# Introduction to Harmonic analysis

A chapter for the Mathematics++ Lecture Notes

Robert Šámal

Rev. 2014-06-03 RS

**Class 1 (RS) Oct 3, 2013**

These are lecture notes for a class at the Charles University in Prague, but hopefully interesting for more general audience. They should serve as an introduction to several areas of mathematics which are frequently omitted from the curriculum of a computer science student, but which are often used in modern computer science and discrete mathematics.

The introduction is rather condensed, we cover two or three chapters per term, in an upper level undergraduate class. However, it should give the reader enough courage to learn more – either from the literature we reference, or XXX

The authors are grateful to many readers who helped to find mistakes in earlier versions of this text: Vojta Tůma, ... Teaching this material to a group of eager students at Charles University was great motivation in writing these notes.

# Chapter 1

# Harmonic analysis

The word 'harmonic' in the name of this field goes back to analysis of sounds, which were considered harmonious if they were multiples of a basic frequency. Most people have heard of Fourier analysis, either because of the use in compressing sound recordings, or more classical use (which was the original motivation for Joseph Fourier) for solving differential equation for heat dissipation.

While these topics definitely belong to harmonic analysis, and we will return to them later in Section 1.5, our point of view will be much different. We will study the vector space of (complex-valued) functions on a given group (mostly finite abelian). We will observe that choosing the right basis for this vector space will be of great advantage – transforming to this new base is what is frequently called the Fourier transform (and denoted by $\hat{f}$ for a function $f$). The classical analogue of this is expressing periodic functions $[0, 2\pi] \to \mathbb{R}$ as a sum of functions $\sin nx$, $\cos nx$; we will meet this example later as a special case. However, concentrating on the finite case first will make many things easier, in particular convergence of infinite series will not be an issue.

Before we get to any detail, let us outline some of the highlights of this chapter.

- Given a function $f : \mathbb{Z}_2^n \to Z_2$, how fast can you check if it is linear?
  We suppose that $f$ is given as a blackbox – we can evaluate it at any point but have no information about its structure. The obvious algorithm (go over all pairs $x$, $y$ and check if $f(x+y) = f(x) + f(y)$ is much too slow. In 1993 Blum, Luby and Rubinfeld [**?**] found a clever way around: if we are able to tolerate a small probability of mistake, we can do this in linear time. (In fact very simply, by checking several randomly chosen pairs $x$, $y$.) We will analyze this algorithm in Section 1.3.1 by looking at Fourier coefficients of the function $(-1)^{f(x)}$.

- Given a set $A \subseteq \mathbb{Z}_n$, when can we say that $A$ must contain a 3-term arithmetic progression?
  Erdős and Turán conjectured in 1936, that any set $A$ of 'positive density' will work, even more generally: for any integer $k$, real $d > 0$ and all sufficiently large $n$ every set of at least $dn$ elements of $\mathbb{Z}_n$ contains a $k$-term arithmetic progression. In 1953 Roth proved this for $k = 3$ and we will see the proof in Section 1.3.2. (Later, in 1975, Szemerédi proved the conjecture for every $k$, developing the now famous regularity lemma in the process.) We will only prove the case of 3-term arithmetic progressions in $\mathbb{Z}_3^n$, as this illustrates better our central theme.

- Consider a function $f : \{0, 1\}^n \to \{0, 1\}$. We will think of it as of a voting system: $n$ voters try to decide about a "yes/no" question, $f$ says how are their votes combined. Some voting systems are rather extreme (say, constant 0, or the *dictatorship* $f(x) = x_1$), some are more balanced (take the majority among inputs). To measure the influence of each voter for various systems we define $\mathrm{Inf}_i(f)$ to be the probability that $x_i$ has an influence on $f$, when the other coordinates are chosen randomly. A natural question is to measure how small can the maximal (or average) influence be (and what effect it has on the system). Again, we will solve the question by considering Fourier transform of $f$, following Kahn, Kalai, Linial, and Friedgut.

- Among applications of harmonic analysis on infinite groups we will briefly mention Hurwitz' proof of the isoperimetric inequality and a particularly elegant proof of the central limit theorem.

We close this introduction by a simple motivation. A real function $f$ can be even (if $f(-x) = f(x)$) or odd (if $f(-x) = -f(x)$). Obviously, most functions are none of these. However, every function can be written as a sum of two functions, one of them even (namely $\frac{f(x)+f(-x)}{2}$) and the other odd (namely $\frac{f(x)-f(-x)}{2}$). Moreover, this decomposition is unique.

Compared to the above plans this is more of a toy, but hopefully it illustrates the ideas behind what we will be doing.

## 1.1 Characters

We assume that the reader is familiar with the definition of abelian group, vector space, scalar product, and norm. We will be mostly interested in "usual" groups: reals $\mathbb{R}$, complex numbers $\mathbb{C}$, integers $\mathbb{Z}$. One slightly unusual group is the so-called *torus*[1] $T = \{z \in \mathbb{C} : |z| = 1\}$, with multiplication as the operation and (obviously) 1 as the identity. An isomorphic group is $\mathbb{R}/\mathbb{Z}$ (real numbers, where we identify numbers that differ by an integer – reals modulo 1, if you wish). The natural isomorphism $e : \mathbb{R}/\mathbb{Z} \to \mathbb{T}$ is defined by $e(x) = e^{2\pi i x}$. The group $\mathbb{R}/\mathbb{Z}$ is the natural place to study periodic functions of a real variable, which is the topic of classical Fourier analysis.

We will study mainly finite abelian groups,[2] most frequently the group $\mathbb{Z}_n := \mathbb{Z}/n\mathbb{Z}$, integers modulo $n$. We shall assume the reader is familiar with computation in this group. Fortunately, we don't need much more: every finite abelian group is isomorphic to a product

$$\prod_{i=1}^{k} \mathbb{Z}_{n_i} \tag{1.1}$$

for some integer $k \geq 1$ and sequence of integers $n_1, \ldots, n_k \geq 1$.

We will study functions on abelian groups. A particularly important class of such functions are characters.

**Definition 1.1.1.** *Let $G = (G, +, 0)$ be a finite abelian group. We say that $\chi : G \to \mathbb{T}$ is a* character *of $G$ if it is a group homomorphism, that is, if*

1. *$\chi(0) = 1$ and*
2. *$\chi(x + y) = \chi(x)\chi(y)$ for all $x, y \in G$.*

*We will denote the set of all characters of $G$ by $\widehat{G}$.*

[1] Although the geometrically this is a circle, while the geometric torus – tube – is actually $\mathbb{T}^2$.

