5. The large sieve

Thursday Feb 20.

Sieve theory is a classical topic in number theory. With hindsight, it is closely parallel to projection theory. In particular, the large sieve, developed by Linnik in the 1940s, is closely parallel to the Fourier method in projection theory, developed by Kaufman and Falconer in the 1960s and 70s.

5.1. The Large Sieve.

Let $[N] = \{1, 2, \ldots, N\}$ and $f : [N] \to \mathbb{C}$. We define a projection of f for many different p as follows: let $\pi_p f : \mathbb{Z}_p \to \mathbb{C}$ be defined as

$$\pi_p f(a) = \sum_{n \equiv a \bmod p} f(n)$$

It's often helpful to separate a function into its constant part and mean zero part:

$$f_0 = \Big[\frac{1}{N}\sum_{n=1}^{N} f(n)\Big]\mathbf{1}_{[N]}$$

$$f_H = f - f_0 \text{ and we have } \sum_n f_H(n) = 0$$

We do the same thing with the projections:

$$(\pi_p f)_0 = \frac{1}{p}\sum_{a\in\mathbb{Z}_p} \pi_p f(a) = \text{ constant fn}$$

$$(\pi_p f)_H = \pi_p f - (\pi_p f)_0$$

Remark. We have

- $(\pi_p f)_H = \pi_p f_H$
- $(\pi_p f)_0 = \pi_p f_0$

so the order of those operations does not matter.

The main theme of the large sieve is that for an almost arbitrary function, if we take many different projections $\pi_p f$, then for most p, the oscillating high-frequency part of $\pi_p f$ is smaller than the constant part. We make this precise in the following theorem.

Let $P_M = \{p \text{ prime}, \frac{M}{2} \le p \le M\}$.

Theorem 5.1 (Linnik)**.** *If $f : [N] \to \mathbb{C}$ and $M \le N^{1/2}$ then*

$$\sum_{p\in P_M} ||(\pi_p f)_H||^2_{L^2} \lesssim \frac{N}{M}\sum_n |f_H(n)|^2$$

Remark. Background result from analytic number theory: $|P_M| \sim \frac{M}{\log M} \approx M$

Corollary 5.2.

$$Avg_{p\in P_M}||(\pi_p f)_H||_{L^2}^2 \lessapprox \frac{N}{M^2}\sum_n |f_H(n)|^2$$

Let us first see an application of this result before we move on to the proof. Last time we gave the example of *square numbers*, which have the interesting property that they leave only $\frac{p+1}{2}$ different residues mod p (that is, the quadratic residues) for any prime p. So let us think about such a set, i.e. a set where if you project it via mod p you get significantly less than all p residue classes. We ask the question "What does that tell us about the set?"

Corollary 5.3. *If $A \subset [N]$, $|\pi_p A| \leq (.99)p$ for any $p \in P_{N^{1/2}}$ then $|A| \lessapprox N^{1/2}$.*

Proof. Let $f = \mathbf{1}_A$. Assume $p \in P_{N^{1/2}}$, we get

$$\sum_{a\in\mathbb{Z}_p} |\pi_p f(a)|^2 \gtrsim \left(\frac{|A|}{p}\right)^2 \cdot p \sim |A|^2 N^{-1/2}$$

by Cauchy-Schwarz. Now, let's analyze the high-frequency part. Because $\operatorname{supp}(\pi_p f) \subseteq \pi_p(A)$, $|\operatorname{supp}(\pi_p f)| \leq .99p$. Hence

$$\sum_{a\in\mathbb{Z}_p} |(\pi_p f)_H(a)|^2 \sim \sum_{a\in\mathbb{Z}_p} |\pi_p f(a)|^2 \gtrsim |A|^2 N^{-1/2}$$

where we are using the following lemma:

Lemma 5.4. *If $g : \mathbb{Z}_p \to \mathbb{C}$ and $|\operatorname{supp}(g)| \leq .99p$ then $||g_H||_{L^2}^2 \sim ||g||_{L^2}^2$.*

