. ■

2.4 The Rayleigh quotient of Operators

In this section we compute the Rayleigh quotient of the operator $T_1 + T_2$ and $T_1 \circ T_1 + T_2 \circ T_2$ from the section 2.3 following the proof of Gabber and Galil [GG79].

Given a bounded linear operator T its **spectral norm** is

$$||T||_2 = \sup\Big\{||Tx||_2\,;\,||x||_2 = 1\Big\},$$

its **spectral radius** is

$$\rho(T) = \lim_{n\to\infty} ||T^n||_2^{1/n},$$

and its **Rayleigh quotient** is

$$r(T) = \sup\Big\{|\langle Tx, x\rangle|\,;\,||x||_2 = 1\Big\}.$$

The Cauchy-Schwarz inequality yields

$$\rho(T) \leq r(T) \leq ||T||_2.$$

If T is self-adjoint then $\rho(T) = r(T) = ||T||_2$.

The linear operator T on $L^2(\mathbb{T}^n)$ gives us a linear operator $\widehat{T} = F \circ T \circ F^{-1}$ on $L^2(\mathbb{Z}^n)$ where F is the Fourier transform operator. By Parseval's identity the operators $\widehat{T}$ and T have the same spectral norm, radius and Rayleigh quotient. In our application it will turn out to be easier to analyze the operator $\widehat{T}$.

Lemma 2.4.1 *Let A be an integral $n \times n$ matrix with determinant 1 and $(Tf)(x) = f(Ax)$ a linear operator on $L^2(\mathbb{T}^n)$. Then*

$$(\widehat{T}\widehat{f})(x) = \widehat{f}(A^{-T}x). \tag{2.7}$$

Proof :
Since $\widehat{T}$ is linear, it is enough to see (2.7) for $\widehat{f} = \mathbf{1}_a$

$$(\widehat{T}\mathbf{1}_a)(x) = \int_{\mathbb{T}^n} e^{-2\pi i x^T z} T(e^{2\pi i a^T z})\, d^n z = \int_{\mathbb{T}^n} e^{2\pi i (a^T A - x^T) z}\, d^n z = (x = A^T a) = \mathbf{1}_a(A^{-T}x).$$

■

Let X be the subspace of $L^2(\mathbb{T}^2)$ containing the functions satisfying $\widehat{f}(0) = \int f\, d^2x = 0$. Our goal now is to compute the Rayleigh quotient of the operator $T = T_1 + T_2$ (defined in (2.4)) on the space X. Recall that

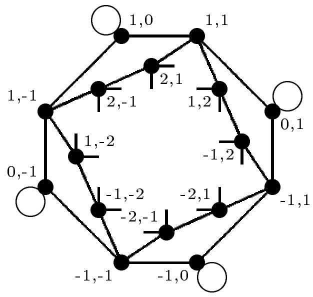
$$(Tf)(x, y) = f(x - y, y) + f(x, y - x).$$

Hence by Lemma 2.4.1

$$(\widehat{T}\widehat{f})(x, y) = \widehat{f}(x, x + y) + \widehat{f}(x + y, y)$$

where $\widehat{T}$ is operator on the space Y which is a subspace of $L^2(\mathbb{Z}^2)$ consisting of functions with $f(0) = 0$.

Consider an undirected graph H_1 with vertices $\mathbb{Z}^2 - \{0\}$ and (a, b) being connected to $(a + b, b)$ and $(a, a + b)$. We define a function r on undirected locally finite graphs so that $r(H_1) = r(\widehat{T})$.



A connected component of H_1.

Definition 2.4.2 Let $G = (V, E)$ be a locally finite undirected graph. Let

$$r(G) = \sup \Big\{ \sum_{\{u,v\} \in E} f(u) f(v); \sum_{v \in V} f(v)^2 = 1,\ f : V \to \mathbb{R} \Big\}.$$

The following Theorem gives us an upper bound for $r(G)$ of a graph G.

Theorem 2.4.3 *Let λ be a labeling of arcs of G with positive real numbers such that $\lambda(u, v) = 1/\lambda(v, u)$. Then*

$$r(G) \le \frac{1}{2} \sup_{u \in V} \sum_{v \in N(u)} \lambda(u, v).$$

Proof :
Using the inequality between arithmetic and geometric mean and the property of λ we obtain

$$\sum_{\{u,v\}\in E} f(u)f(v) \leq \sum_{\{u,v\}\in E} \frac{1}{2}\left(f(u)^2\lambda(u,v) + f(v)^2\lambda(v,u)\right) = (*).$$

Now

$$(*) = \frac{1}{2}\sum_{u\in V} f(u)^2 \sum_{v\in N(u)} \lambda(u,v) \leq \frac{1}{2}\left(\sup_{u\in V} \sum_{v\in N(u)} \lambda(u,v)\right)\sum_{u\in V} f(u)^2.$$