[2] Infinite groups will be discussed in Section 1.5; the main theme is the same, but there are substantial technical difficulties as we need to somehow deal with "infinite sums". For nonabelian groups things change even more. It turns out that the right setting for this case is to replace $\mathbb{C}$ by $\mathrm{GL}(V)$ for some vector space $V$. (When $V$ is 1-dimensional, we recover the abelian case.) This is the start of *Representation theory*, we will discuss it in Chapter **??**.

**Notes:** For a finite group it is enough to require that $\chi$ maps $G$ to $\mathbb{C} \setminus \{0\}$, see Exercise **??**. For infinite groups, on the other hand, we will add the assumption that $\chi$ is continuos – but this must wait till Section 1.5.

**Theorem 1.1.2** (Pontryagin dual)**.** *The set $\widehat{G}$ forms an abelian group – with pointwise multiplication as the operation and the constant* 1 *function (denoted by $1_G$) as the identity element. The group $\widehat{G}$ is called the* Pontryagin dual of $G$*, and the character $1_G$ is the* trivial character[3]*.*

*Proof.* Obviously, $1_G$ satisfies the conditions for a character. Properties like commutativity, associativity follow immediately from those of $(\mathbb{T}, \cdot, 1)$. Thus all we need to do is to verify that for characters $\chi, \xi$ the mapping $\alpha = \chi\bar{\xi}$ is also a character. (Note that $\bar{\xi}\xi = 1_G$ and thus $\bar{\xi}$ is the group inverse of $\xi$ whenever $\bar{\xi}$ is a character.) Consider $x, y \in G$ and calculate:

$$\begin{aligned}
\alpha(x+y) &= \chi(x+y)\overline{\xi(x+y)} && \text{by the definition of } \alpha \\
&= \chi(x)\chi(y)\overline{\xi(x)\chi(y)} && \text{since } \chi,\ \xi \text{ are characters} \\
&= \chi(x)\overline{\xi(x)}\chi(y)\overline{\xi(y)} \\
&= \alpha(x)\alpha(y)\,.
\end{aligned}$$

□

Now we can say more precisely what is the main theme of harmonic analysis: to study function in the vector space $\mathbb{C}^G$ by expressing them as a linear combination of characters. To do this, we will need to find all of the characters. We start with several definitions to simplify notaion.

Let $f, g$ be functions $G \to \mathbb{C}$.

- *Expectation of $f$* is the expected value $f(x)$ when $x$ is a uniformly random element of $G$. As we have a finite group $G$, we can write simply[4]

$$\mathbb{E}[f] = \mathbb{E}_{x\in G}[f(x)] := \frac{1}{|G|}\sum_{x\in G} f(x)$$

- We can use expectation to define a slightly unusual *scalar product.*

$$\langle f, g\rangle := \mathbb{E}[f\overline{g}] = \frac{1}{|G|}\sum_{x\in G} f(x)\overline{g(x)}$$

  It is easy to verify that this is indeed a scalar product – it satisfies $\langle cf, g\rangle = c\langle f, g\rangle$, $\langle f_1 + f_2, g\rangle = \langle f_1, g\rangle + \langle f_2, g\rangle$, and, finally, $\langle f, g\rangle = \overline{\langle g, f\rangle}$; the word for these properties is *sesquilinear.* The difference from the usual scalar product is in the $\frac{1}{|G|}$ factor.

- The $L_2$ norm is defined in the usual way from the above-defined scalar product.

$$\|f\|_2 := \sqrt{\langle f, f\rangle} = \sqrt{\mathbb{E}_x[|f(x)|^2]}$$

- At times we will also use other $L_p$-norms. Again, it differs from the usual definition only by a constant factor – so the truth of triangle inequality follows from triangle inequality for the usual $L_p$ norm.

$$\|f\|_p := \sqrt[p]{\mathbb{E}_x[|f(x)|^p]}$$

[3] Also called principal character or unit character.

[4] For infinite groups we will replace sums by integrals, see Section 1.5.

We will denote by $L_2(G)$ the vector space $\mathbb{C}^G$ with the scalar product and norm defined above.

**Lemma 1.1.3.** *Let $G$ be a finite abelian group and $\chi$ a nontrivial character of $G$. Then $\mathbb{E}[\chi] = 0$.*

*Proof.* Take any $y \in G$ and observe that

$$\chi(y)\mathbb{E}_{x\in G}[\chi(x)] = \mathbb{E}_{x\in G}[\chi(y)\chi(x)] = \mathbb{E}_{x\in G}[\chi(y+x)] = \mathbb{E}_{x\in G}[\chi(x)] \,.$$

The first equality is the distributivity: $\chi(y)$ is a constant. The second one is part of the definition of a character and the last one follows, because as $x$ goes over all elements of $G$, so does $y + x$.

By rearranging we have

$$(\chi(y) - 1)\mathbb{E}[\chi] = 0 \,.$$

It follows that either $\chi(y) = 1$ for every $y \in G$ (which cannot be the case by the assumptions) or $E\chi = 0$. □

**Lemma 1.1.4.** *The characters are an orthonormal set of vectors in $L_2(G)$. Explicitly:*

1. *$\|\chi\|_2 = 1$ for every $\chi \in \widehat{G}$ and*
2. *$\langle\chi, \xi\rangle = 0$ for every $\chi \neq \xi \in \widehat{G}$.*

*Proof.* The first part is easy: for every $\chi \in \widehat{G}$ and $x \in G$ we have $|\chi(x)| = 1$ and thus also $\mathbb{E}_x|\chi(x)|^2 = 1$. For the second part, let $\chi, \xi$ be two distinct characters. Theorem 1.1.2 implies that $\chi\xi^{-1}$ is also a character; moreover $\chi\xi^{-1} \neq 1_G$. Lemma 1.1.3 finishes the proof. □

**Examples:** 1. Let $G = \mathbb{Z}_n$. For $a, x \in G$ denote

$$\chi_a(x) := e\left(\frac{a \cdot x}{n}\right) = e^{2\pi i \frac{ax}{n}} \,.$$

(Tricky question: does the product in $a \cdot x$ mean the product in $\mathbb{Z}_n$ or in $\mathbb{Z}$?) For every $a \in G$, $\chi_a$ is a character of $\mathbb{Z}_n$. (Exercise!)

2. Let $G = Z_2^n$. For $S \subseteq \{1, \ldots, n\}$ and $x \in G$ we let

$$\chi_S(x) := (-1)^{\sum_{i\in S} x_i} \,.$$

For every $S \subseteq [n]$, $\chi_S$ is a character of $\mathbb{Z}_2^n$. (Exercise!)

We invite the readers to stop now and discover for themselves what are the characters for a general finite abelian group, that is, for a group of form (1.1). For the impatient, the answer is given away in the exercises at the end of this section.