Proof. Recall that $g = g_0 + g_H$ and we know $g_0 \perp g_H$. So $||g||_{L^2}^2 = ||g_0||_{L^2}^2 + ||g_H||_{L^2}^2$. If $||g_0||_{L^2}^2 \leq \frac{1}{2}||g||_{L^2}^2$ then we are done, so assume the contrary. Let $S = (\operatorname{supp}(g))^c$, by the given condition we have $|S| \geq .01p$. On S we have $g_H = -g_0$ and thus

$$||g_H||_{L^2}^2 \geq \sum_{a\in S} |g_H(a)|^2 = \sum_{a\in S} |g_0|^2 = \frac{|S|}{p}\sum_{a\in\mathbb{Z}_p} |g_0|^2 \geq \frac{1}{100}||g_0||_{L^2}^2$$

This gives $||g_H||_{L^2}^2 \sim ||g||_{L^2}^2$, as desired. □

Now we go back to our proof of the Corollary 5. We know that the L^2 norm of the high-frequency part of $\pi_p f$ is comparable to the L^2 norm of $\pi_p f$ itself. But we can upper bound the former by our Theorem:

$$\text{Avg}_{p\in P_{N^{1/2}}}||(\pi_p f)_H||_{L^2}^2 \lessapprox \frac{N}{(N^{1/2})^2}\sum_n |f_H(n)|^2 \lesssim |A|$$

In conclusion, $|A|^2 N^{-1/2} \lessapprox |A|$ and thus $|A| \lessapprox N^{1/2}$. □

It is interesting that this result matches the example of square numbers. In that sense, the bound proven above is sharp. However, it would be helpful to look at more examples. For that purpose, we look at the following.

Reference point. Random set: take a subset $A \subseteq [N]$ randomly by choosing n in A with probability $1/2$ independently. Then we see

$$\pi_p \mathbf{1}_A(a) = \#\{n \in [N], n \equiv a \bmod p, n \in A\}$$

and thus

$$\mathbb{E}_A \pi_p \mathbf{1}_A(a) = \frac{1}{2}\#\{n \in [N], n \equiv a \bmod p\} \sim \frac{1}{2}\frac{N}{p}$$

However, we don't expect it to always be $\frac{1}{2}\frac{N}{p}$. So we consider the variance, which is the square root of $\frac{1}{2}\frac{N}{p}$. Hence

$$\text{with high probability } \left|\pi_p \mathbf{1}_A(a) - \frac{N}{2p}\right| \lessapprox \sqrt{\frac{N}{2p}}$$

In particular, if $p \in P_{N^{1/2}}$ then for all $a \in \mathbb{Z}_p$

$$\text{w.h.p } \left|\pi_p \mathbf{1}_A(a) - \frac{N}{2p}\right| \lessapprox N^{1/4}$$

Now, let us compare this with what our theorem says about an arbitrary set.

Corollary 5.5. *If $A \subseteq [N]$ then*

$$\mathrm{Avg}_{p \in P_{N^{1/2}}} \mathrm{Avg}_{a \in \mathbb{Z}_p} \left|\pi_p \mathbf{1}_A(a) - \frac{|A|}{p}\right| \lessapprox N^{1/4}$$

Proof. We plug in Corollary 4 and get

$$\mathrm{Avg}_{p \in P_{N^{1/2}}} \sum_{a \in \mathbb{Z}_p} \left|\pi_p \mathbf{1}_A(a) - \frac{|A|}{p}\right|^2 \lessapprox |A| \leq N$$

Since the size of p is around $N^{1/2}$ we find that

$$\mathrm{Avg}_{p \in P_{N^{1/2}}} \mathrm{Avg}_{a \in \mathbb{Z}_p} \left|\pi_p \mathbf{1}_A(a) - \frac{|A|}{p}\right|^2 \lessapprox N^{1/2}$$

Replace the average of the squares by the square of the average (by using Cauchy-Schwartz):

$$\mathrm{Avg}_p \mathrm{Avg}_a \left|\pi_p \mathbf{1}_A(a) - \frac{|A|}{p}\right| \lessapprox N^{1/4}$$

□

So the large sieve tells us that if you take an arbitrary set A and look at a random residue class $\{n \in A : n \equiv a \bmod p\}$ with a random p and a random a, the size of the intersection is similar to what occurs for random sets A.