■

Now we use Theorem 2.4.3 to compute $r(H_1)$

Theorem 2.4.4

$$r(H_1) \leq 1/2 + \sqrt{2}.$$

Proof :
From Euclid's algorithm it follows that for any $d \in \mathbb{N}$ the vertices (a,b) such that $\gcd(a,b) = d$ form a connected component of H_1. These connected components are isomorphic and hence it is enough to consider only one of them, e.g. the one with $d = 1$. The mappings $(x,y) \to (-y,x)$ and $(x,y) \to (y,x)$ are automorphisms of H_1. Consider the following labeling λ of arcs (the arcs in the same orbits have the same labels)

- for the self loop e at $(0,1)$ let $\lambda(e) = 1$
- for the arc e from $(0,1)$ to $(1,1)$ let $\lambda(e) = a$ (the opposite arcs have label $1/a$)
- for an arc e from u to v which was not labeled yet let

$$\lambda(e) = \begin{cases} 1 & \text{if } ||u||_\infty = ||v||_\infty \\ b & \text{if } ||u||_\infty < ||v||_\infty \\ 1/b & \text{otherwise} \end{cases}$$

The labeling satisfies the conditions of Theorem 2.4.3. For $v = (0,1)$ we have $\sum_{u\in N(v)} \lambda(v,u) = 2a + 2$, for $v = (1,1)$ we have $\sum_{u\in N(v)} \lambda(v,u) = 2/a + 2b$. For the other vertices two neighbors have larger, one has smaller and one has equal infinity norm (for vertex (x,y), w.l.o.g. $x > y > 0$, then $(x+y,y), (x,x+y)$ have larger, $(x,y-x)$ has equal and $(x-y,y)$ has smaller infinity norm). Hence $\sum_{u\in N(v)} \lambda(v,u) = 1 + 2b + 1/b$. For $b = 1/\sqrt{2}$ and $a = \sqrt{2} - 1/2$ we have $\max(2a+2, 2/a + 2b, 1 + 2b + 1/b) = 1 + 2\sqrt{2}$ and hence $r(H_1) \leq 1/2 + \sqrt{2}$. ■

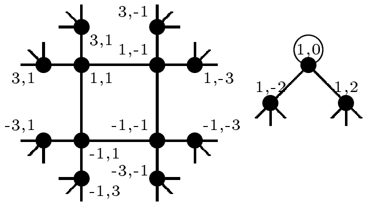
The operator T' mentioned at the end of the proof of Theorem 2.3.3 was defined as

$$(T'f)(x,y) = f(x-2y,y) + f(x,y-2x).$$

Hence by Lemma 2.4.1

$$(\widehat{T'}\widehat{f})(x,y) = \widehat{f}(x, x+2y) + \widehat{f}(x+2y, y).$$

Again we can show an upper bound on the Rayleigh quotient of T' in the space X by considering graph H_2 for which $r(H_2) = r(\widehat{T'}) = r(T')$. Let H_2 be the undirected graph with vertices $\mathbb{Z}^2 - \{0\}$ and (a,b) connected with $(a+2b, b)$ and $(a, b+2a)$.



A part of graph H_2.

Theorem 2.4.5

$$r(H_2) \leq \sqrt{3}.$$

Proof :
The mappings $(x,y) \to (-y,x)$ and $(x,y) \to (y,x)$ are automorphisms of H_1. Every vertex $u = (x,y)$ where both x, y are non-zero, $|x| \neq |y|$ is connected to three vertices with larger and one with smaller ∞ norm. If one of the x, y is zero or $|x| = |y|$ then u is connected to two vertices with larger and two vertices with equal ∞ norm (possibly via a self-loop). Label the arc e between the vertices u, v

$$\lambda(e) = \begin{cases} 1 & \text{if } ||u||_\infty = ||v||_\infty \\ a & \text{if } ||u||_\infty < ||v||_\infty \\ 1/a & \text{otherwise} \end{cases}$$

For any u we have either $\sum_{v\in N(u)} \lambda(u,v) = 3a + 1/a$ or $\sum_{v\in N(u)} \lambda(u,v) = 2 + 2a$. For $a = \sqrt{3}$ we obtain $r(H_2) \leq \sqrt{3}$. ■

It is not hard to find labelings of the vertices of H_1 and H_2 which show that the constants in Theorems 2.4.4 and 2.4.5 are optimal.

2.5 Lattice Duality: Banaszczyk's Transference Theorem

Given an an $n \times n$ regular matrix B a lattice L is

$$L = \{Bx\,;\, x \in \mathbb{Z}^n\}.$$

Alternatively a lattice can be viewed as a discrete additive subgroup of $\mathbb{R}^n$. Define the successive minima $\lambda_1, \ldots, \lambda_n$ of the lattice L

$$\lambda_i = \min\{r > 0; \dim \operatorname{span}(L \cap rB_n) \geq i\},$$

where B_n is the unit ball in $\mathbb{R}^n$. The dual lattice L^* of the lattice L is the lattice with matrix B^{-T}. Our goal is to prove the Transference Theorem of Banaszczyk [Ban93]

Theorem 2.5.1 *For any lattice L in $\mathbb{R}^n$*

$$\lambda_i(L)\lambda_{n+1-i}(L^*) \leq n.$$

The Theorem is tight up to a multiplicative constant as there exist a self dual lattice L such that $\lambda_1(L)^2 \geq \frac{n}{2\pi e}(1+o(1))$ as $n \to \infty$, a result of Conway and Thompson (see [Mil73], p. 42). A transference theorem was used in [LLS90] to show that $O(n)$-approximation of shortest lattice vector in L^2 norm cannot be NP-hard unless NP=co-NP.

