18.785: Analytic Number Theory, MIT, spring 2007 (K.S. Kedlaya)
Introduction to large sieve inequalities

In this unit, we consider a relatively simple example of a large sieve inequality, of the sort introduced by Linnik. This is a setup for the multiplicative large sieve inequality we will need for Bombieri-Vinogradov.

1 Overview

The purpose of a "large sieve" is to allow sieving over a range of primes not possible with the traditional sieve methods we considered earlier. The price to be paid is that one only gets results of an aggregate nature. For instance, in the Bombieri-Vinogradov theorem, we will consider the error terms in the prime number theorem in arithmetic progression for *all* moduli in some range, and show that the sum of the errors cannot be too large.

This said, a "large sieve inequality" does not itself involve a sieve, at least not the way we look at these things nowadays; the sieves only appear in the application. The general *large sieve problem*: given a finite set V of vectors $v \in \mathbb{C}^n$, find the smallest constant $C = C(V)$ such that for any vector $x \in \mathbb{C}^n$,

$$\sum_{v \in V} |v \cdot \overline{x}|^2 \leq C x \cdot \overline{x}. \tag{1}$$

(Note that $v \cdot \overline{x}$ is the usual Hermitian inner product.) Of course one has $C \leq \sum_{v \in V} v \cdot \overline{v}$ by Cauchy-Schwarz term by term, but this is nowhere near optimal if the vectors v are pointing in all different directions, as then the vector x cannot simultaneously be nearly parallel to all of them. A trivial example is given by an orthonormal set of vectors, in which case $C = 1$; see the exercises for another simple example.

In number theory applications, we tend to view the same setup as follows. Given a finite set X of complex-valued sequences, and a cutoff N, find a constant $C = C(X, N)$ such that for any $a_n \in \mathbb{C}$,

$$\sum_{x \in X} \left| \sum_{n \leq N} a_n x(n) \right|^2 \leq C \sum_{n \leq N} |a_n|^2.$$

2 An additive large sieve

In the additive large sieve, we take the sequences $x \in X$ to be of the form $\exp(2\pi i \alpha n)$ for some $\alpha = \alpha_x \in \mathbb{R}$ (or better, in $\mathbb{R}/\mathbb{Z}$). In order for these to be "not too parallel", we insist that the corresponding α_x be δ-*spaced* for some $\delta > 0$, i.e., if $x, y \in X$ are distinct, then $\alpha_x - \alpha_y$ must have distance at least δ from the nearest integer. The following inequality is due independently to Selberg, and to Montgomery and Vaughan; it refines a result of Davenport and Halberstam.

Theorem 1. *Fix $\delta \in (0, 1/2]$. Let $S \subset \mathbb{R}$ be a δ-spaced set (necessarily finite). Then for any $a_n \in \mathbb{C}$ for $M < n \leq M + N$,*

$$\sum_{\alpha \in S} \left| \sum_{M<n\leq M+N} a_n \exp(2\pi i \alpha n) \right|^2 \leq (\delta^{-1} + N - 1) \sum_{M<n\leq M+N} |a_n|^2.$$

The key input is the following inequality, a variation of a classic inequality of Hilbert.

Lemma 2. *Let $\lambda_1, \ldots, \lambda_n$ be real numbers with $|\lambda_i - \lambda_j| \geq \delta$ whenever $i \neq j$. Then for any $z_1, \ldots, z_n \in \mathbb{C}$,*

$$\left| \sum_{i\neq j} \frac{z_i \overline{z_j}}{\lambda_i - \lambda_j} \right| \leq \frac{\pi}{\delta} \sum_{i=1}^{n} |z_i|^2.$$

Proof. Exercise. □

Corollary 3. *For $S = \{\alpha_1, \ldots, \alpha_n\}$ a δ-spaced set and $z_1, \ldots, z_n \in \mathbb{C}$,*

$$\left| \sum_{i\neq j} \frac{z_i \overline{z_j}}{\sin \pi(\alpha_i - \alpha_j)} \right| \leq \delta^{-1} \sum_{i=1}^{n} |z_i|^2.$$

Proof. Let K be a large positive integer. By the previous lemma applied to the set of $M + \alpha_i$ and the numbers $(-1)^M z_i$ for $i = 1, \ldots, n$ and $M = 1, \ldots, K$, we get

$$\left| \sum_{(i,M)\neq(j,N)} (-1)^{M-N} \frac{z_i \overline{z_j}}{M - N + \alpha_i - \alpha_j} \right| \leq \frac{\pi K}{\delta} \sum_{i=1}^{n} |z_i|^2.$$

