

Lecture 8

Instructor: Sruthi Sekar

Scribe: Nidhi Jha

1 Boolean functions and the Fourier expansion

1.1 Some basic Fourier formula

Theorem 1. Parseval's Theorem: For any $f : \{-1, 1\}^n \rightarrow \mathbb{R}$,

$$\langle f, f \rangle = \mathbb{E}_{\mathbf{x} \sim \{-1, 1\}^n} [f(\mathbf{x})^2] = \sum_{S \subseteq [n]} (\hat{f}(S))^2$$

In particular, if $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$ is Boolean-valued then

$$\sum_{S \subseteq [n]} \hat{f}(S)^2 = 1$$

Theorem 2. Plancherel's Theorem: For any $f, g : \{-1, 1\}^n \rightarrow \mathbb{R}$,

$$\langle f, g \rangle = \mathbb{E}_{\mathbf{x} \sim \{-1, 1\}^n} [f(\mathbf{x})g(\mathbf{x})] = \sum_{S \subseteq [n]} \hat{f}(S)\hat{g}(S)$$

Remark: For $f, g : \{-1, 1\}^n \rightarrow \{-1, 1\}$ (Boolean valued), $\langle f, g \rangle$ can be interpreted as some kind of correlation between f and g according to the following corollary.

Corollary: If $f, g : \{-1, 1\}^n \rightarrow \{-1, 1\}$ (Boolean valued), then

$$\langle f, g \rangle = \mathbb{E}_{\mathbf{x} \sim \{-1, 1\}^n} [f(\mathbf{x})g(\mathbf{x})] = \Pr[f(\mathbf{x}) = g(\mathbf{x})] - \Pr[f(\mathbf{x}) \neq g(\mathbf{x})] = 1 - 2\text{dist}(f, g)$$

where $\text{dist}(f, g)$ = Relative Hamming distance.

In particular,

$$\langle f, f \rangle = 1 \implies \sum_{S \subseteq [n]} (\hat{f}(S))^2 = 1$$

Definition 3. For $f, g : \{-1, 1\}^n \rightarrow \{-1, 1\}$, relative Hamming distance $\text{dist}(f, g) := \Pr[f(\mathbf{x}) \neq g(\mathbf{x})]$.

Exercise 4. Prove the uniqueness of the Fourier expansion using Parseval's theorem.

Proof. Let $f(x)$ have two distinct Fourier expansions:

$$f(x) = \sum_{S \subseteq [n]} \hat{f}(S)\chi_S(x)$$

and

$$f(x) = \sum_{S \subseteq [n]} \tilde{\hat{f}}(S)\chi_S(x)$$

As $f(x)$ is the same function in both expansions. Then, subtract the two expansions:

$$0 = \sum_{S \subseteq [n]} (\hat{f}(S) - \tilde{\hat{f}}(S))\chi_S(x)$$

Now, apply Parseval's theorem to the difference of the two expansions:

$$\sum_{S \subseteq [n]} |\hat{f}(S) - \tilde{f}(S)|^2 = \mathbb{E}[0^2] = 0$$

Since the sum of squares of the Fourier coefficients equal zero, each individual term in the sum must be zero:

$$|\hat{f}(S) - \tilde{f}(S)|^2 = 0 \quad \text{for all } S \subseteq [n]$$

This implies:

$$\hat{f}(S) = \tilde{f}(S) \quad \text{for all } S \subseteq [n]$$

Thus, we have shown that the Fourier coefficients must be identical in both expansions. This proves the uniqueness of the Fourier expansion. $\square$

1.2 Some interesting combinatorial properties of f

- **"Bias of f ":** Mean of $f : \{-1, +1\}^n \rightarrow \mathbb{R}$ is $\mathbb{E}[f]$.
 f is unbiased or balanced if $\mathbb{E}[f] = 0$.
 In particular, for $f : \{-1, +1\}^n \rightarrow \{-1, +1\}$

$$0 = \hat{f}(\emptyset) = \mathbb{E}[f] = \Pr[f = 1] - \Pr[f = -1]$$

$\implies f$ is unbiased if and only if it takes value 1 on exactly half the points of the Hamming cube.
 Similarly, f is unbiased if and only if $\hat{f}(\emptyset) = 0$

- **Variance of f :** For $f : \{-1, +1\}^n \rightarrow \mathbb{R}$,

$$\text{Var}[f] = \langle f - \mathbb{E}[f], f - \mathbb{E}[f] \rangle = \mathbb{E}[f^2] - (\mathbb{E}[f])^2 \stackrel{\text{Parseval's}}{=} \sum_{S \neq \emptyset} (\hat{f}(S))^2$$

For $f : \{-1, +1\}^n \rightarrow \{-1, +1\}$,

$$\text{Var}[f] = 1 - (\mathbb{E}[f])^2 = 4\Pr_{\mathbf{x}}[f(\mathbf{x}) = 1]\Pr_{\mathbf{x}}[f(\mathbf{x}) = -1] \in [0, 1]$$

So, $\text{Var}[f] = 1$ if f is unbiased or balanced and $\text{Var}[f] = 0$ if f is a constant function.