In addition to the material covered in section 2.1 we will only need that characters of $\mathbb{R}^n$ are

$$\{\chi_b | b \in \mathbb{R}^n\}$$

where $\chi_b(x) = \exp(b^T x)\}$. The Fourier transform of a finite measure μ is defined as

$$\widehat{\mu}(x) = \int_{y \in \mathbb{R}^n} \chi_x(y)\, d\mu(y).$$

For $A \subseteq \mathbb{R}^n$ let

$$\rho(A) = \sum_{x \in A} e^{-\pi ||x||^2}. \tag{2.8}$$

Later in section 2.6 we will prove the following Lemma.

Lemma 2.5.2 *Let $n \geq 2$. Let A be a ball of diameter $\frac{3}{4}\sqrt{n}$ centered around the origin and let $u \in \mathbb{R}^n$. Then*

$$\frac{\rho\big((L+u) \setminus A\big)}{\rho(L)} \leq 0.285$$

Given a lattice L we define a discrete measure on $\mathbb{R}^n$ by

$$\sigma_L(A) = \frac{\rho(A \cap L)}{\rho(L)}.$$

Let

$$\phi_L(u) = \frac{\rho(L+u)}{\rho(L)}.$$

Later in section 2.6 we will prove

Lemma 2.5.3 *For a lattice L and its dual lattice L^**

$$\widehat{\sigma_L} = \phi_{L^*}.$$

Proof of Theorem 2.5.1

For $n = 1$, $\lambda_1(L)\lambda_1(L^*) = 1$. Assume that $n \geq 2$. Suppose that there is a lattice L such that $\lambda_i(L)\lambda_{n+1-i}(L^*) > n$. We can scale the lattice so that $\lambda_i(L) > \frac{3}{4}\sqrt{n}$ and $\lambda_{n+1-i}(L^*) > \frac{3}{4}\sqrt{n} + \frac{4}{5}$. Let A and A^* be the balls of diameters $\frac{3}{4}\sqrt{n}$ and $\frac{3}{4}\sqrt{n} + \frac{4}{5}$ centered around the origin. We have $\dim \operatorname{span}(L \cap A) < i$ and $\dim \operatorname{span}(L^* \cap A^*) < n + 1 - i$. Hence there is a vector u which is perpendicular to all vectors in $L \cap A$ and all vectors in $L^* \cap A^*$. We can chose u such that $||u|| = \frac{4}{5}$.

- Since u is perpendicular to all vectors in $L \cap A$,

$$\phi_{L^*}(u) = \widehat{\sigma}_L(u) = \sum_{x \in L} \sigma_L(x) e^{-2\pi i u^T x} = \sum_{x \in L \cap A} \sigma_L(x) + \sum_{x \in L \setminus A} \sigma_L(x) e^{-2\pi i u^T x} \geq$$
$$\sum_{x \in L \cap A} \sigma_L(x) - \sum_{x \in L \setminus A} \sigma_L(x) = 1 - 2 \sum_{x \in L \setminus A} \sigma_L(x) \geq 0.43. \quad (2.9)$$

- We have

$$\phi_{L^*}(u) = \frac{\rho\big((L^* \cap A^*) + u\big)}{\rho(L^*)} + \frac{\rho\big((L^* \setminus A^*) + u\big)}{\rho(L^*)}.$$

  Since u is perpendicular to vectors in $L^* \cap A^*$

$$\frac{\rho\big((L^* \cap A^*) + u\big)}{\rho(L^*)} = e^{-\pi ||u||^2} \frac{\rho(L^* \cap A)}{\rho(L^*)} \leq 0.135. \quad (2.10)$$

  The length of vectors in $(L^* \setminus A^*) + u$ is at least $\frac{3}{4}\sqrt{n}$ and hence

$$\frac{\rho\big((L^*/A^*) + u\big)}{\rho(L^*)} \leq \frac{\rho\big((L^* + u) \setminus \frac{3}{4}\sqrt{n}B\big)}{\rho(L^*)} \leq 0.285. \quad (2.11)$$

From (2.9),(2.10),(2.11) we obtain $0.43 \leq \phi_{L^*}(u) \leq 0.42$, a contradiction. ■

2.6 Gaussian-like Measures on Lattices

In this section we will prove Lemmas 2.5.3 and 2.5.2. Our only tool from harmonic analysis on $\mathbb{R}^n$ will be the Poisson Summation Formula.