One cute application of this idea is to count the number of primes in an arithmetic progression. Specifically, if we take A as the set of primes up to N, then $\pi_p \mathbf{1}_A(a)$ is the number of primes $\leq N$ and congruent to a modulo p. So, the question is "How evenly distributed are the primes among those arithmetic progressions?". One might conjecture that for every p and every $a \neq 0$ the following holds:

$$\left| \pi_p \mathbf{1}_A(a) - \frac{|A|}{p} \right| \lessapprox N^{1/4}$$

The above corollary makes some progress towards this conjecture, since it implies that the conjecture is true for most residue classes. However, it is somewhat silly to call this a progress towards counting primes in arithmetic progressions, since the proof uses nothing about the prime numbers and only uses the fact that the primes are a set of numbers. That being said, this line of reasoning is still important, and in the next class we will come back to this question. We will discuss the Bombieri-Vinogradov theorem, which uses those ideas in a crucial way.

Lastly, we mention the following before we move onto the proof of the large sieve inequality. Imagine that the set A had cardinality $N/2$. Then $\pi_p \mathbf{1}_A$ would have size around N/p and since $p \in P_{N^{1/2}}$ we have that $N/p \sim N^{1/2}$. Also $|A|/p$ has size $\sim N^{1/2}$ as well, and we know the error (on average) is around $N^{1/4}$. In particular this means $(\pi_p \mathbf{1}_A)_0$ is much higher than $(\pi_p \mathbf{1}_A)_H$ at most of the points. Hence, when we take a set A of size $N/2$ look at all the projections, a typical projection looks almost constant - it's a constant function plus something much smaller. So the projection process takes something with no structure and produces something that's almost constant. People often describe this as "the projections get smoother. " In the next lecture, we will work out analogous ideas for orthogonal projections in $\mathbb{R}^2$, and we will see that the word "smoother" is just the right word in that context.

5.2. **Proof of Linnik's Large Sieve inequality.** The main idea of the proof is to study f and $\pi_p f$ by taking their Fourier transforms. So, let us first state how the Fourier transform of the functions $f : \mathbb{Z} \to \mathbb{C}$ and $\pi_p f : \mathbb{Z}_p \to \mathbb{C}$ are defined.

First, for the function $f : \mathbb{Z} \to \mathbb{C}$ with $\text{supp} f \subseteq [N]$ we define $\widehat{f} : \mathbb{R}/\mathbb{Z} \to \mathbb{C}$ as

$$\widehat{f}(\xi) = \sum_n f(n) e^{-2\pi i \xi \cdot n}$$

and we can check that $\widehat{f}(\xi)$ is 1-periodic, showing that it is well-defined. Also the two main theorems of Fourier analysis of functions over the reals hold in our case as well:

(i) Fourier Inversion:

$$f(n) = \int_0^1 \widehat{f}(\xi) e^{2\pi i n \cdot \xi} d\xi$$

(ii) Plancherel:

$$\sum_n |f(n)|^2 = \int_0^1 |\widehat{f}(\xi)|^2 d\xi$$

Secondly, for a function $g : \mathbb{Z}_p \to \mathbb{C}$ we define the Fourier transform $\widehat{g} : \mathbb{Z}_p \to \mathbb{C}$ as

$$\widehat{g}(\alpha) = \sum_{a \in \mathbb{Z}_p} g(a) e^{-2\pi i \frac{a\alpha}{p}}$$

Similarly, if we plug in $\alpha + p \cdot t$ for integer t into the definition we get that $\widehat{g}(\alpha + p \cdot t) = \widehat{g}(\alpha)$. Hence the Fourier transform $\widehat{g}$ is a well defined function on the cosets $\alpha + p\mathbb{Z}$ and thus is well defined on $\mathbb{Z}_p$. Simiarly, the Fourier Inversion and Planchárel hold as well:

(i) Fourier Inversion:

$$g(a) = \frac{1}{p} \sum_{\alpha \in \mathbb{Z}_p} \widehat{g}(\alpha) e^{2\pi i \frac{a\alpha}{p}}$$

(ii) Plancherel:

$$\sum_a |g(a)|^2 = \frac{1}{p} \sum_\alpha |\widehat{g}(\alpha)|^2$$

Now we introduce a lemma that connects the Fourier transforms of f and $\pi_p f$. We call this the *Dictionary* between the integer world and the mod$p$ world.