**Exercises:**

1. Let $G$ be group $\prod_{i=1}^k \mathbb{Z}_{n_i}$ for some $k$ and $n_i > 1$. Show that every character is of form

$$\chi_a(x) = e(\sum_{i=1}^{k} \frac{a_i x_i}{n_i}) \tag{1.2}$$

for some $a \in G$. Moreover, the mapping $a \mapsto \chi_a$ is 1–1.

## 1.2 Fourier transform

The following definition will be our main topic in this chapter.

**Definition 1.2.1.** *Let $G$ be a finite abelian group, and let $f$ be a mapping $G \to \mathbb{C}$. The* Fourier transform *of $f$ is the mapping $\widehat{f} : \widehat{G} \to C$ defined by*

$$\widehat{f}(\chi) := \langle f, \chi \rangle = \frac{1}{|G|} \sum_{x \in G} f(x) \overline{\chi(x)} \,.$$

**Class 2 (RS) Oct 10, 2013**

**Theorem 1.2.2.** $|G| = |\widehat{G}|$ *for any finite abelian group $G$.*

*Proof.* Lemma 1.1.4 implies that $|\widehat{G}| \leq |G|$, while the fact that every mapping of form (1.2) is a character (and these characters are distinct) proves the other inequality. □

Before we get any further, one remark about notation. As we saw in Exercise 1 groups $G$ and $\widehat{G}$ are of the same size, there is even a natural bijection given by (1.2). For this reason some authors identify $\widehat{G}$ with $G$ using this bijection and consider the also the Fourier transform of $f$ as a function on $G$: so they write $\widehat{f}(a)$ where we write $\widehat{f}(\chi_a)$. While this is very pleasent notationally, it is also somewhat misleading. In particular, we note here that for infinite groups, $G$ and $\widehat{G}$ may not be isomorphic (see Section 1.5). Also, the bijection $a \mapsto \chi_a$ is not canonical – in the sense that we may just as naturally consider bijection $a \mapsto \chi_{-a}$ or possibly other.

**Observation 1.2.3.** $\widehat{f}(\chi_0) = \mathbb{E} f$

**Theorem 1.2.4** (Inverse Fourier transform)**.** *Let $G$ be a finite abelian group, $f \in \mathbb{C}^G$ any function. Then*

$$f = \sum_{\chi \in \widehat{G}} \widehat{f}(\chi) \chi$$

*(Note that $f$ and $\chi$ are functions $G \to \mathbb{C}$, while $\widehat{f}(\chi)$ are (complex) coefficients.)*

*Proof.* This is a basic linear algebra result, but we prove it for completeness. Put $g = f - \sum_\chi \widehat{f}(\chi)\chi$. For any character $\xi$ we have $\langle g, \xi \rangle = 0$, as the characters are orthonormal. Because the characters generate the space $L_2(G)$ (as there is enough of them and they are orthonormal), it follows that $g = 0$. □

**Theorem 1.2.5** (Plancherel theorem)**.** *For any $f, g \in L_2(G)$ we have*

$$\langle f, g \rangle = \sum_{\chi \in \widehat{G}} \widehat{f}(\chi) \overline{\widehat{g}(\chi)} \,.$$

Note that the right-hand side of Theorem 1.2.5 defines a scalar product on $\widehat{G}$— the "usual one", different from the scalar product we defined on $G$; another good reason to distinguish carefully between $G$ and $\widehat{G}$.

**Theorem 1.2.6** (Parseval theorem)**.** *For any $f \in L_2(G)$ we have*

$$\|f\|_2 = \sqrt{\sum_{\chi \in \widehat{G}} |\widehat{f}(\chi)|^2} \,.$$

As above we note, that the right-hand side of Theorem 1.2.6 defines a norm product on $\widehat{G}$—the "usual $L_2$–norm".

**Theorem 1.2.7.** *The mapping $\mathcal{F} : f \mapsto \widehat{f}$ (the Fourier transform) is linear, invertible, preserves scalar products (isometry)*

- $\|\widehat{f}\|_\infty \leq \|f\|_1$
- *convolution* $(f*g)(a) := \mathbb{E}_{x\in G} f(x)g(a-x) = E_{x,y}[f(x)g(y) \mid x+y=a]$ Notes about importance, etc. Examples:

  1. For a $S \subseteq G$ consider the characteristic function $1_S$ on $G$. Then for any $f \in L_2(G)$ we have
  $$(f * 1_S)(a) = E_x f(x) 1_S(a-x) = \frac{1}{|G|} \sum_{x\in a-S} f(x)\,.$$
  So convolution with a characteristic function provides an averaging operator on functions.
  2. Consider sets $A, B \subseteq G$. Then
  $$(1_A * 1_B)(x) = \mathbb{E}_{a,b}[1 \mid a+b=x] = \frac{\text{the number of solutions to } a+b=x}{|G|^2}\,.$$

- properties: $f*g = g*f$, $f*(g*h) = (f*g)*h$, linear, etc.
- properties: $\widehat{f*g} = \widehat{f}\widehat{g}$

**Exercises:**

1. Let $a$ a fixed element of $G$, $c$ a fixed complex number. We define operators $T_a$ and $S_c$ to map a function $f : G \to \mathbb{C}$ to another function defined as follows

   (a) $(T_a f)(x) = f(x+a)$
   (b) $(P_c)f(x) = cf(x)$
   (c) In case $G$ is a field, we also define $(S_a)f(x) = f(ax)$

   Prove that

   (a) $\widehat{T_a f}(\chi_y) = \chi(a)\widehat{f}(\chi_y)$
   (b) $\widehat{P_c f}(\chi_y) = c\widehat{f}(\chi_y)$
   (c) $\widehat{S_c f}(\chi_y) = \widehat{f}(\chi_{y/c})$

2. Let $G$ be a group and $S \subseteq G$ its symmetric subset ($S = -S$). *Cayley graph* $Cay(G,S)$ is the graph $(G,E)$, where $ab \in E$ whenever $b - a \in S$.

   Let now $G$ be finite and abelian, let $\chi$ a character of $G$ and suppose $0 \notin S$. Furhter, let $M$ be normed adjacency matrix: $M_{i,j} = 1/|S|$ if $ij$ is an edge, $M_{i,j} = 0$ otherwise.

   - Consider a vector $\mathbf{x} \in \mathbb{C}^{|G|}$ such that $x_a = \chi(a)$. Prove that $\mathbf{x}$ is an eigenvecotr of $Cay(G,S)$ (i.e., of matrix $M$).