Theorem 2.6.1 (Poisson Summation Formula) *Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a continuous function, such that for some $\varepsilon > 0$ and $c \in \mathbb{R}$*

$$f(x) \leq c(1 + |x|)^{-n-\varepsilon} \quad (2.12)$$
$$\widehat{f}(x) \leq c(1 + |x|)^{-n-\varepsilon} \quad (2.13)$$

where $\widehat{f}$ is the Fourier transform of f over $\mathbb{R}^n$. Then

$$\sum_{a \in \mathbb{Z}^n} f(a) = \sum_{a \in \mathbb{Z}^n} \widehat{f}(a).$$

Proof :
Define a function $F : \mathbb{T}^n \to \mathbb{R}$

$$F(x) = \sum_{a \in \mathbb{Z}^n} f(x+a). \tag{2.14}$$

The right-hand side of (2.14) converges uniformly because of (2.12). Clearly F is continuous and integrable. We have

$$\widehat{F}(y) = \int_{\mathbb{T}^n} F(x) e^{-2\pi i y^T x}\, d^n x = \int_{\mathbb{T}^n} \sum_{a \in \mathbb{Z}^n} f(x+a) e^{-2\pi i y^T (x+a)}\, d^n x = \int_{\mathbb{R}^n} f(x) e^{2\pi i y^T x}\, d^n x = \widehat{f}(y).$$

Condition (2.13) gives us $\widehat{F} \in L^1(\mathbb{Z}^n)$. Hence for any $x \in \mathbb{T}^n$ the inversion formula holds

$$F(x) = \sum_{a \in \mathbb{Z}^n} \widehat{f}(a) e^{-2\pi i a^T x}$$

and for $x = 0$ we obtain the result. ■

Lemma 2.6.2 *Let B be an $n \times n$ matrix and $u \in \mathbb{R}^n$. For $f(x) = \exp\big(-(Bx+u)^T(Bx+u)\big)$*

$$\widehat{f}(y) = \frac{\pi^{n/2}}{\det B} \exp\big(-\pi^2 (B^{-T}y)^T (B^{-T}y) + 2\pi i u^T (B^{-T}y)\big).$$

Remark 3 For $f(x) = \exp\big(-\pi(Bx+u)^T(Bx+u)\big)$ we obtain

$$\widehat{f}(y) = \frac{1}{\det B} \exp\big(-\pi (B^{-T}y)^T (B^{-T}y) + 2\pi i u^T (B^{-T}y)\big).$$

Proof of Lemma 2.6.2

$$\widehat{f}(y) = \int_{\mathbb{R}^n} \exp\big(-(Bx+u)^T(Bx+u) - 2\pi i y^T x\big)\, d^n x =$$
$$\int_{\mathbb{R}^n} \exp\big(-(Bx+u+\pi i B^{-T}y)^T(Bx+u+\pi i B^{-T}y) + 2\pi i y^T B^{-1} u - \pi^2 (B^{-T}y)^T(B^{-T}y)\big)\, d^n x =$$
$$\frac{\pi^{n/2}}{\det B} \exp\big(-\pi^2 (B^{-T}y)^T (B^{-T}y) + 2\pi i u^T (B^{-T}y)\big).$$

■

For notational convenience we define for $A \subseteq \mathbb{R}^n$

$$\rho'(A) = \sum_{x \in A} e^{-||x||^2}.$$

Note that $\rho'(A) = \rho(\pi^{-1/2} A)$ where ρ is defined by (2.8).

Lemma 2.6.3 *For a lattice $L \subseteq \mathbb{R}^n$, any $0 < t \leq 1$ and any $u \in \mathbb{R}^n$*

$$\rho'(tL+u) \leq \frac{1}{t^n}\rho'(L).$$

Proof :
Let B be the matrix of the lattice L. Using the Poisson summation formula

$$\rho'(tL+u) = \sum_{x\in\mathbb{Z}^n} \exp\left(-(tBx+u)^T(tBx+u)\right) = \frac{\pi^{n/2}}{t^n \det B}\sum_{y\in\mathbb{Z}^n} \exp\left(-\frac{\pi^2}{t^2}(B^{-T}y)^T(B^{-T}y)+\frac{2\pi i}{t}y^tB^{-1}u\right).$$

Again by the Poisson summation formula

$$\frac{1}{t^n}\rho'(L) = \frac{1}{t^n}\sum_{x\in\mathbb{Z}^n} \exp(-(Bx)^T(Bx)) = \frac{\pi^{n/2}}{t^n \det B}\sum_{y\in\mathbb{Z}^n} \exp\Big(-\pi^2(B^{-T}y)^T(B^{-T}y)\Big).$$

Now it is enough to notice that for any $y \in \mathbb{Z}^n$

$$\left|\exp\Big(-\frac{\pi^2}{t^2}(B^{-T}y)^T(B^{-T}y)+\frac{2\pi i}{t}y^tB^{-1}u\Big)\right| \leq \exp\Big(-\pi^2(B^{-T}y)^T(B^{-T}y)\Big).$$

■

Lemma 2.6.4 *For any lattice L any $u \in \mathbb{R}^n$ and $c \geq \sqrt{n/2}$*

$$\frac{\rho'\big((L+u)\setminus cB\big)}{\rho'(L)} \leq \left(\frac{2c^2}{n}\right)^{n/2}\exp\left(\frac{n}{2}-c^2\right).$$