It changes nothing to run the sum over pairs of pairs in which only $i \neq j$, since the terms $(i, M), (i, N)$ and $(i, N), (i, M)$ cancel each other. Put $k = M - N$ and divide by K to obtain

$$\left| \sum_{i\neq j} z_i \overline{z_j} \sum_{k=-K}^{K} \left(1 - \frac{|k|}{K}\right) \frac{(-1)^k}{k + \alpha_i - \alpha_j} \right| \leq \frac{\pi}{\delta} \sum_{i=1}^{n} |z_i|^2.$$

Taking $K \to \infty$ and recalling that

$$\frac{1}{\alpha} + \sum_{k=1}^{\infty} \left(\frac{(-1)^k}{k + \alpha} + \frac{(-1)^{-k}}{-k + \alpha} \right) = \frac{\pi}{\sin \pi\alpha}$$

yields the claim. □

Corollary 4. *With notation as in the previous corollary, for any $x \in \mathbb{R}$,*

$$\left| \sum_{i\neq j} z_i \overline{z_j} \frac{\sin 2\pi x(\alpha_i - \alpha_j)}{\sin \pi(\alpha_i - \alpha_j)} \right| \leq \delta^{-1} \sum_{i=1}^{n} |z_i|^2.$$

Proof. Apply the previous corollary twice, multiplying z_i by $\exp(\pm 2\pi i x \alpha_i)$. $\square$

We also need the following "duality" lemma.

Lemma 5 (Duality)**.** *Let $A_{m,n} \in \mathbb{C}$ and $C \in \mathbb{R}$ be constants such that for any $\beta_n \in \mathbb{C}$,*

$$\sum_m \left| \sum_n \beta_n A_{m,n} \right|^2 \leq C \sum_n |\beta_n|^2.$$

Then for any $\alpha_m \in \mathbb{C}$,

$$\sum_n \left| \sum_m \alpha_m A_{m,n} \right|^2 \leq C \sum_m |\alpha_m|^2.$$

Proof. Exercise. $\square$

Proof of Theorem 1. We prove here only the bound with the factor $\delta^{-1} + N - 1$ replaced by $\delta^{-1} + N$; there is a fun trick to pick up the extra -1 (see exercises).

By duality, we may reduce to showing that for any $z_\alpha \in \mathbb{C}$,

$$\sum_{M < n \leq M+N} \left| \sum_{\alpha \in S} z_\alpha \exp(2\pi i n \alpha) \right|^2 \leq (\delta^{-1} + N) \sum_{\alpha \in S} |z_\alpha|^2.$$

When we expand the square on the left side, the diagonal terms contribute $N \sum_\alpha |z_\alpha|^2$. The off-diagonal terms give

$$\sum_{\alpha \neq \beta} z_\alpha \overline{z_\beta} \exp(2\pi i K(\alpha - \beta)) \frac{\sin \pi N(\alpha - \beta)}{\sin \pi(\alpha - \beta)}$$

for $K = M + (N+1)/2$. By Corollary 4, this is bounded by $\delta^{-1} \sum_\alpha |z_\alpha|^2$. $\square$

Exercises

1. Find the optimal constant in the large sieve inequality (1) when the vectors in V are taken to be unit vectors forming the corners of a regular simplex in $\mathbb{R}^n$ with center at the origin. (Hint: it may simply matters to view the situation inside an $(n+1)$-dimensional space.)

2. Prove Lemma 2. (Hint: by Cauchy-Schwarz, it is enough to prove

$$\sum_{i=1}^n \left| \sum_{j \neq i} \frac{z_j}{\lambda_i - \lambda_j} \right|^2 \leq \frac{\pi^2}{\delta^2} \sum_{i=1}^n |z_i|^2.$$

Do this by extremizing an appropriate Hermitian (quadratic) form, and noting that the extremal vector must be an eigenvector.)

3. Prove Lemma 5.

4. (Cohen) Prove Theorem 1 as stated, assuming the version in which the factor $\delta^{-1}+N-1$ is replaced by $\delta^{-1}+N$. (Hint: apply the weak version to the δK-spaced points $(\alpha+k)/K$ for α running over S and k running over $\{1,\dots,K\}$, and the values b_m being related to the original a_n via

$$\sum_m b_m \exp(2\pi i \alpha m) = \sum_n a_n \exp(2\pi i K \alpha n).$$

Then take the limit as $K \to \infty$.)