- **Covariance of f, g :** For $f, g : \{-1, +1\}^n \rightarrow \mathbb{R}$, $\text{Cov}[f, g]$ measures the variance between f & g

$$\text{Cov}[f, g] := \langle f - \mathbb{E}[f], g - \mathbb{E}[g] \rangle = \mathbb{E}[fg] - \mathbb{E}[f] \cdot \mathbb{E}[g] = \sum_{S \neq \emptyset} \hat{f}(S) \hat{g}(S)$$

1.3 Probability densities and convolution

Definition 5. Probability density function on Hamming cube $\mathbb{F}_2^n$ is any non-negative function $\phi : \mathbb{F}_2^n \rightarrow \mathbb{R}^{\geq 0}$ satisfying:

$$\mathbb{E}_{\mathbf{x} \sim \{-1, 1\}^n} [\phi(\mathbf{x})] = 1$$

Notation: $Y \sim \phi$ meaning Y is drawn from associated distribution:

$$\Pr_{Y \sim \phi} [Y = y] = \phi(y) \cdot \frac{1}{2^n} \forall y \in \mathbb{F}_2^n$$

Example 6. If ϕ is a density function and $g : \mathbb{F}_2^n \rightarrow \mathbb{R}$ then

$$\mathbb{E}_{Y \sim \phi} [g(Y)] = \langle \phi, g \rangle = \mathbb{E}_{\mathbf{X} \sim \mathbb{F}_2^n} [\phi(\mathbf{X})g(\mathbf{X})]$$

Example 7. $\phi = 1$ (Constant function)

Then,

$$\Pr_{Y \sim \phi} [Y = y] = \frac{1}{2^n} \quad \forall y \in \mathbb{F}_2^n \quad (\text{Uniform distribution on } \mathbb{F}_2^n)$$

Example 8. Uniform distribution over some subset $A \subseteq \mathbb{F}_2^n$

If $A \subseteq \mathbb{F}_2^n$, define indicator function of A , $1_A : \mathbb{F}_2^n \rightarrow \{0, 1\}$, as:

$$1_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$$

Assume $A \neq \emptyset$, then density function associated with A :

$$\phi(A) = \frac{1}{\mathbb{E}[1_A]} \cdot 1_A$$

• **Notation:** $Y \sim \phi_A$ is written as $Y \sim A$

Example 9. For $A = \{0\}$, $\phi_{\{0\}}(x) = \begin{cases} 2^n & x = \mathbf{0} \\ 0 & x \in \mathbb{F}_2^n \setminus \{0\} \end{cases}$

Definition 10. Convolution of f, g : For $f, g : \mathbb{F}_2^n \rightarrow \mathbb{R}$

$$(f * g)(\mathbf{x}) = \mathbb{E}_{\mathbf{y} \sim \mathbb{F}_2^n} [f(\mathbf{y})g(\mathbf{x}-\mathbf{y})] = \mathbb{E}_{\mathbf{y} \sim \mathbb{F}_2^n} [f(\mathbf{x}-\mathbf{y})g(\mathbf{y})]$$

Since in $\mathbb{F}_2^n$, addition is same as subtraction,

$$\implies (f * g)(\mathbf{x}) = \mathbb{E}_{\mathbf{y} \sim \mathbb{F}_2^n} [f(\mathbf{y})g(\mathbf{x}+\mathbf{y})] = \mathbb{E}_{\mathbf{y} \sim \mathbb{F}_2^n} [f(\mathbf{x}+\mathbf{y})g(\mathbf{y})]$$

Exercise 11. Prove associativity and commutativity of $*$.