Proof :
Let $x \in (L+u)\setminus cB$ and let $0 < t \leq 1$. The corresponding element $tx \in t(L+u)$ contributes to $\rho'\big(t(L+u)\big)$ by $\exp(-t^2||x||^2)$. We have

$$\frac{\exp(-t^2||x||^2)}{\exp(-||x||^2)} = \exp\big((1-t^2)||x||^2\big) \geq \exp\big((1-t^2)c^2\big).$$

Hence

$$\rho'\big((L+u)\setminus cB\big) \leq \rho'\big(t(L+u)\big)\exp\big((t^2-1)c^2\big).$$

By Lemma 2.6.3, $\rho'\big(t(L+u)\big) \leq (1/t)^n\rho'(L)$ and hence we have

$$\rho'\big((L+u)\setminus cB\big) \leq (1/t)^n\exp\big((t^2-1)c^2\big)\rho'(L).$$

For $t^2 = \frac{n}{2c^2}$ (optimal value) we obtain the result. ■

Corollary 2.6.5 (of Lemma 2.6.4) *For $c \geq \sqrt{n/(2\pi)}$*

$$\frac{\rho\big((L+u)\setminus cB\big)}{\rho(L)} \leq \left(\frac{2\pi c^2}{n}\right)^{n/2}\exp\left(\frac{n}{2}-\pi c^2\right).$$

Proof of Lemma 2.5.2
Plugging in $c = \frac{3}{4}\sqrt{n}$ into Corollary 2.6.5 we obtain

$$\frac{\rho\big((L+u)\setminus\frac{3}{4}\sqrt{n}B\big)}{\rho(L)} \leq \left(\frac{3}{2}\pi e^{1-\frac{9}{8}\pi}\right)^{n/2} \leq (0.285)^{n/2}.$$

which for $n \geq 2$ proves 2.5.2. ■

Proof of Lemma 2.5.3
Let B be the matrix of the lattice L. We have

$$\phi_L(u) = \frac{\rho(L+u)}{\rho(L)} = \frac{1}{\rho(L)}\sum_{x\in\mathbb{Z}^n}\exp\big(-\pi(Bx+u)^T(Bx+u)\big). \tag{2.15}$$

By Lemma 2.6.2 and the Poisson summation formula

$$(2.15) = \frac{1}{\rho(L)\det B}\sum_{y\in\mathbb{Z}^n}\exp\big(-\pi(B^{-T}y)^T(B^{-T}y)\big)\exp\big(2\pi i u^T(B^{-T}y)\big) =$$
$$\frac{1}{\rho(L)\det B}\sum_{y\in L^*}\exp\big(-\pi||y||^2\big)\exp\big(-2\pi i u^T y\big) = \frac{\rho(L^*)}{\rho(L)\det B}\widehat{\sigma_{L^*}}(u).$$

To finish the proof note

$$\rho(L) = \sum_{x\in\mathbb{Z}^n}\exp\big(-\pi(Bx)^T(Bx)\big) = \frac{1}{\det(B)}\sum_{y\in\mathbb{Z}^n}\exp\big(-\pi(B^{-T}y)^T(B^{-T}y)\big) = \frac{\rho(L^*)}{\det(B)}.$$

■

Maximum of $x^n e^{-x^2}$ is attained for $x = \sqrt{n/2}$.

Chapter 3

Generalizations of Harmonic Analysis

3.1 Introduction

In this section we will consider the theory from section 1 with the field of complex numbers replaced by a finite field. Most of the theorems from 1 remain valid even in this setting. We omit the proofs because they are identical with those in section 1.

Let G be a finite abelian group, let $n = |G|$. Let t be the exponent of G i.e. the smallest positive number such that $t \cdot g = 0$ for every $g \in G$. Let $\mathbb{F}$ be the finite field with q elements where $t|q-1$. Note that we have primitive t-th roots of unity and $1/n$ in $\mathbb{F}$. (To see that $(n,q) = 1$ note that any prime p that divides n divides t and hence does not divide q.)

We consider the space $\mathbb{F}^G$ of functions with

- pointwise multiplication $(fg)(a) = f(a)g(a),\ a \in G$,
- convolution $(f * g)(x) = \frac{1}{n}\sum\limits_{a \in G} f(a)g(x-a),\ x \in G$, and
- inner product $\langle f, g\rangle = \frac{1}{n}\sum\limits_{a\in G} f(a)\overline{g(a)}$. where $\overline{x} = \begin{cases} x^{-1} & \text{for } x \neq 0 \\ x & \text{for } x = 0 \end{cases}$.

Characters are homomorphisms from G to the multiplicative group of $\mathbb{F}$. The set of all characters is an orthonormal basis of $\mathbb{F}^G$. They also form a group (called the dual group of G over $\mathbb{F}$). Fix $G = \mathbb{Z}_{n_1} \oplus \cdots \oplus \mathbb{Z}_{n_k}$ and ω_i be a primitive n_i-th root of unity in $\mathbb{F}$ (since $n_i|t$ we have $\omega_i \in \mathbb{F}$). For $b = (b_1, \ldots, b_k) \in G$ let

$$\chi_b(x) = \prod_{i=1}^{k} \omega_i^{b_i x_i}.$$

The functions χ_b are characters, they are distinct and by a dimension argument they are all characters of G. Hence

$$\widehat{G} = \Big\{ \chi_b \mid b \in G, \ \Big\}. \tag{3.1}$$

Any function $f \in \mathbb{F}^G$ can be expressed as a linear combination of characters. The coefficient of χ_b is denoted by $\widehat{f}(\chi_b)$. The function $\widehat{f} : \widehat{G} \to \mathbb{F}$ is called the Fourier transform of f.