     Using the preivous exercise use all eigenvalues of the following graphs:
   - $C_n$ (a cycle with $n$ vertices, **[1]**
   - $Q_d$ (a $d$-dimensional hypercube): $V(H_d) = \{0,1\}^d$ and $ab$ is an edge whenever $|a-b| = 1$. **[2]**

3. * Show that the graphs from the prevous exercise show the tightness of the following estimates: valid for a $d$-regular graph $G=(V,E)$.

$$\frac{1-\lambda_2}{2} \le h(G) \le \sqrt{2(1-\lambda_2)}\,,$$

where $\lambda_2$ is the second largest eigenvalue of $M$ and

$$h(G) = \min_{S \subseteq V} \frac{|E(S, V-S)|}{d \min(|S|, |V-S|)},$$

4. Find the matrix of linear mapping given by Fourier transform on $\mathbb{Z}_n$. Explicitly, find a matrix $M_n$ such that for every $f : \mathbb{Z}_n \to \mathbb{C}$ we have

$$(\widehat{f}(\chi_0), \ldots, \widehat{f}(\chi_{n-1}))^T = M_n (f(0), \ldots, f(n-1))^T\,.$$

Compute $\det M_n$ and conclude that Fourier transform is a bijection.

5. We define *support* of a mapping $f\colon G \to \mathbb{C}$ to be the set $\mathrm{Supp}(f)$ of all $x \in G$ for which $f(x) \neq 0$. Let $G$ be a finite abelian group and $f, g\colon G \to \mathbb{C}$. Prove the following:

   - $\mathrm{Supp}(f * g) \subseteq \mathrm{Supp}(f) + \mathrm{Supp}(g)$, **[1]**
   - $\|f * g\|_\infty \le \|f\|_p \cdot \|g\|_q$, kde $1/p + 1/q = 1$, **[2]**
   - $\|f * g\|_1 \le \|f\|_1 \cdot \|g\|_1$
   - $\widehat{f \cdot g}(\chi) = \sum_{\zeta \in \widehat{G}} \widehat{f}(\chi - \zeta)\widehat{g}(\zeta)$. **[1]**

Further exercises: examples of inverse transform, ...

**Class 3 (RS) Oct 17, 2013**

## 1.3 Two applications

### 1.3.1 Linearity testing

Suppose we have a black-box that computes a function $f : \mathbb{Z}_2^n \to \mathbb{Z}_2$ and wish to test whether $f$ is linear, that is, whether $f(0) = 0$ and $f(x+y) = f(x) + f(y)$. Obviously, we want to do this as fast as possible, say by evaluating $f$ as few times as possible. If we have no prior knowledge about $f$, then we need to test $f(x)$ for every $x \in \mathbb{Z}_2^n$ to be certain. As $2^n$ is often too high number, we may want to try some sort of approximation. An interesting variant was suggested by Blum, Luby and Rubinfeld in 1993 [**?**]. They allowed for a small probability of error, and did not care about border-line cases: mappings that are close to being linear. This approach lead to the development of so-called *property testing*, which is now an important research field.

As a historic remark, the original motivation for Blum, Luby and Rubinfeld was to study generally how to test correctness of a computer implementation of some function. However, few years later this became an important step in the proof of the famous PCP theorem [**?**].

**Theorem 1.3.1.** *Let $f : \mathbb{Z}_2^n \to \mathbb{Z}_2$ be a mapping, $\varepsilon, \delta > 0$. There is a linear-time*[5] *(randomized) algorithm that evaluates $f$ at $\lceil \frac{\log \delta}{\log(1-\varepsilon)} \rceil$ points with the following properties:*

[5] We do not count the time spend by evaluating the function $f$.

- *If $f$ is linear, the algorithm always confirms it.*
- *If $f$ is such that every linear mapping $g$ differs from $f$ on at least $\varepsilon 2^n$ points then the algorithm finds that $f$ is not linear with probability at least $1-\delta$.*

*Proof.* Our algorithm is as simple as it could be: we repeatedly test the linearity. Specifically, take $N$ large enough (based on $\varepsilon$ and $\delta$. We will $N$-times choose random $x, y \in \mathbb{Z}_2^n$ (all choices independent and uniform), test whether $f(x+y) = f(x) + f(y)$. If we find no counterexample, we declare $f$ linear.

The algorithm behaves as it should for linear mappings. To analyze it's behaviour for nonlinear ones, let us denote $q = \Pr_{x,y\in\mathbb{Z}_2^n}[f(x+y) = f(x)+f(y)]$ the probability we are fooled at one step. The probability that the algorithm declares $f$ linear equals to $q^N$. Our goal is to upper bound $q$ for mappings that are $\varepsilon$-far from every linear mapping.

To achieve that we define mapping $F(x) = (-1)^{f(x)}$. It is easy to see, that $f$ is linear if and only if $F$ is a character; also the number of points at which $f$ needs to be changed to become linear coincides with the number of points where $F$ needs to be changed to become a character. For a character $\chi$ of $\mathbb{Z}_2^n$ we put $p(\chi) = \Pr_x[F(x) \neq \chi(x)]$. We only care about behaviour of the algorithm when the given mapping satisfies $p(\chi) \geq \varepsilon$ for every character $\chi$.

We will consider the Fourier transform of $F$, and use several times the basic formula

$$F(x) = \sum_{a\in\mathbb{Z}_2^n} \widehat{F}(\chi_a)\chi_a(x)\,. \tag{1.3}$$

We start by observing, that (writing $p = p(\chi)$)

$$\widehat{F}(\chi) = \mathbb{E}_x F(x)\overline{\chi(x)} = 1\cdot(1-p) + (-1)\cdot p = 1-2p \leq 1-2\varepsilon\,.$$

Next, we express similarly $q$. We will again use the easy fact that two numbers in $\{\pm 1\}$ are equal if their product is 1, otherwise the product is $-1$. Consequently

$$\mathbb{E}_{x,y}[F(x+y)F(x)F(y)] = 1\cdot q + (-1)\cdot(1-q) = 2q-1\,.$$

On the other hand, using (1.3) we obtain

$$\begin{aligned}\mathbb{E}[F(x+y)F(x)F(y)] &= \mathbb{E}_{x,y}\sum_{a\in\mathbb{Z}_2^n}\widehat{F}(\chi_a)\chi_a(x+y)\sum_{b\in\mathbb{Z}_2^n}\widehat{F}(\chi_b)\chi_b(x)\sum_{c\in\mathbb{Z}_2^n}\widehat{F}(\chi_c)\chi_c(y)\\ &= \sum_{a,b,c}\widehat{F}(\chi_a)\widehat{F}(\chi_b)\widehat{F}(\chi_c)\mathbb{E}_x\chi_a(x)\chi_b(x)\mathbb{E}_y\chi_b(y)\chi_c(y)\end{aligned}$$