Lemma 5.6 (Dictionary)**.** $\widehat{\pi_p f}(\alpha) = \widehat{f}(\frac{\alpha}{p})$

Proof. The proof is clear if we unwind all the definitions:

$$\begin{aligned} \widehat{\pi_p f}(\alpha) &= \sum_{a \in \mathbb{Z}_p} \pi_p f(a) e^{-2\pi i \frac{a\alpha}{p}} \\ &= \sum_{a \in \mathbb{Z}_p} \left(\sum_{n \equiv a \bmod p} f(n) \right) e^{-2\pi i a \frac{\alpha}{p}} \end{aligned}$$

Notice that $n \equiv a \bmod p$ implies $e^{-2\pi i a \frac{\alpha}{p}} = e^{-2\pi i n \frac{\alpha}{p}}$. Thus we get

$$\widehat{\pi_p f}(\alpha) = \sum_n f(n) e^{-2\pi i n \frac{\alpha}{p}} = \widehat{f}\left(\frac{\alpha}{p}\right)$$

□

Lemma 5.7 (previous)**.** $||(\pi_p f)_H||_{L^2}^2 = \sum_{\substack{\alpha \in \mathbb{Z}_p \\ \alpha \neq 0}} \left|\widehat{\pi_p f}(\alpha)\right|^2$

Remark. Since Lemma 5.2 applies to any function, we also have $\widehat{\pi_p f_H}(\alpha) = \widehat{f_H}(\alpha/p)$.

Now let us write the left hand side of the Linnik's inequality using the *Dictionary* lemma:

$$\begin{aligned} \text{LHS of Thm.} &= \sum_{p \in P_M} ||(\pi_p f)_H||_{L^2}^2 \\ &= \sum_{p \in P_M} \frac{1}{p} \sum_{\substack{\alpha \neq 0 \\ \alpha \in \mathbb{Z}_p}} \left|\widehat{\pi_p f_H}(\alpha)\right|^2 \\ &\sim \frac{1}{M} \sum_{p \in P_M} \sum_{\substack{\alpha \neq 0 \\ \alpha \in \mathbb{Z}_p}} \left|\widehat{f_H}\Big(\frac{\alpha}{p}\Big)\right|^2 \end{aligned} \tag{17}$$

Let's now visualize this set of points $Q_M = \{\frac{\alpha}{p} : p \in P_M \text{ and } 0 < \alpha \leq p-1\}$. Note that $|Q_M| \approx M^2$.

Lemma 5.8. *If* $\frac{\alpha_1}{p_1}, \frac{\alpha_2}{p_2} \in Q_M$ *are not equal, then* $|\frac{\alpha_1}{p_1} - \frac{\alpha_2}{p_2}| \geq \frac{1}{M^2}$.

Proof.

$$\left|\frac{\alpha_1}{p_1} - \frac{\alpha_2}{p_2}\right| = \left|\frac{\alpha_1 p_2 - \alpha_2 p_1}{p_1 p_2}\right| \geq \frac{1}{p_1 p_2} \geq \frac{1}{M^2}$$

□

Remark. If $\frac{\alpha_1}{p_1} = \frac{\alpha_2}{p_2}$ in Q_M, then $p_1 = p_2$ and $\alpha_1 = \alpha_2$.

In Figure 7 below, we have the interval $[0, 1]$ with the points of Q_M on it. Q_M is not perfectly evenly spaced out but is very close to perfect. In orange is the graph of the function $|\widehat{f_H}|^2$ and we have highlighted the value of $|\widehat{f_H}|^2$ on the set Q_M. What we are interested in is taking the sum of this function $|\widehat{f_H}|^2$ on the set Q_M. This reminds us of Riemann integration. Indeed, we will compare this to the integral $\int_{[0,1]} |\widehat{f_H}(\omega)|^2 d\omega$.