Proof. Associativity: $f * (g * h) = (f * g) * h$

We compute the left-hand side:

$$\begin{aligned} (f * (g * h))(x) &= \mathbb{E}_{y \in \mathbb{F}_2^n} [f(y) \cdot (g * h)(x - y)] \\ &= \mathbb{E}_{y \in \mathbb{F}_2^n} [f(y) \cdot \mathbb{E}_{z \in \mathbb{F}_2^n} [g(z)h((x - y) - z)]] \\ &= \mathbb{E}_{y, z \in \mathbb{F}_2^n} [f(y)g(z)h(x - y - z)]. \end{aligned}$$

Now compute the right-hand side:

$$\begin{aligned} ((f * g) * h)(x) &= \mathbb{E}_{z \in \mathbb{F}_2^n} [(f * g)(z) \cdot h(x - z)] \\ &= \mathbb{E}_{z \in \mathbb{F}_2^n} [\mathbb{E}_{y \in \mathbb{F}_2^n} [f(y)g(z - y)] \cdot h(x - z)] \\ &= \mathbb{E}_{y, z \in \mathbb{F}_2^n} [f(y)g(z - y)h(x - z)]. \end{aligned}$$

Now perform a change of variable: let $w = z - y$, so $z = w + y$. Then:

$$\mathbb{E}_{y,z \in \mathbb{F}_2^n} [f(y)g(z-y)h(x-z)] = \mathbb{E}_{y,w \in \mathbb{F}_2^n} [f(y)g(w)h(x-y-w)].$$

Therefore, both expressions are equal:

$$f * (g * h)(x) = (f * g) * h(x),$$

and convolution is associative over $\mathbb{F}_2^n$. □

Commutativity: $f * g = g * f$

$$(f * g)(x) = \mathbb{E}_{\mathbf{y} \sim \mathbb{F}_2^n} [f(\mathbf{y})g(\mathbf{x}+\mathbf{y})] = \mathbb{E}_{\mathbf{y} \sim \mathbb{F}_2^n} [g(\mathbf{x}+\mathbf{y})f(\mathbf{y})] = (g * f)(x) \quad \forall x \in \mathbb{F}_2^n$$

$$\implies f * g = g * f \quad \square$$

- If ϕ & φ are probability density functions, then $\phi * \varphi$ is also a probability density function. Moreover, $\phi * \varphi$ represents the following distribution: For $X \sim \phi$ & $Y \sim \varphi$ sampled independently, $\phi * \varphi$ represents the distribution of $X + Y$.

Theorem 12. For $f, g : \mathbb{F}_2^n \rightarrow \mathbb{R}$

$$\forall S \subseteq [n], \quad \widehat{f * g}(S) = \widehat{f}(S) \widehat{g}(S)$$

Proof.

$$\begin{aligned} \widehat{f * g}(S) &= \mathbb{E}_{\mathbf{x} \sim \mathbb{F}_2^n} [(f * g)(\mathbf{x}) \chi_S(\mathbf{x})] \\ &= \mathbb{E}_{\mathbf{x} \sim \mathbb{F}_2^n} \left[\mathbb{E}_{\mathbf{y} \sim \mathbb{F}_2^n} [(f(\mathbf{y})g(\mathbf{x}-\mathbf{y})) \chi_S(\mathbf{x})] \right] \quad (\text{by definition}) \\ &= \mathbb{E}_{\mathbf{y}, \mathbf{z} \sim \mathbb{F}_2^n} [f(\mathbf{y})g(\mathbf{z}) \chi_S(\mathbf{y}+\mathbf{z})] \quad (\because \mathbf{x}-\mathbf{y} \text{ is uniform on } \mathbb{F}_2^n) \\ &= \mathbb{E}_{\mathbf{y}, \mathbf{z} \sim \mathbb{F}_2^n} [f(\mathbf{y})g(\mathbf{z}) \chi_S(\mathbf{y}) \chi_S(\mathbf{z})] \\ &= \mathbb{E}_{\mathbf{y} \sim \mathbb{F}_2^n} [f(\mathbf{y}) \chi_S(\mathbf{y})] \mathbb{E}_{\mathbf{z} \sim \mathbb{F}_2^n} [g(\mathbf{z}) \chi_S(\mathbf{z})] \quad (\text{by independence}) \\ &= \widehat{f}(S) \cdot \widehat{g}(S) \end{aligned}$$

□

1.4 Application: Almost linear functions and the BLR Test

Definition 13. A function $f : \mathbb{F}_2^n \rightarrow \mathbb{F}_2$ is linear if either of the following equivalent conditions hold:

1. $f(x + y) = f(x) + f(y) \quad \forall x, y \in \mathbb{F}_2^n$;
2. $f(x) = a \cdot x$ for some $a \in \mathbb{F}_2^n$; i.e., $f(x) = \sum_{i \in S} x_i$ for some $S \subseteq [n]$.