To further simplify, we observe that (here, as in every finite abelian group, with our notation for characters)

$$\chi_a(x) = \chi_x(a)\,.$$

Thus, part of the above sum can be written as

$$\chi_x(a)\chi_x(b) = \chi_x(a+b) = \chi_{a+b}(x)\,.$$

Using Lemma 1.1.3 we find, that the above sum only has nonzero terms for $a+b=0$ and $b+c=0$ (dealing with the second product of characters similarly). For those values, the expectation equals to 1, yielding

$$\mathbb{E}[F(x+y)F(x)F(y)] = \sum_a \widehat{F}(\chi_a)^3\,.$$

Finally, we recall that $\widehat{F}(\chi_a) \leq 1 - 2\varepsilon$ and write

$$\mathbb{E}[F(x+y)F(x)F(y)] \leq (1-2\varepsilon)\sum_a \widehat{F}(\chi_a)^2 = (1-2\varepsilon)\|F\|_2^2 \,.$$

We are using the fact, that $\widehat{F}(\chi_a)$ is a real number (thus its square is nonnegative) and we use Parseval theorem.[6] As $F$ only attains value ±1, it's norm is 1. To wrap up, we obtained inequality $2q-1 \leq 1-2\varepsilon$, thus $q \leq 1-\varepsilon$. It remains to choose $N$ so that $(1-\varepsilon)^N \leq \delta$, so $\delta \geq \frac{\log \delta}{\log(1-\varepsilon)}$. For fixed $\varepsilon$ and $\delta$, the running time is linear in $n$ — we need to generate $O(n)$ random bits, the evaluation of $f$ is not counted. □

### 1.3.2 Arithmetic progressions

Our next application is from additive number theory. The history starts with Roth theorem (which once was a special case of Erdős-Turán conjecture, that was finally settled by the Szemerédi theorem): every sufficiently large set $A \subseteq [N]$ contains an arithmetic progression of length 3. What exactly means sufficiently large (i.e., what is the best bound) is not clear yet (see notes TODO), but for every $\delta > 0$ and sufficiently large $N$, $|A| \geq \delta N$ is enough).

To present the main idea of the proof we present analogous result for the group $\mathbb{Z}_3^n$, that was proved by Meshulam using Roth's ideas. (Note that the same proof works for group $\mathbb{Z}_p^n$ for any prime $p$.)

**Theorem 1.3.2.** *Every set $A \subseteq \mathbb{Z}_3^n$ of size $A \geq c\frac{3^n}{n}$ contains an arithmetic progression of length* 3.

*Proof.* With a slight abuse of notation, let $A(x)$ be the characteristic function of $A$. We will use Fourier transform of $A$ to either find an arithmetic progression of length 3 (an AP$_3$) (in fact many of them), or to find a character "along which $A$ oscillates a lot". This will allow us to find a hyperplane in which $A$ has higher density than in the whole $\mathbb{Z}_3^n$, and using induction will finish the proof.

To start slowly, we observe that $\widehat{A}(0)$ is[7] the density $\frac{|A|}{3^n}$ of the set $A$, we put $\delta := \frac{|A|}{3^n}$. By assumption, $\delta \geq \frac{c}{n}$.

The key quantity for the proof is the "density of AP$_3$'s"

$$T(A) = \mathbb{E}_{x,d} A(x-d)A(x)A(x+d)\,. \tag{1.4}$$

The number of AP$_3$'s is $T(A)|\mathbb{Z}_3^n|^2$. Of those, many are trivial: the expression for $T(A)$ allows $d = 0$, thus it counts $3^n$ progressions of form $x, x, x$. It follows, that to prove there is a nontrivial AP$_3$ we need to show

$$T(A) \cdot 3^{2n} > 3^n \,.$$

To this end, we plug the Fourier expansion of $A$ to Equation (1.4) and use tricks

[6] Observe that it is not enough to use the bound $\widehat{F}(\chi)^3 \leq (1-2\varepsilon)^3$ for every $\chi$, as we get a factor of $2^n$.

[7] in this proof we use a convenient shortcut $\widehat{A}(a)$ to stand for $\widehat{A}(\chi_a)$.

similar to those in the above proof of Theorem 1.3.1.

$$\begin{aligned}
T(A) &= \mathbb{E}_{x,d} \sum_{a\in\mathbb{Z}_3^n} \widehat{A}(a)\chi_a(x-d) \sum_{b\in\mathbb{Z}_3^n} \widehat{A}(b)\chi_b(x) \sum_{c\in\mathbb{Z}_3^n} \widehat{A}(c)\chi_c(x+d) \\
&= \sum_{a,b,c\in\mathbb{Z}_3^n} \widehat{A}(a)\widehat{A}(b)\widehat{A}(c)\mathbb{E}_x\chi_a(x)\chi_b(x)\chi_c(x)\mathbb{E}_d\chi_a(-d)\chi_c(d) \\
&= \sum_{a,b,c\in\mathbb{Z}_3^n} \widehat{A}(a)\widehat{A}(b)\widehat{A}(c)\mathbb{E}_x\chi_{a+b+c}(x)\mathbb{E}_d\chi_{c-a}(d) \\
&= \sum_{a\in\mathbb{Z}_3^n} \widehat{A}(a)^2\widehat{A}(-2a)
\end{aligned}$$

To further estimate this, we put $M = \max_{b\neq 0} |\widehat{A}(b)|$. We use triangle inequality and Parseval theorem to obtain

$$|T(A)-\delta^3| = \left| \sum_{0\neq a\in\mathbb{Z}_3^n} \widehat{A}(a)^2\widehat{A}(-2a) \right| \le \sum_{0\neq a\in\mathbb{Z}_3^n} |\widehat{A}(a)^2| \cdot M \le \delta M\,.$$

We distinguish two cases:

**Case 1:** $M \le \delta^2/2$ We conclude $T(A) \ge \delta^3 - \delta M \ge \delta^3/2$. If $3^n\delta^3/2 > 1$, we are done; so this requires $3^n c^3/n^3 > 2$.

Vaguely speaking, if all (nontrivial) Fourier coefficients of $A$ are rather small, then $A$ resembles random set, or a function that is a constant $\delta$. (For both of these cases, the (expected) value of $T(A)$ is exactly $\delta^3$, in our case it is at least half of that.)

**Case 2:** $M \ge \delta^2/2$ In this case the goal is to utilize the large Fourier coefficient to find a hyperplane $H$ such that $|A\cap H|/|H| \ge \delta + \delta^2/4$. This will enable us to use induction to finish the proof.