Notice that there is a way for this sum to be way bigger than the integral: if $|\widehat{f_H}|^2$ has narrow peaks on Q_M. This way, the sum will be big while the peaks don't contribute much to the integral $\int_{[0,1]} |\widehat{f_H}(\omega)|^2 d\omega$. So it is important to understand how wide the peaks are. The following heuristics helps for this task:

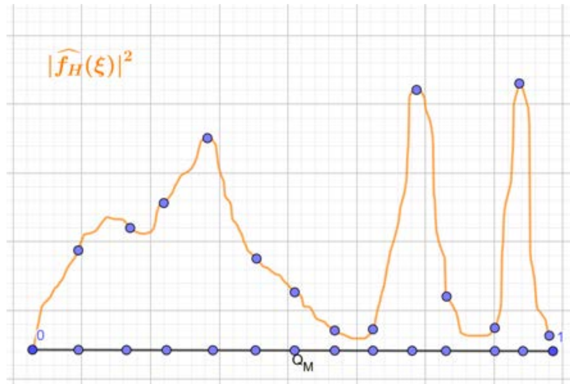

FIGURE 7. Picture.

Heuristic: $|\widehat{f_H}|^2$ is roughly constant on intervals of size $\frac{1}{N}$.

This can be seen from the fact that f is supported on $[0, N]$. We will make this notion precise in a moment, but it means that each peak should be $\frac{1}{N}$ wide. Since we are given $M \leq N^{\frac{1}{2}}$, this guarantees that the spacing between two consecutive points of Q_M is bigger than the width $\frac{1}{N}$.

We will now follow this heuristic and obtain our desired inequality (we shall come back and prove more rigorously later). Heuristic implies

$$\sum_{\xi \in Q_M} |\widehat{f_H}(\xi)|^2 \lesssim N \int_0^1 |\widehat{f_H}(\omega)|^2 d\omega$$

This is because for each $\xi \in Q_M$:

$$|\widehat{f_H}(\xi)|^2 \lesssim N \int_{I_\xi} |\widehat{f_H}(\omega)|^2 d\omega$$

where I_ξ is a length $\frac{1}{N}$ interval around ξ. Then we can see that the intervals I_ξ for $\xi \in Q_M$ doesn't overlap, so we can bound the sum over $\xi \in Q_M$ by the integral over the domain $[0, 1]$.

The rest is just algebra: recall (17) and we get

$$\text{LHS of Thm.} \sim \frac{1}{M} \sum_{\xi \in Q_M} |\widehat{f_H}(\xi)|^2$$
$$\lesssim \frac{N}{M} \int_0^1 |\widehat{f_H}(\omega)|^2 d\omega = \frac{N}{M} \sum_n |f_H(n)|^2$$

as desired.

Remark. We have this theme that if you take one function and project it mod$p$ for many different primes, most of them look nearly constant. So why is the zero frequency special in this story? It's because for primes p the sets $\{\frac{\alpha}{p} : 0 \le \alpha \le p-1\}$ all intersect at 0 but all the other points appear only once. Hence the zero frequency is being counted very differently than all the other frequencies. If $\hat{f}$ is large on a small interval I that does not contain zero, then this part of $\hat{f}$ will contribute to $\pi_p f$ for only a few primes p. But if $\hat{f}$ is large on a small interval I around zero, then this part of $\hat{f}$ will contribute to $\pi_p f$ for every p.

Lastly, we will rigorously prove our heuristic. We will take a function $\psi_N : \mathbb{Z} \to \mathbb{C}$ such that

$$\psi_N(n) = 1 \text{ for } n \in [N] \text{ and } \psi_N \text{ smooth, rapidly decaying}$$

The Fourier Transform of ψ_N behaves like this:

$$\widehat{\psi_N}(\xi) = \begin{cases} \sim N & \text{if } |\xi| \le \frac{1}{N} \\ \lesssim N(N|\xi|)^{-1000} & \text{if } |\xi| > \frac{1}{N} \end{cases} \tag{18}$$

Refer to the figure below for a visualization of $|\widehat{\psi_N}|$.

Audience Question: What does smoothness mean for a function on $\mathbb{Z}$? **Answer:** You can think of ψ_N as a smooth function on the real line being restricted to $\mathbb{Z}$.