Mapping $b \to \chi_b$ is an isomorphism of G and $\widehat{G}$ and hence we view $\widehat{f}$ as a function in $\mathbb{F}^G$ and write $\widehat{f}(b)$ instead of $\widehat{f}(\chi_b)$.

The space $\mathbb{F}^{\widehat{G}}$ is endowed with

- inner product $\hat{\langle} f, g \hat{\rangle} = \sum_{a \in G} f(a)\overline{g(a)}$, and
- convolution $(f \hat{*} g)(x) = \sum_{a \in G} f(a) g(x-a)$.

By the orthogonality of characters

$$\widehat{f}(b) = \langle f, \chi_b \rangle = \frac{1}{n} \sum_{a \in G} f(a)\overline{\chi_a(b)}. \tag{3.2}$$

Theorem 3.1.1 *The Fourier transform satisfies*

- *linearity* $\widehat{f+g} = \widehat{f} + \widehat{g}$, $\widehat{af} = a\widehat{f}$, $f, g \in \mathbb{F}^G$, $a \in \mathbb{C}$
- $\widehat{fg} = \widehat{f} \hat{*} \widehat{g}$, $\widehat{f * g} = \widehat{f}\widehat{g}$
- $\langle f, g \rangle = \hat{\langle} \widehat{f}, \widehat{g} \hat{\rangle}$. *(the Plancherel formula)*

3.2 Sums of matrix columns (mod m)

We are going to prove a Theorem of Thérien [Thé94]. It was used to show that a circuit with MOD_m gates where m is composite needs at least $\Omega(n)$ gates to compute the AND of n inputs. A MOD_m gate is a gate which outputs 0 iff the sum of its inputs is divisible by m.

Theorem 3.2.1 *Let m, s, t be positive integers. If $t > c \cdot s \cdot m^{11/2} \ln m$ where c is an absolute constant, then for any $s \times t$ integer matrix there exists a set of columns which sum to the zero column modulo m.*

We shall use the following explicit version of Dirichlet's Theorem about primes in arithmetic progressions.

Theorem 3.2.2 ([Hea90]) *For any a coprime to m there is a prime $p \equiv a \pmod m$ such that $p \leq cm^{11/2}$ where c is an absolute constant.*

The estimate in Theorem 3.2.2 can be improved to $2(m \ln m)^2$ under the assumption of the Extended Riemann Hypothesis [BS94]. Theorem 3.2.1 follows from Theorem 3.2.2 and the following Lemma

Lemma 3.2.3 *Let p be a prime and A be an $s \times t$ integer matrix. If $t > s(p-1)\ln(p-1)$ then there exists a set of columns which sum to the zero column modulo $p-1$.*

We use harmonic analysis on the space of functions $G \to \mathbb{F}_p$ where $G = \mathbb{Z}_{p-1}^t$. Define the weight $wt(f)$ of f as the number of non-zero Fourier coefficients of f. When we write f and g as linear combination of characters we see

- $wt(f \cdot g) \le wt(f) \cdot wt(g)$
- $wt(f + g) \le wt(f) + wt(g)$

For example $\widehat{\mathbf{1}_0}(x) = \frac{1}{n}$ for any x and hence

$$wt(\mathbf{1}_0) = (p-1)^t. \tag{3.3}$$

Let $B = \{0,1\}^t \subseteq \mathbb{Z}_{p-1}^t$. We have

$$\widehat{\mathbf{1}_B}(x) = \frac{1}{n}\sum_{y \in B} \omega^{x^T y} = \frac{1}{n}(1+\omega^{x_1})\dots(1+\omega^{x_t}) \tag{3.4}$$

Since (3.4) is non-zero for $(p-2)^t$ choices of x we obtain

$$wt(\mathbf{1}_B) = (p-2)^t. \tag{3.5}$$

Proof of Lemma 3.2.3
Let A be an $s \times t$ integer matrix such that no set of columns sums to the zero column modulo $p-1$. For each row of the matrix define a function

$$f_i(x) = \omega^{x_1 a_{i,i} + \dots + x_t a_{i,t}}, \quad i \in [s].$$

Consider the function

$$f = \prod_{i=1}^{s} \left(1 - (1-f_i)^{p-1}\right).$$

If $x = 0$ then each $f_i(x) = 1$ and hence $f(x) = 1$. Since no set of columns sums to the zero column modulo $p-1$, for any $x \in B \setminus \{0\}$ there is i such that $f_i(x) \ne 1$ and hence $f(x) = 0$. Thus $f|_B \equiv \mathbf{1}_0|_B$.