To start with the details, let $a \in \mathbb{Z}_3^n$ be such that $|\widehat{A}(\chi_a)| \ge \delta^2/2$ and $a \neq 0$. We recall that $\chi_a(x) = e(\sum_i a_i x_i/3)$. We write $a\cdot x = \sum_i a_i x_i$ and $\omega = e(1/3)$. With this notation, we may write $\widehat{A}(\chi_a)$ in a succint form

$$\begin{aligned}
\widehat{A}(\chi_a) &= \mathbb{E}_x A(x)\overline{\chi_a(x)} \\
&= \mathbb{E}_x A(x)\omega^{-a.x} \\
&= \mathbb{E}_{c\in\mathbb{Z}_3}\mathbb{E}_{x:a.x=c}A(x)\omega^c \\
&= \mathbb{E}_{c\in\mathbb{Z}_3}\omega^c\alpha_c\,,
\end{aligned}$$

where $\alpha_c := |\{x\in A : a\cdot x = c\}|/3^{n-1}$ is the density of $A$ in a certain hyperplane $H_c$. To recap, we have

$$\tfrac{1}{3}|\sum_{c=0}^{2}\alpha_c\omega^c| \ge \frac{\delta^2}{2} \qquad \text{and} \qquad \tfrac{1}{3}\sum_{c=0}^{2}\alpha_c = \delta \tag{1.5}$$

(the second equality is the density of $A$ expressed in two ways). By rewriting this slightly we get

$$\tfrac{1}{3}\sum_{c=0}^{2}(\alpha_c-\delta) = 0 \qquad \text{and}$$

$$\tfrac{1}{3}|\sum_{c=0}^{2}(\alpha_c-\delta)\omega^c| \ge \frac{\delta^2}{2} \qquad (\text{as } \textstyle\sum_c \omega^c = 0).$$

Using the triangle inequality and adding the two equations we obtain

$$\tfrac{1}{3}\sum_c |\alpha_c - \delta| + (\alpha_c - \delta) \geq \delta^2/2$$

and, consequently, for some $c$ we have $|\alpha_c - \delta| + (\alpha_c - \delta) \geq \delta^2/2$, and thus $\alpha_c \geq \delta + \delta^2/4$, as promised.

Now we get to the big picture. We plan to use induction. For $n = 1$ the result is true, if $c3^1/1 \geq 1$, or $c \geq 1/3$. For $n \geq 1$ we use the computations above and conclude that either the result is true (in Case 1, if $c \geq XXX$, or that we can apply induction assumption.[8]. For this we need that

$$\delta + \delta^2/4 \geq \frac{c}{n-1}$$

or $c \geq n/(n-1)$. Putting all conditions together, $c = 8$ in the statement is sufficient. □

**Notes:** 1) Another way to state the proof is perhaps intuitively more convincing: if we don't find the desired progression by direct counting, we find a subspace where the set $A$ is denser that in the whole space. If we repeat this long enough, we reach the condition that the density of $A$ in some set is more than 1, which is a contradiction.

2) To provide motivation for the slight trick where we replaced $\alpha_c$ by $\alpha_c - \delta$, consider a function $B(x) = A(x) - \delta$. This function has average 0, thus should be easier to work with. However, $\widehat{A}(\chi_a) = \widehat{B}(\chi_a)$ for every $a \neq 0$. Anyway, the fact that Equation (1.5) implies that one of $\alpha_c$ is quite a bit larger than $\delta$ should be obvious, at least on qualitative level: the average of $\alpha_c$ is $\delta$, and if all of them were equal, then the first sum in Equation (1.5) would be zero.

3) The original Roth's theorem speaks about arithmetic progressions in a starting segment of $\mathbb{Z}$. It uses the same main idea: small Fourier coefficients imply many arithmetic progressions, while at least one large coefficient will enable us to find a place, where $A$ is denser than in the whole set. However, one cannot use Fourier tranform over the infinite set $\mathbb{Z}$, and also using Fourier transform over $\mathbb{Z}_N$ (with the set contained in $\{1, \ldots, N\}$) is not easy: if $N$ is even, we may find that $A$ is very dense on a subset of two points. The hints to solve these issues are given in the exercises.

4) TODO: add HA approach to Szemerédi's theorem, Gowers' uniformity norm – "Fourier uniformity" and GI conjectures?

5) TODO: add PAC learning?

**Class 4 (RS) Oct 24, 2013**

### 1.3.3 Some more theory

We start with a motivating example. Let $G = \mathbb{Z}_k \times \mathbb{Z}_l$, consider characteristic function $1_S$ of a set $S = \{(x, 0) : x \in \mathbb{Z}_k\}$. We may also write this function explicitly using the Dirac notation

$$1_S(x, y) = \delta_0(y)\,.$$

[8] Here we use the fact that $\mathbb{Z}_3$ is a field, thus each hyperplane in $\mathbb{Z}_3^n$ is isomorphic to $\mathbb{Z}_3^{n-1}$

Let us compute the Fourier coefficients; we will use the shorthand $\widehat{f}(a) = \widehat{f}(\chi_a)$. By definition,

$$\begin{aligned}\widehat{1_S}(a,b) &= \mathbb{E}_{(x,y)} 1_S(x,y)\chi_{a,b}(x,y)\\ &= \mathbb{E}_{(x,y)}\delta_0(y)\chi_a(x)\chi_b(y)\\ &= \tfrac{1}{l}\mathbb{E}_{x\in\mathbb{Z}_k}\chi_a(x)\\ &= \tfrac{1}{l}\delta_0(x)\,.\end{aligned}$$

Thus, the Fourier coefficients are a multiple of the characteristic function of the set $\{(0,y) : y \in \mathbb{Z}_l\}$. Similar result, for a group $\mathbb{Z}_n$ with $n = kl$ appeared earlier in Exercise **??**. Next, we will present a general definition that captures this phenomenon.

Let $S$ be a subset of a (finite abelian) group $G$. Its *orthogonal complement* is the set

$$S^\perp = \{a \in G : \chi_a(x) = 1 \quad \text{for every } x \in S\}\,.$$

**Lemma 1.3.3.** *Let $G$ be a finite abelian group, let $H$ be a subgroup of $G$. Then*

$$\widehat{1_H}(a) = \begin{cases} \frac{|H|}{|G|} & \textit{if } a \in H^\perp \\ 0 & \textit{otherwise.}\end{cases}$$

*Proof.* Put $E = \widehat{1_H}(a) = \mathbb{E}_x 1_H(x)\overline{\chi_a(x)}$. If $a \in H^\perp$, then the product can be simplified to $1_H(x)$ yielding $E = \mathbb{E}_{x\in G}1_H(x) = \frac{H}{G}$.

