Geometry of Numbers Lecture 11

Reminder: $\mathcal{X}_n = SL_n(\mathbb{R})/SL_n(\mathbb{Z})$ - space of lattices in $\mathbb{R}^n$ of covolume 1.

$m_{\mathcal{X}_n}$ - the unique $SL_n(\mathbb{R})$-inv. Borel prob. measure on $\mathcal{X}_n$.

Given F on $\mathcal{X}_n$, want to know the distribution of its values - is $F(L)$ close to $\mathbb{E}(F)$ $\leftarrow$ average of F w.r.t. $m_{\mathcal{X}_n}$ for most L?

Specifically for $F = \hat{f}$, $f \in L^1(\mathbb{R}^n)$,

e.g. $f = \mathbb{1}_B$ $B \subset \mathbb{R}^n$ a ball.

(Recall: $\hat{f} : \mathcal{X}_n \longrightarrow \mathbb{R}$, $\hat{f}(L) = \sum_{v \in L \setminus 0} f(v)$,

$\hat{f} \in L^1(\mathcal{X}_n, m_{\mathcal{X}_n})$ whenever $f \in L^1(\mathbb{R}^n, \mathrm{Vol})$,

and $$\int_{\mathcal{X}_n} \hat{f}\, dm_{\mathcal{X}_n} = \int_{\mathbb{R}^n} f\, d\mathrm{Vol} \quad (\mathrm{SSF2})$$).

Q1 Can we compute higher moments of $\hat{f}$ (thought of as a random variable on the prob. space $(\mathcal{X}_n, m_{\mathcal{X}_n})$).

Recall: $\int \hat{f}\, dm_{\mathcal{X}_n} = \mathbb{E}(X)$

when $X = \hat{f}$

For $k \geq 1$, the (raw) k-th moment of X

is $\mathbb{E}(X^k) = \int_{\mathcal{X}_n} (\hat{f})^k\, dm_{\mathcal{X}_n}$

the centered k-th moment

$$\mathbb{E}\left((X - \mathbb{E}(X))^k \right) = \int_{\mathcal{X}_n} \left(\hat{f} - \int_{\mathbb{R}^n} f\, d\mathrm{Vol}\right)^k dm_{\mathcal{X}_n}$$

The centered 2nd moment is the variance

$\mathrm{Var}(X) = \sigma^2$ $\quad \sigma$ is the standard deviation.

Prop (Markov/Chebyshev inequality)

Suppose $k \geq 1$, $X \geq 0$, $a > 1$, and $\mathbb{E}(X^k) < \infty$.

Then $\mathbb{P}(X > a\,\mathbb{E}(X)) \leq \dfrac{\mathbb{E}(X^k)}{(a\,\mathbb{E}(X))^k}$.

If $k=1$, $\mathbb{P}(X > a\,\mathbb{E}(X)) \leq \frac{1}{a}$.

Pf: $\mathbb{E}(X^k) \geq \mathbb{P}(X^k > a^k\,\mathbb{E}(X)^k)\, a^k\,\mathbb{E}(X)^k$

$= \mathbb{P}(X > a\,\mathbb{E}(X))\,(a\,\mathbb{E}(X))^k$

Moving terms, get the Prop.

Markov's inequality for $Y = (X - \mathbb{E}(X))^2$, $k=1$,

$E(Y) = \mathrm{Var}(X)$, for $a > 1$,

$\mathbb{P}(Y > a\,\mathrm{Var}(X)) \leq \frac{1}{a}$.

Prop (Borel-Cantelli Lemma) If μ is a prob. measure, on a space X, $A_1, A_2, \ldots$ measurable sets with $\sum_{n=1}^{\infty} \mu(A_n) < \infty$ then for a.e. $x \in X$,

$\exists\, n_0 = n_0(x)$ $\forall\, n \geq n_0$ $x \notin A_n$.

Pf Define $A_\infty = \{x \in X :$ there are inf. many n for which $x \in A_n \}$

$$= \bigcap_{m} \bigcup_{n \geq m} A_n$$

Then $\mu(A_\infty) = \lim_{m \to \infty} \mu\left(\bigcup_{n \geq m} A_n\right) \leq \lim_{m \to \infty} \sum_{n=m}^{\infty} \mu(A_n)$

$$\xrightarrow[m \to \infty]{} 0$$

Thm (Rogers '55) If $f \in C_c(\mathbb{R}^n)$, $k \in \mathbb{N}$, $k < n$, then $\hat{f} \in L^k(\mathcal{X}_n, m_{\mathcal{X}_n})$ (i.e. the raw k-th moment of $\hat{f}$ is finite).

Moreover Rogers obtained explicit bounds for the kth moment.