Each f_i is a character. The function $1-(1-f_i)^{p-1}$ is a linear combination of powers of the character f_i and hence has weight at most $p-1$. Hence by the submultiplicativity of weight $wt(f) \le (p-1)^t$.

Since $f \cdot \mathbf{1}_B = \mathbf{1}_0$

$$(p-1)^t = wt(\mathbf{1}_0) = wt(f \cdot \mathbf{1}_B) \le (p-1)^s (p-2)^t$$

and hence

$$e^{t/(p-1)} \le \left(1 + \frac{1}{p-2}\right)^t \le (p-1)^s$$

which yields $t \le s(p-1)\ln(p-1)$. ■

Bibliography

[AC88] N. Alon and F.R.K. Chung. Explicit construction of linear sized tolerant networks. *Discrete Mathematics*, 72:15–19, 1988.

[AGHP92] N. Alon, O. Goldreich, J. Håstad, and R. Peralta. Simple construction of almost k-wise independent random variables. *Random Structures and Algorithms*, 3(3):289–304, 1992.

[Ajt83] M. Ajtai. Σ_1^1-formulae on finite structures. *Annals of Pure and Applied Logic*, 24:1–48, 1983.

[AKS83] M. Ajtai, J. Komlós, and E. Szemerédi. Sorting in $c \log n$ parallel steps. *Combinatorica*, 3(1):1–19, 1983.

[AKS87] M. Ajtai, J. Komlós, and E. Szemerédi. Deterministic simulation in logspace. *19th Annual ACM Symposium on the The Theory of Computing (STOC)*, pages 132–140, 1987.

[ASE92] N. Alon, J.H. Spencer, and P. Erdős. *The Probabilistic Method*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley, 1992.

[Bab89] L. Babai. Fourier transforms and equations over finite abelian groups. *(manuscript)*, 1989.

[Ban91] W. Banaszczyk. *Additive Subgroups of Topological Vector Spaces*. Lecture notes in Mathematics, Vol. 1466. Springer Verlag, 1991.

[Ban93] W. Banaszczyk. New bounds in some transference theorems in the geometry of numbers. *Mathematische Annalen*, 296(4):625–635, 1993.

[Bec75] W. Beckner. Inequalities in Fourier analysis. *Annals of Mathematics*, 102(1):159–182, 1975.

[Ber98] A. Bernasconi. Harmonic analysis and Boolean function complexity. *Calcolo*, 35(3):149–186, 1998.

[BFJ$^+$94] A. Blum, M. Furst, J. Jackson, M. Kearns, Y. Mansour, and S. Rudich. Weakly learning DNF and characterizing statistical query learning using Fourier analysis. *26th Annual ACM Symposium on the Theory of Computing (STOC)*, pages 63–72, 1994.

[BKK$^+$92] J. Bourgain, J. Kahn, G. Kalai, Y. Katznelson, and Nathan Linial. The influence of variables in product spaces. *Israel Journal of Mathematics*, 77(1-2):55–64, 1992.

[BL90] M. Ben-Or and N. Linial. Collective coin flipping. *Randomness and Computation (S. Micali ed.)*, Academic Press:91–115, 1990.

[BOH90] Y. Brandman, A. Orlitsky, and J. Hennessy. A spectral lower bound technique for the size of decision trees and two-level and/or circuits. *IEEE Transactions on Computers*, 39(2):282–287, 1990.

[BS94] E. Bach and J. Sorenson. Explicit bounds for primes in residue classes. *Proceedings of Symposia in Applied Mathematics*, 48:535–539, 1994.

[BT96] N.H. Bshouty and C. Tamon. On the Fourier spectrum of monotone functions. *Journal of ACM*, 43(4):747–770, 1996.

[Cai99] J.Y. Cai. A new transference theorem and applications to Ajtai's connection factor. *Computing an Combinatorics (Tokyo, 1999)*, Lecture notes in Computer Science, 1627, Springer:113–122, 1999.

[CDG87] F.R.K. Chung, P. Diaconis, and R.L. Graham. Random walks arising in random number generation. *Annals of Probability*, 15(3):1148–1165, 1987.

[Cha87] D.C. Champeney. *A Handbook of Fourier Theorems.* Cambridge University Press, 1987.

[Che66] E.W. Cheney. *Introduction to Approximation Theory.* McGraw-Hill, 1966.

[CT65] J.W. Cooley and J.W. Tukey. An algorithm for the machine calculation of complex Fourier series. *Mathematics of Computation*, 90(19):297–301, 1965.

[CW89] A. Cohen and A. Wigderson. Dispersers, deterministic amplification, and weak random sources. *30th Annual IEEE Symposium on The Foundations of Computer Science (FOCS)*, pages 14–19, 1989.

[DM72] H. Dym and H.P. McKean. *Fourier Series and Integrals.* Academic Press, 1972.

[FSS81] M. Furst, J. Saxe, and M. Sipser. Parity, circuits, and the polynomial time hierarchy. *22nd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 260–270, 1981.

[FSS84] M. Furst, J. Saxe, and M. Sipser. Parity, circuits, and the polynomial time hierarchy. *Mathematical Systems Theory*, 17(1):13–27, 1984.