If $a \notin H^\perp$, pick some $x_0 \in G$ such that $\chi_a(x_0) \neq 1$. As $x + x_0$ runs over all elements of $G$ when $x$ does, we may write

$$\begin{aligned}E &= \mathbb{E}_x 1_H(x+x_0)\overline{\chi_a(x+x_0)} \\ &= \mathbb{E}_x 1_H(x+x_0)\overline{\chi_a(x)\chi_a(x_0)} && \text{as } \chi_a \text{ is a character}\\ &= \mathbb{E}_x 1_H(x)\overline{\chi_a(x)\chi_a(x_0)} && \text{as } H \text{ is a subgroup}\\ &= E\cdot\overline{\chi_a(x_0)}\,.\end{aligned}$$

By our choice of $x_0$, the last quantity is not equal to $E$ unless $E = 0$, which finishes the proof. □

**Theorem 1.3.4** (Poisson sumation formula)**.** *Let $G$ be a finite abelian group and $H$ a subgroup of $G$. Consider $f : G \to \mathbb{C}$ and $a \in G$. Then*

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

**Notes:** It is instructive to put extreme values of $H$ in this theorem. 1) If $H = \{0\}$ is the trivial group, then $H^\perp = G$ and we obtain the basic inversion formula

$$f(a) = \sum_{y\in G}\widehat{f}(y)\chi_y(a)\,.$$

2) If $H = G$ then $H^\perp = \{0\}$ and we get the often-used result

$$\mathbb{E}_x f(x) = \widehat{f}(0)\,.$$

**Application: MacWilliams identities** TODO expand

- error correcting codes, codes as subsets of $\mathbb{Z}_2^n$ with large pairwise Hamming distance
- polynomial associated to a code
- dual code
- the theorem
- the proof

TODO: more classical application
TODO: explain that this is the special case of Selberg trace formula (?)

**Class 5 (RS) Oct 31, 2013**

## 1.4 Influence of variables – KKL theorem

In this section we will study Boolean functions of many variables. Consider a function $f : \{0,1\}^n \to \{0,1\}$. We want to understand, how much does each of these variables affect the value of the function. The natural way to formalize this is to set $\mathrm{Inf}_i(f)$ to be the probability that $x_i$ has an influence on $f$, when the other coordinates are chosen randomly. Explicitly,

$$\mathrm{Inf}_i(f) = \Pr_{x \in \mathbb{Z}_2^n}[f(x) \neq f(x+e_i)]$$

where $e_i$ is the vector with a single 1 at the $i$-th position. (Note that we identified $\{0,1\}^n$ with $\mathbb{Z}_2^n$ to utilize the group structure. The range of $f$, on the other hand, will be conveniently understood as a subset of $\mathbb{C}$.)

It is obvious, that if $f(x) = \sum_i x_i$ (sum in $\mathbb{Z}_2$) that the influence of each variable is 1. For the *dictatorship* $g(x) = x_1$, the first variable has influence 1, all the rest 0. In the exercises, you are welcome to check that the influences of each variable on the *majority function* (suppose $n$ is odd) is $\Theta(\frac{1}{\sqrt{n}})$.

Hopefully, by now the reader is motivated to find out what can be said for influences of variables on other boolean functions. The most natural question seems to be, how small can be the influence of the "most important" variable. The "sociologic implication" of this is that such function would contitute the most just voting system. To avoid trivialities (a constant function), we want to express the answer in terms of $p = \Pr[f(x) = 1]$. Kahn, Kalai, and Linial **??** proved the following influential result. Among other consequences this sparkled the interest about harmonic analysis among computer scientists.

**Theorem 1.4.1.** *Let $f : \mathbb{Z}_2^n \to \mathbb{Z}_2$ be a function, let $p = \Pr[f(x) = 1] \leq 1/2$. Then*

$$\sum_{i=1}^n \mathrm{Inf}_i(f)^2 \geq \Omega\left(p^2 \frac{(\log n)^2}{n^2}\right) .$$

*Consequently, for some $i$ we have*

$$\mathrm{Inf}_i(f) \geq \Omega\left(p \frac{\log n}{n}\right) .$$

*Proof.* It is immediate that the second part follows from the first one. That will take more effort. We start slowly, by defining $f_i(x) = f(x) - f(x + e_i)$ and noting that

$$\mathrm{Inf}_i(f) = \mathbb{E}|f_i(x)|^2 = \|f_i\|_2^2 \,.$$

An application of Exercise **??** gives us that $\widehat{f_i}(\chi_s) = \widehat{f}(\chi_s) - \chi_s(e_i)\widehat{f}(\chi_s)$. This equals to 0 if $i \notin S$ and to $2\widehat{f}(\chi_s)$ otherwise. Putting all together and using Parseval's theorem, we obtain

$$\begin{aligned}\sum_i \mathrm{Inf}_i(f)^2 = \ldots &= \sum_i \|f_i\|_2^2 \\ &= \sum_i \sum_S \widehat{f_i}(S)^2 \\ &= \sum_i \sum_{S \ni i} (2\widehat{f}(S))^2 \\ &= \sum_S \sum_{i \in S} (2\widehat{f}(S))^2 \\ &= \sum_S |S|(2\widehat{f}(S))^2\end{aligned}$$

TODO: finish this ...

□

Inequalities for norm of convolution. Riesz-Thorin interpolation theorem (without proof)

**Class 6 (RS) Nov 7, 2013**

## 1.5 Infinite groups

In this section we will generalize the theory to infinite groups. We will have to do with a brief introduction to the topic, as the theory is vast, for reasons soon to become clear. The basic idea is still the same: we wish to express function $f : G \to \mathbb{C}$ (with $G$ an abelian group) as certain combination of symmetric functions – characters.

A function $\chi : G \to \mathbb{T}$ is a *character of* $G$ if it is a group homomorphisms (it satisfies the conditions of Definition 1.1.1) and it is continuous.[9] As before, we let $\widehat{G}$ be the set of all characters of $G$ with the obvious group structure.

**Exercises:** Prove the following:

1. $\widehat{\mathbb{Z}} \cong \mathbb{T}$.
2. $\widehat{\mathbb{T}} \cong \mathbb{Z}$.
3. $\widehat{\mathbb{R}} \cong \mathbb{R}$.

Each of these tasks has two parts: one is guessing the characters (and how are they naturally "indexed" by the dual group), and the other is proving that there are no other characters. All of these examples can be obtained from the knowledge of solutions to the so-called Cauchy functional equation – a real function $f$ satisfying $f(x + y) = f(x) + f(y)$. It is easy to prove (as was done by Cauchy in 1821) that all continuous solutions are of form $f(x) = cx$ (with $c$ any real contant).

[9] Every mapping from a finite set is continuous, so we did not have to add this condition before.

TODO: present a solution?