This function is helpful because

$$f = f\psi_N \text{ if supp} f \subseteq [N]$$

By taking the Fourier Transform, we get $\widehat{f} = \widehat{f} * \widehat{\psi_N}$. By the triangle inequality we obtain $|\widehat{f}| \le |\widehat{f}| * \left|\widehat{\psi_N}\right|$. Noting that

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} |\widehat{\psi_N}(\xi)| d\xi \lesssim 1$$

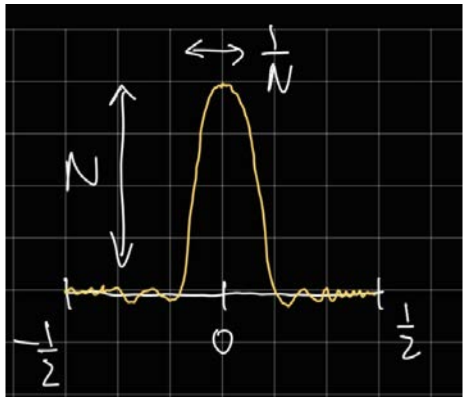

FIGURE 8. Graph of $|\widehat{\psi_N}|$

we can show by Cauchy-Schwartz that

$$|\widehat{f}|^2 \lesssim |\widehat{f}|^2 * \left|\widehat{\psi_N}\right|$$

Audience Question: The Fourier Transform of functions on $\mathbb{Z}$ and $\mathbb{R}$ are not the same. Which one do you mean when you say $\widehat{\psi_N}$?

Answer: So we mean that we first take a function $\psi_{N,\mathbb{R}} : \mathbb{R} \to \mathbb{C}$ smooth with $\psi_{N,\mathbb{R}} = 1$ on $[-N, N]$ and rapidly decaying outside. Then we define $\psi_{N,ZZ}$ as the restriction of $\psi_{N,\mathbb{R}}$ to $\mathbb{Z}$. To analyze the Fourier transform of these functions, we start with $\psi_{N,\mathbb{R}}$. By standard integration by parts, we get: for $\xi \in \mathbb{R}$

$$|\widehat{\psi_{N,\mathbb{R}}}(\xi)| = \begin{cases} \sim N & \text{if } |\xi| \leq \frac{1}{N} \\ \lesssim N(N|\xi|)^{-1000} & \text{if } |\xi| > \frac{1}{N} \end{cases} \tag{19}$$

Now $\widehat{\psi_{N,\mathbb{Z}}}$ is related to $\widehat{\psi_{N,\mathbb{R}}}$ by the equation below, which boils down to Poisson summation:

$$\widehat{\psi_{N,\mathbb{Z}}}(\xi) = \sum_{z\in\mathbb{Z}} \widehat{\psi_{N,\mathbb{R}}}(\xi + z)$$

for $\xi \in \mathbb{R}/\mathbb{Z}$. Now the bounds for $|\widehat{\psi_{N,\mathbb{R}}}|$ in (19) combined with this equation give the desired bounds for $|\widehat{\psi_{N,\mathbb{Z}}}|$ in (18).

Now let's do a slightly more rigorous proof of the Linnik's large sieve inequality. Recall the statement:

Theorem 5.9 (Linnik)**.** *If* $f : [N] \to \mathbb{C}$ *and* $M \leq N^{1/2}$ *then*

$$\sum_{p \in P_M} ||(\pi_p f)_H||_{L^2}^2 \lesssim \frac{N}{M} \sum_n |f_H(n)|^2$$

Proof. Remember that

$$\text{LHS} \sim \frac{1}{M} \sum_{\xi \in Q_M} |\widehat{f_H}(\xi)|^2$$

To relate this sum to an integral, we use the fact that $|\widehat{f_H}|^2 \lesssim |\widehat{f_H}|^2 * \left|\widehat{\psi_N}\right|$. This fact encodes the locally constant property of $|\widehat{f}|^2$. We get

$$\begin{aligned} \frac{1}{M} \sum_{\xi \in Q_M} |\widehat{f_H}(\xi)|^2 &\lesssim \frac{1}{M} \sum_{\xi \in Q_M} \int_{\mathbb{R}/\mathbb{Z}} |\widehat{f_H}(\omega)|^2 \left|\widehat{\psi_N}(\xi - \omega)\right| d\omega \\ &= \frac{1}{M} \int_{\mathbb{R}/\mathbb{Z}} |\widehat{f_H}(\omega)|^2 \left(\sum_{\xi \in Q_M} |\widehat{\psi_N}(\xi - \omega)| \right) d\omega \end{aligned}$$