<u>Thm</u> (Eskin-Margulis-Mozes '98) If $p<n$

$1\le p\in\mathbb{R}$, $f\in C_c(\mathbb{R}^n)$, then $\hat{f}\in L^p(\mathcal{X}_n, m_{\mathcal{X}_n})$.

OTOH there are $f\in C_c(\mathbb{R}^n)$ for which $\hat{f}\notin L^n(\mathcal{X}_n, m_{\mathcal{X}_n})$.

We will give a sketch of proof of the Rogers bound (more detailed for $k=2$).

Let $k<n$ $k,n\in\mathbb{N}$, and write

$$\mathbb{R}^{kn}\cong M_{k,n}(\mathbb{R})\cong\{(u_1,\dots,u_k) : u_i\in\mathbb{R}^n \text{ for } i=1,\dots,k\}$$

<u>Theorem</u> (Siegel '45) Let $f\in L^1(\mathbb{R}^{kn}, \mathrm{Vol})$

Define ${}^{n,k}\hat{f}:\mathcal{X}_n\longrightarrow\mathbb{R}$

$$ {}^{n,k}\hat{f} = \sum_{\substack{u_1,\dots,u_k\in L^{\text{-prot}} \\ \text{lin. ind.}}} f(u_1,\dots,u_k)$$

note: $u_1,\dots,u_k$ linearly independent

(what we denoted by $\hat{f}$ is ${}^{n,1}\hat{f}$).

Then $$\int_{\mathcal{X}_n} {}^{n,k}\hat{f}\, dm_{\mathcal{X}_n} = \int_{\mathbb{R}^{kn}} f\, d\mathrm{Vol} \qquad (\mathrm{SSF2})_{n,k}$$

(for $k=1$ we recover (SSF2)).

<u>Sketch of proof of Siegel's theorem</u>:

Denote $\mathbb{R}^{kn}_{l.i.} = \{(u_1,\dots,u_k)\in\mathbb{R}^{kn} : u_i \text{ lin. ind.}\}$

$\mathbb{R}^{kn}_{l.i.}$ is an open dense subset of $\mathbb{R}^{kn}$

$$\mathbb{R}^{kn}_{l.i.} = \mathbb{R}^{kn} \smallsetminus \bigcap_{i_1<\dots<i_k} \{\det \text{ of } k\times k \text{ minor } i_1,\dots,i_k \text{ is } 0\}$$

So enough to prove $(\mathrm{SSF2})_{n,k}$ for

$f\in C_c(\mathbb{R}^{kn}_{l.i.})$. This uses the lin. ind requirement in def of ${}^{n,k}\hat{f}$

$SL_n(\mathbb{R})$ acts transitively on $\mathbb{R}^{kn}_{l.i.}$,

$$\mathrm{Stab}_{SL_n(\mathbb{R})}(e_1,\dots,e_k) = \left\{\left(\begin{array}{c|c} I_k & * \\ \hline 0 & * \end{array}\right)\right\}$$

So $\mathbb{R}^{kn}_{l.i.}$ admits a <u>unique</u> $SL_n(\mathbb{R})$-inv.

Radon measure (up to scaling), namely restriction of Vol on $\mathbb{R}^{kn}$.

$*$ Show that for $f \in C_c(\mathbb{R}^{nk}_{l.i.})$

$$\widehat{{}^{n,k}f} \in L^1(\mathcal{X}_n, m_{\mathcal{X}_n}).$$

We will not do this, this is the main technical part of the proof.

Assuming $*$, the map $C_c(\mathbb{R}^{kn}_{l.i.}) \longrightarrow \mathbb{R}$

$$f \longmapsto \int_{\mathcal{X}_n} \widehat{{}^{n,k}f} \, dm_{\mathcal{X}_n}.$$

is a continuous positive linear functional on $C_c(\mathbb{R}^{kn}_{l.i.})$ so defines a Radon measure μ on $\mathbb{R}^{kn}_{l.i.}$.

<u>Claim:</u> $\forall f \in C_c(\mathbb{R}^{kn}_{l.i.})$ $\forall g \in SL_n(\mathbb{R})$,

$$\widehat{{}^{n,k}(f\circ g)} = \left(\widehat{{}^{n,k}f}\right)\circ g$$

$$\widehat{{}^{n,k}(f\circ g)}(L) = \sum_{\substack{u_1,\dots,u_k \in L\setminus\{0\} \\ \text{l.i.}}} f\circ g(u_1,\dots,u_k) = \sum_{\substack{u_1,\dots,u_k \in L\setminus\{0\} \\ \text{l.i.}}} f(gu_1,\dots,gu_k)$$

$$\underset{v_i = gu_i}{=} \sum_{\substack{v_1,\dots,v_k \in gL\setminus\{0\} \\ \text{l.i.}}} (v_1,\dots,v_k) = \left(\widehat{{}^{n,k}f}\right)(gL)$$

This proves claim.