Proof. (1) $\Rightarrow$ (2) Assume that f is additive: $f(x + y) = f(x) + f(y)$ for all $x, y \in \mathbb{F}_2^n$.

Let $e_1, \dots, e_n$ be the standard basis vectors of $\mathbb{F}_2^n$, where e_i has a 1 in the i -th coordinate and 0 elsewhere. Define $a_i := f(e_i)$, and let $a = (a_1, \dots, a_n) \in \mathbb{F}_2^n$. Now, take any $x = (x_1, \dots, x_n) \in \mathbb{F}_2^n$. Then:

$$x = \sum_{i=1}^n x_i e_i$$

Using additivity of f :

$$f(x) = f\left(\sum_{i=1}^n x_i e_i\right) = \sum_{i=1}^n f(x_i e_i)$$

Since $x_i \in \mathbb{F}_2$, $x_i e_i$ is either 0 or e_i , and $f(0) = 0$. So:

$$f(x_i e_i) = x_i f(e_i) = x_i a_i$$

Hence:

$$f(x) = \sum_{i=1}^n x_i a_i = a \cdot x$$

Therefore, $f(x)$ can be written as a dot product with some $a \in \mathbb{F}_2^n$, proving the second condition.

(2) $\Rightarrow$ (1) Assume that $f(x) = a \cdot x = \sum_{i=1}^n a_i x_i$ for some fixed $a \in \mathbb{F}_2^n$.

Let $x, y \in \mathbb{F}_2^n$. Then:

$$f(x+y) = \sum_{i=1}^n a_i (x_i + y_i) = \sum_{i=1}^n (a_i x_i + a_i y_i) = \sum_{i=1}^n a_i x_i + \sum_{i=1}^n a_i y_i = f(x) + f(y)$$

So f satisfies the additivity condition.

The two definitions of linearity for functions $f : \mathbb{F}_2^n \rightarrow \mathbb{F}_2$ are equivalent:

$$f(x+y) = f(x) + f(y) \quad \Leftrightarrow \quad f(x) = a \cdot x \text{ for some } a \in \mathbb{F}_2^n$$

□

Let's think of what it might mean for a function $f : \mathbb{F}_2^n \rightarrow \mathbb{F}_2$ to be approximately linear.

1'. $f(x+y) = f(x) + f(y)$ for almost all pairs $x, y \in \mathbb{F}_2^n$;

2'. there is some $S \subseteq [n]$ such that $f(x) = \sum_{i \in S} x_i$ for almost all $x \in \mathbb{F}_2^n$.

Proof. $2' \Rightarrow 1'$ is immediate from the previous proof.

$1' \Rightarrow 2'$ Motivation for this problem comes from an area of theoretical computer science called property testing. □

1.4.1 Property Testing for Boolean Functions

Imagine that you have “black-box” access to a function $f : \mathbb{F}_2^n \rightarrow \mathbb{F}_2$, meaning that the function f is unknown to you but you can “query” its value on inputs $x \in \mathbb{F}_2^n$ of your choosing. The function f is “supposed” to be a linear function, and you would like to try to verify this.

The only way you can be certain f is indeed a linear function is to query its value on all 2^n inputs; unfortunately, this is very expensive. The idea behind “property testing” is to try to verify that f has a certain property— in this case, linearity— by querying its value on just a few random inputs. In exchange for efficiency, we need to be willing to only approximately verify the property.

Definition 14. If f and g are Boolean-valued functions we say they are ε -close if $\text{dist}(f, g) \leq \varepsilon$; otherwise we say they are ε -far. If $\mathcal{P}$ is a (nonempty) property of n -bit Boolean functions we define $\text{dist}(f, \mathcal{P}) = \min_{g \in \mathcal{P}} \text{dist}(f, g)$. We say that f is ε -close to $\mathcal{P}$ if $\text{dist}(f, \mathcal{P}) \leq \varepsilon$; i.e., f is ε -close to some g satisfying $\mathcal{P}$.

Example 15. Check if $f = 1$ or not?