We claim that this sum is bounded by $\lesssim N$:

$$\sum_{\xi \in Q_M} |\widehat{\psi_N}(\xi - \omega)| \lesssim N$$

This is because the function $g(\xi) = |\widehat{\psi_N}(\xi - \omega)|$ has a peak around ω with height N and width $1/N$ and is extremely small away from this peak. The distance between any two distinct points in Q_M is $\gtrsim \frac{1}{M^2} \geq \frac{1}{N}$, and so at most $O(1)$ points of Q_M lie under the peak of $g(\xi)$. Hence, we find that

$$\frac{1}{M} \sum_{\xi \in Q_M} |\widehat{f_H}(\xi)|^2 \lesssim \frac{N}{M} \int_{\mathbb{R}/\mathbb{Z}} |\widehat{f_H}(\omega)|^2 d\omega = \frac{N}{M} \sum_n |f_H(n)|^2$$

finishing the proof of Linnik's large sieve. □

In the last five minutes of the class, we want to give a quick teaser on how these ideas come up in the setting of projection theory over $\mathbb{R}^d$. We have this theme that functions on $[1, 2, \ldots, N]$ look almost constant after projecting mod$p$ for most primes p. And there is a totally analogous phenomenon for functions on $\mathbb{R}^d$. Specifically, if you project those functions onto lower subspaces, almost all of them look smoother than the original function. We have mentioned on the first day that if you are in a high enough dimension, even L^2 functions that are nowhere continuous has the property that its projection on a typical line are C^1 are even C^2.

So here is a setup that is analogous to the large sieve. Let $f : \mathbb{R}^d \to \mathbb{C}$ and $V \subset \mathbb{R}^d$ be a subspace. Then we have the projection $\pi_V f : V \to \mathbb{C}$.

Remark. For any function $g : V \to \mathbb{C}$ on a vector space V, the Fourier Transform $\widehat{g} : V \to \mathbb{C}$ is also defined on V.

We also have the Dictionary lemma:

Lemma 5.10 (Dictionary)**.** *We have $\widehat{\pi_V f} = \widehat{f}\big|_V$. Notice that $\widehat{\pi_V f}$ is a function on V while $\widehat{f}$ is a function on $\mathbb{R}^d$.*

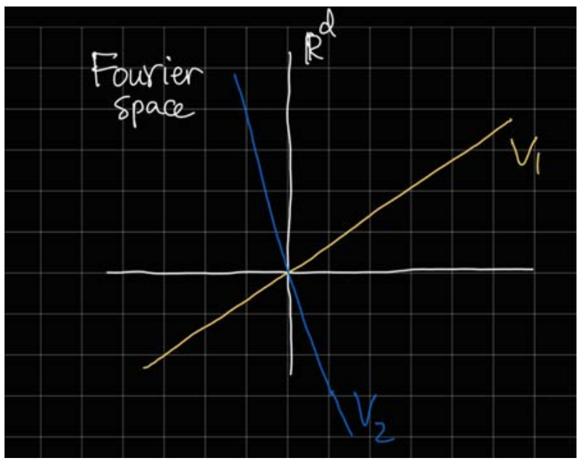

FIGURE 9. Picture.

In the figure, we see two subspaces V_1 and V_2 (among others) of $\mathbb{R}^d$. Notice that the origin lies in every subspace V. On the other hand, a non-zero frequency $\omega \in \mathbb{R}^d$ only lies in a small fraction of subspaces V. Therefore, if $\hat{f}$ is large on a small ball B far away from zero, then this contributes to $\pi_V f$ for only a small fraction of subspaces V. On the other hand, if $\hat{f}$ is large on a small B around zero, then this contributes to $\pi_V f$ for every subspace V. If we compare f with a typical $\pi_V f$, the high-frequency parts of the Fourier transform are "damped" in $\pi_V f$ compared to f. This causes $\pi_V f$ to be smoother than f. We will explore these ideas more fully next class.