Before returning to the case of general abelian group, we will study the case of functions defined on $\mathbb{T}$ in detail. These functions correspond naturally to periodic functions defined on $\mathbb{R}$, and are the topic that most people learn first when studying Fourier analysis. We will now explain the connection.

As we have seen in the exercises, the characters of $\mathbb{T}$ are functions $\chi_n : \mathbb{T} \to \mathbb{C}$, defined by $\chi_n(t) = e^{2\pi int} = \cos(2\pi int) + i\sin(2\pi int)$. For a function $f : \mathbb{T} \to \mathbb{C}$ the Fourier coefficients are again defined as in Definition 1.2.1. Obviously, the definition of expectation, and thus the scalar product, must be updated. Rather naturally, we put

$$\mathbb{E}f = \int_0^1 f(t)\,\mathrm{d}t$$

and (as before) scalar product is defined by means of expectation

$$\langle f, g\rangle = \mathbb{E}f\overline{g} = \int_0^1 f(t)\overline{g(t)}\,\mathrm{d}t\,.$$

For notational convenience we write $\widehat{f}(n)$ for $\widehat{f}(\chi_n)$: thus we let $\widehat{f}$ be a function defined on $\mathbb{Z}$, instead of the (isomorphic) group $\widehat{\mathbb{T}}$. Thus, we finally get to the formula (*Fourier transformation*)

$$\widehat{f}(n) = \int_0^1 f(t)e^{2\pi int}\,\mathrm{d}t\,. \tag{1.6}$$

As before we may want to express $f$ using $\widehat{f}$ (*inverse Fourier transformation*). Natural formula to expect is

$$f(t) = \sum_{n\in\mathbb{Z}} \widehat{f}(n)e^{2\pi int}\,. \tag{1.7}$$

However, here is an important divergence between the finite and infinite theory. In the finite case, all scalar products are defined (thus we may compute Fourier transformation of any function) and the truth of the inverse Fourier transformation is a simple linear algebra. In the infinite case though, there are integrals/infinite sums involved, and they may not be defined. Indeed, larger part of the Fourier theory is involved with understanding the conditions under which Equations (1.6) and (1.7) stand true – or in what sense are the equations true.

Frequently, we deal with real-valued functions and it may be inconvenient to use complex exponential to express them. It is fortunately easy to go between complex exponentials and combination of sines and cosines. We recall the following formulas:

$$e^{it} = \cos t + i\sin t \qquad \cos t = \frac{e^{it}+e^{-it}}{2} \qquad \sin t = \frac{e^{it}-e^{-it}}{2i}$$

Consequently,

$$\widehat{f}(n) = \int_0^1 f(t)\cos 2\pi nt\,\mathrm{d}t + i\int_0^1 f(t)\sin 2\pi nt\,\mathrm{d}t$$

We may write $\widehat{f}(n) = a_n + ib_n$, where

$$a_n = \int_0^1 f(t)\cos 2\pi nt\,\mathrm{d}t \qquad b_n = \int_0^1 f(t)\sin 2\pi nt\,\mathrm{d}t$$

For the inverse direction, we may write Equation (1.7) in the form

$$f(t) = \sum_{n\in\mathbb{Z}} \widehat{f}(n)(\cos 2\pi nt + i\sin 2\pi nt)\,.$$

Plugging in $a_n + ib_n$ for $\widehat{f}(n)$ and joining the terms with $n$ and $-n$ we get

$$f(t) = a_0 + \sum_{n=1}^{\infty} (a_n + a_{-n}) \cos 2\pi nt + (b_n - b_{-n}) \sin 2\pi nt$$

Finally, if $f$ is real-valued, we have $a_{-n} = -a_n$ and $b_{-n} = -b_n$, and these are real numbers. We get

$$f(t) = a_0 + \sum_{n=1}^{\infty} 2a_n \cos 2\pi nt + 2b_n \sin 2\pi nt\,. \tag{1.8}$$

This is the formula that most people recognize as "the Fourier series", expressing real-valued periodic function in terms of sines and cosines.[10]

After witnessing that Equations (1.7) and (1.8) are expressing the same statement in different terms, we return back to complex exponentials, as they are more elegant and also generalize easier to other groups in place of $\mathbb{T}$.

TODO:

- some convergence theorems
- analogues of finite case – now with a measure
- Fourier transform and derivative (note on the original motivation: how to solve ODE, PDE with Fourier)
- application: isoperimetric inequality
- application: Central Limit Theorem

sketch: We recall again Exercise **??**: Let $f : \mathbb{T} \to \mathbb{C}$, for $h \in T$ get $g_h(t) = f(t+h)$. Then $\widehat{g}_h(n) = \chi_n(h)\widehat{f}(n)$. Plugging in the definition of $\chi_n$ and using linearity of the Fourier transform we get

$$\frac{\widehat{f(t+h) - f(t)}}{h} = \frac{e^{2\pi inh} - 1}{h} \widehat{f}(n)\,.$$

Passing to the limit (Exercise: why can we do this?) we find that

$$\widehat{f'}(n) = 2\pi in \widehat{f}(n)\,.$$

Next, we are going to apply this to prove the well-known isoperimetric inequality: the fact that, given a fixed perimeter, the planar set with the largest area is the circle. In fact, we will prove it with a small restriction: the boundary of the set is a differentiable curve. TODO: more general version? TODO: reference to other proofs

While the theorem itself is ancient (TODO: reference), it took long before a first proof was found. This proof is by Hurwitz **??**.

**Theorem 1.5.1.** *Let $\varphi : [0,1] \to \mathbb{R}^2$ be a simple*[11] *$C^1$ curve with $\varphi(0) = \varphi(1)$. Let $L$ be the length of $\varphi$, let $A$ be the area of the interior of $\varphi$. Then $L^2 \geq 4\pi A$, with equality only if $\varphi$ is a circle.*

[10] Usually, the scaling factors are different – people study $2\pi$-periodic, instead of 1-periodic functions, and there is some scaling factor in the definition of $a_n$, and $b_n$.

[11] no self-intersections

*Proof.* Sketch: Let $\varphi = (x, y)$, express $L$ and $A$ using $x$, $y$, $x'$ and $y'$. Then use Harmonic analysis to show, that

$$L^2 - 4\pi A = 2\pi^2 \left( \sum_{n \neq 0} |n\widehat{x}(n) - i\widehat{y}(n)|^2 + |n\widehat{y}(n) + i\widehat{x}(n)|^2 + (n^2 - 1)(|\widehat{x}(n)|^2 + |\widehat{y}(n)|^2) \right) .$$

As the right-hand side is a sum of squares of real numbers, it is non-negative, thus $L^2 \geq 4\pi A$ as claimed. Analyzing when all summands above are zero gives us infinite system of equation, that imply $(x, y)$ bound a circle. □