It follows that μ defined above is $SL_n(\mathbb{R})$-inv, because $\forall g \in SL_n(\mathbb{R})$

$$\int_{\mathbb{R}^{n,k}_{\text{l.i.}}} (f\circ g)\, d\mu \underset{\text{def}}{=} \int_{\mathfrak{X}_n} \widehat{{}^{n,k}f\circ g}\, dm_{\mathfrak{X}_n} \underset{\text{claim}}{=}$$

$$\int_{\mathfrak{X}_n} \widehat{{}^{n,k}f}\circ g\, dm_{\mathfrak{X}_n} \underset{m_{\mathfrak{X}_n} \text{ is } SL_n(\mathbb{R})\text{-inv}}{=} \int_{\mathfrak{X}_n} \widehat{{}^{n,k}f}\, dm_{\mathfrak{X}_n} = \int_{\mathbb{R}^{n,k}_{\text{l.i.}}} f\, d\mu$$

By uniqueness, $\exists c>0$ s.t. $\mu = c\,\mathrm{Vol}$

$$\int_{\mathcal{X}_n}^{n,k} \hat{f}\, dm_{\mathcal{X}_n} = c \int_{\mathbb{R}^{nk}_{c \dots}} f\, d\mathrm{Vol} \quad (\ast\ast)$$

Remains to show $c=1$, for this use

$$f_R = \frac{\mathbb{1}_{B_R}}{\mathrm{Vol}(B_R)} \qquad B_R = [-R,R]^{kn}$$

As we did for $k=1$, not hard to show

$$\forall L \in \mathcal{X}_n, \qquad \hat{f}_R(L) \xrightarrow[R\to\infty]{} 1.$$

Using the dominated convergence theorem (need to justify), and plugging $(\ast\ast)$, get $c=1$.

!!!

$\underline{\text{Cor}}$ For $f \in L^1(\mathbb{R}^n)$, $\hat{f}(L) = \sum_{v \in L\setminus\{0\}} f(v)$ satisfies $\hat{f} \in L^k(\mathcal{X}_n, m_{\mathcal{X}_n})$, for all $k \in \mathbb{N}$, $k<n$.

Pf Given $f \in L^1(\mathbb{R}^n)$, define $f_1 \in L^1(\mathbb{R}^{kn})$ by $f_1(u_1, \dots, u_k) = f(u_1) f(u_2) \cdots f(u_k)$.

$$\int_{\mathcal{X}_n} (\hat{f})^k(L)\, dm_{\mathcal{X}_n}(L) = \int_{\mathcal{X}_n} \Big(\sum_{u \in L \setminus 0} f(u) \Big)^k dm_{\mathcal{X}_n}(L) =$$

$$= \int_{\mathcal{X}_n} \Big(\sum_{u_1, \dots, u_k \in L \setminus 0} f(u_1) \cdots f(u_k) \Big) dm_{\mathcal{X}_n}(L) = \int_{\mathcal{X}_n} \sum_{\substack{u_1, \dots, u_k \\ \in L \setminus 0}} f_1(u_1, \dots, u_k)\, dm_{\mathcal{X}_n}(L)$$

$$\cancel{=} \int {}^{n,k}\hat{f}_1\, dm_{\mathcal{X}_n} < \infty$$

This is a mistake, def of ${}^{n,k}\hat{f}$ involves lin. ind. vectors. Need to separate into independent and dependent k-tuples, and work inductively on k, see next proof.

Want bounds on $\int_{\mathcal{X}_n} (\hat{f})^k\, dm_{\mathcal{X}_n}$.

changed const

Thm (Rogers second moment bound '55) $\forall n\, \exists C$ $\forall f \in C_c(\mathbb{R}^n)$, $f \geq 0$, $\hat{f}(L) = \sum_{x \in L \setminus 0} f(x)$, $\hat{f} : \mathcal{X}_n \to \mathbb{R}$,

added condition

satisfies $\hat{f} \in L^2(\mathcal{X}_n, m_{\mathcal{X}_n})$ and

$$\mathrm{Var}(\hat{f}) = \int_{\mathcal{X}_n} \Big(\hat{f} - \int_{\mathcal{X}_n} \hat{f}\, dm_{\mathcal{X}_n} \Big)^2 dm_{\mathcal{X}_n} \leq C \|f\|_\infty \int_{\mathcal{X}_n} |\hat{f}|\, dm_{\mathcal{X}_n}$$

$\underline{\text{Pf}}$: $Var(X) = \mathbb{E}((X-\mathbb{E}(X))^2) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$

So we compute $\int_{\mathcal{X}_n} (\hat{f})^2 \, dm_{\mathcal{X}_n}$, using

$(SSF2)_{n,k}$ with $k=2$.