[GG79] O. Gabber and Z. Galil. Explicit constructions of linear size superconcentrators. *20th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 364–370, 1979.

[GL94] C. Gotsman and N. Linial. Spectral properties of threshold functions. *Combinatorica*, 14(1):35–50, 1994.

[Hås86] J. Håstad. Computational limitations for small depth circuits. *Ph.D. dissertation*, MIT Press, 1986.

[Hea90] D.R. Heath-Brown. Siegel zeros and the least prime in an arithmetic progression. *The Quarterly Journal of Mathematics, Oxford Series 2*, 41(164):405–418, 1990.

[HLP34] G.H. Hardy, J.E. Littlewood, and G. Pólya. *Inequalities.* Cambridge University Press, 1934.

[IZ89] R. Impagliazzo and D. Zuckermann. How to recycle random bits. *30th Annual IEEE Symposium on The Foundations of Computer Science (FOCS)*, pages 222–227, 1989.

[Jac95] J. Jackson. The harmonic sieve: A novel application of Fourier analysis to machine learning theory and practice. *(Ph.D. Dissertation), Carnegie Mellon University*, 1995.

[JM87] S. Jimbo and A. Maruoka. Expanders obtained from affine transformations. *Combinatorica*, 7(4):343–355, 1987.

[Kat68] Y. Katznelson. *An Introduction To Harmonic Analysis.* Wiley, 1968.

[KKL88] J. Kahn, G. Kalai, and N. Linial. The influence of variables on Boolean functions. *29th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 68–80, 1988.

[KKS95] J. Kahn, J. Komlós, and E. Szemerédi. On the probability that a random ± 1-matrix is singular. *Journal of AMS*, 8(1):223–240, 1995.

[KM93] E. Kushilevitz and Y. Mansour. Learning decision trees using the Fourier spectrum. *SIAM Journal of Computing*, 22(6):1331–1348, 1993.

[LLS90] J.C. Lagarias, H.W. Lenstra, and C.P. Schnorr. Korkin-Zolotarev bases and successive minima of lattice and its reciprocal lattice. *Combinatorica*, 10(4):333–348, 1990.

[LMN93] N. Linial, Y. Mansour, and N. Nisan. Constant depth circuits, Fourier transform, and learnability. *Journal of ACM*, 40(3):607–620, 1993.

[LPS88] A. Lubotzky, R. Phillips, and P. Sarnak. Ramanujan graphs. *Combinatorica*, 8(3):261–277, 1988.

[Mar73] G.A. Margulis. Explicit constructions of concentrators. *Problems of Information Transmission (translated from Russian)*, 9(4):325–332, 1973.

[Mar88] G.A. Margulis. Explicit group theoretic constructions of combinatorial schemes and their applications for construction of expanders and concentrators. *Problems of Information Transmission (translated from Russian)*, 24(1):39–46, 1988.

[Mat99] J. Matoušek. *Geometric Discrepancy.* Springer Verlag, 1999.

[Mil73] J. Milnor. *Symmetric Bilinear Forms.* Springer Verlag, 1973.

[MR95] R. Motwani and P. Raghavan. *Randomized Algorithms.* Cambridge University Press, 1995.

[Nis91] N. Nisan. CREW PRAMs and decision trees. *SIAM Journal of Computing*, 20(6):999–1007, 1991.

[NN93] J. Naor and M. Naor. Small-bias probability spaces: efficient constructions and applications. *SIAM Journal of Computing*, 22(4):838–856, 1993.

[NS94] N. Nisan and M. Szegedy. On the degree of Boolean functions as real polynomials. *Computational Complexity (special issue on circuit complexity)*, 4(4):301–313, 1994.

[Rei68] H. Reiter. *Classical Harmonic Analysis and Locally Compact Groups.* Clarendon Press, Oxford, 1968.

[Rud87] W. Rudin. *Real And Complex Analysis.* McGraw-Hill, 1987.

[Sak93] M.E. Saks. Slicing the hypercube. *Surveys in Combinatorics*, London Mathematical Society Lecture notes Series, 187:211–255, 1993.

[Sch76] Wolfgang M. Schmidt. *Equations over Finite Fields: An Elementary Approach.* Lecture notes in Mathematics, Vol. 536. Springer Verlag, 1976.

[SS71] A. Schönhage and V. Strassen. Schnelle Multiplikation grosser Zahlen. *Computing*, 7:281–292, 1971.

[Thé94] D. Thérien. Circuits constructed with mod_q gates cannot compute "AND" in sublinear size. *Computational Complexity (Special Issue on Circuit Complexity)*, 4(4):383–388, 1994.

[Yao85] A. C. Yao. Separating the polynomial-time hierarchy by oracles. *26th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1–10, 1985.

[Zyg59] A. Zygmund. *Trigonometric Series I-II (second edition).* Cambridge University Press, 1959.

[Zyg76] A. Zygmund. Notes on the history of Fourier series. *Studies in Harmonic Analysis (J.M. Ash, editor), Studies in Mathematics v. 13*, pages 1–19, 1976.