Given black-box access to f

1. Sample k independent random $x_1, x_2, \dots, x_k$
2. IF $f(x_i) = 1 \quad \forall \quad i \in [k]$ then
 return Y
 Else return N

Analysis: If $f = 1$, $\Pr[\text{Algorithm outputs } Y] = 1$

If f is ε -far from 1 $\implies \Pr[f(x) \neq 1] \geq \varepsilon$

$f(x) = 1$ for at most $(1 - \varepsilon)$ -fraction of inputs

$\Pr[\text{Algorithm outputs } Y] \leq (1 - \varepsilon)^k \leq \frac{1}{e} \quad \text{if } k \geq \frac{1}{\varepsilon}$

Example 16. Linearity Testing:

BLR Test

Given: Black-box access to $f : \mathbb{F}_2^n \rightarrow \mathbb{F}_2$

Goal: Output Y iff f is approximately linear.

1. Sample $x \sim \mathbb{F}_2^n$, and $y \sim \mathbb{F}_2^n$ independently
2. Query f at $x, y, x + y$
 If $f(x + y) = f(x) + f(y)$ return Y
 Else return N

Analysis: If f is linear, then $\Pr[\text{BLR outputs } Y] = 1$

Instead of $f : \mathbb{F}_2^n \rightarrow \mathbb{F}_2$, we can look at $g : \mathbb{F}_2^n \rightarrow \{-1, +1\}$ such that $g(x) = (-1)^{f(x)}$

Now the BLR test for g would be outputting Y iff $g(x + y) = g(x)g(y)$

Now to complete the analysis, consider

$$\begin{aligned}
 \Pr[\text{BLR accepts } g] &= \mathbb{E}_{x,y} \left[\frac{1}{2} + \frac{1}{2} g(x)g(y)g(x+y) \right] \\
 &= \frac{1}{2} + \frac{1}{2} \mathbb{E}_x [g(x) \mathbb{E}_y [g(y)g(x+y)]] \\
 &= \frac{1}{2} + \frac{1}{2} \mathbb{E}_x [g(x) \cdot (g * g)(x)] \quad (\text{from definition}) \\
 &= \frac{1}{2} + \frac{1}{2} \sum_{S \subseteq [n]} \hat{g}(S) \cdot \widehat{g * g}(S) \quad (\text{from Plancherel's theorem}) \\
 &= \frac{1}{2} + \frac{1}{2} \sum_{S \subseteq [n]} (\hat{g}(S))^3 \quad (1)
 \end{aligned}$$

Now if $\Pr[\text{BLR accepts } g] = 1 - \varepsilon$

then (1) $\implies 1 - 2\varepsilon = \sum_{S \subseteq [n]} (\hat{g}(S))^3 \leq \max_{S \subseteq [n]} \{\hat{g}(S)\} \cdot \sum_{S \subseteq [n]} (\hat{g}(S))^2 = \max_{S \subseteq [n]} \{\hat{g}(S)\} \cdot 1$

But $\hat{g}(S) = \langle g, \chi_S \rangle = 1 - 2 \cdot \text{dist}(g, \chi_S)$

$\implies \exists \text{ some } S^* \subseteq [n] \text{ such that } 1 - 2\varepsilon \leq 1 - 2 \cdot \text{dist}(g, \chi_{S^*})$

$\implies \text{dist}(g, \chi_{S^*}) \leq \varepsilon$

i.e. g is ε -close to a linear function.

BLR Test shows: Just given three queries to f , you can show if f is close to some linear function χ_S or not.

But, can we find this χ_S ?

1.5 Local Correctability

Theorem 17. Suppose $f : \mathbb{F}_n^2 \rightarrow \{-1, 1\}$ is ε -close to the linear function χ_S . Then for every $x \in \mathbb{F}_2^n$, the following algorithm outputs $\chi_S(x)$ with probability $\geq 1 - 2\varepsilon$:

1. Choose $y \sim \mathbb{F}_2^n$.
2. Query f at y and $x + y$.
3. Output $f(y)f(x + y)$.

Proof. Since y and $x + y$ are both uniformly distributed on $\mathbb{F}_2^n$ (though not independently), we have $\Pr[f(y) \neq \chi_S(y)] \leq \varepsilon$ and $\Pr[f(x + y) \neq \chi_S(x + y)] \leq \varepsilon$ by assumption. By the union bound, the probability of either event occurring is at most 2ε ; when neither occurs,

$$f(y)f(x + y) = \chi_S(y)\chi_S(x + y) = \chi_S(x)$$

as desired. □

References

- [ODo] Ryan O'Donnell. *Fourier Analysis of Boolean functions*. https://www.youtube.com/watch?v=1XJP-UkTl-4&list=PLm3J0oaFux3ZYpFLwwrlv_EHH9wtH6pnX&index=27.
- [ODo14] Ryan O'Donnell. *Analysis of boolean functions*. Cambridge University Press, 2014.