Define $f_1(x,y) = f(x)f(y)$ $\quad f_1 : \mathbb{R}^{2n} \to \mathbb{R}$

$$\int_{\mathcal{X}_n} (\hat{f})^2 dm_{\mathcal{X}_n} = \int_{\mathcal{X}_n} \Big(\sum_{u \in L \setminus 0} f(u)\Big)\Big(\sum_{v \in L \setminus 0} f(v)\Big) dm_{\mathcal{X}_n}$$

$$= \int_{\mathcal{X}_n} \Bigg[\Big(\sum_{\substack{u,v \in L \setminus 0 \\ \text{l.i.}}} f(u)f(v)\Big) + \Big(\sum_{\substack{u,v \in L \setminus 0 \\ \text{lin. dep.}}} f(u)f(v)\Big)\Bigg] dm_{\mathcal{X}_n}$$

$$= \int_{\mathcal{X}_n} \widehat{{}^{n,2}f_1} \, dm_{\mathcal{X}_n} + \int_{\mathcal{X}_n} \Big(\sum_{\substack{w \in L \\ \text{primitive}}} \sum_{\substack{k_1,k_2 \in \mathbb{Z} \setminus 0 \\ k_1 > 0}} f(k_1 w) f(k_2 w)\Big) dm_{\mathcal{X}_n}$$

uses corrected form of $\widehat{{}^{n,k}f}$

divide into two cases, $k_1 \leq |k_2|$ and $k_1 > |k_2|$, they are identical

$$\underset{(SSF)_{2,n}}{\leq} \int_{\mathbb{R}^{2n}} f_1 \, dVol + 2 \int \sum_{\substack{k_1,k_2 \in \mathbb{Z} \\ 1 \leq k_1 \leq |k_2|}} \sum_{\substack{w \in L \\ \text{primitive}}} f(k_1 w) f(k_2 w) \, dm_{\mathcal{X}_n}$$

p

divide into cases, $k_2>0$ and $k_2<0$. Write $k=k_2$ and bound $f(kw) \le \|f\|_\infty$

$$\le \left(\int_{\mathbb{R}^n} f\, d\mathrm{Vol}\right)^2 + 4\|f\|_\infty \int_{\mathcal{X}_n} \sum_{k\in\mathbb{N}} k \sum_{\substack{w\in L \\ \text{primitive}}} f(kw)\, dm_{\mathcal{X}_n}$$

Using $f_k(x) = f(kx)$ and $\hat{f}^p(L) = \sum_{\substack{w\in L \\ \text{primitive}}} f(w)$

$$= \left(\int_{\mathbb{R}^n} f\, d\mathrm{Vol}\right)^2 + 4\|f\|_\infty \sum_{k\in\mathbb{N}} k \int_{\mathcal{X}_n} \hat{f}_k^{\,p}\, dm_{\mathcal{X}_n}$$

(SSF1) for $\hat{f}^p$

$$= \left(\int_{\mathbb{R}^n} f\, d\mathrm{Vol}\right)^2 + 4\|f\|_\infty \sum_{k\in\mathbb{N}} \frac{k}{\zeta(n)} \int_{\mathbb{R}^n} f_k(x)\, d\mathrm{Vol}(x)$$

$kx=y$, $d\mathrm{Vol}(x) = \frac{1}{k^n} d\mathrm{Vol}(y)$

$$= \left(\int_{\mathbb{R}^n} f\, d\mathrm{Vol}\right)^2 + \frac{4}{\zeta(n)} \|f\|_\infty \sum_{k\in\mathbb{N}} \frac{k}{k^n} \int_{\mathbb{R}^n} f\, d\mathrm{Vol}$$

$$= \left(\int_{\mathbb{R}^n} f\, d\mathrm{Vol}\right)^2 + \frac{4\zeta(n-1)}{\zeta(n)} \|f\|_\infty \int_{\mathbb{R}^n} f\, d\mathrm{Vol}.$$

Subtracting $\left(\int_{\mathcal{X}_n} \hat{f}\, dm_{\mathcal{X}_n}\right)^2 = \left(\int_{\mathbb{R}^n} f\, d\mathrm{Vol}\right)^2$ from both sides, get conclusion desired, with $C = \frac{4\zeta(n-1)}{\zeta(n)}$

Application to counting lattice points in growing sets (following Schmidt '60)

Recall (ex. 3) $\forall L \in \mathcal{X}_n$, any bounded convex set S with non-empty interior, rS is the dilation of S by factor r

$$\#(L \cap rS) = \underbrace{\mathrm{Vol}(rS)}_{\text{main term}} + \underbrace{O(r^{n-1})}_{\text{error term}} \quad \text{as } r \to \infty,$$

where implicit const. depends on L and on S.

$$= \mathrm{Vol}(rS) + O\left(\left(\mathrm{Vol}(rS)\right)^{1-\frac{1}{n}} \right).$$

Q: can error term be improved?

for which choices of L, S, maybe for more general sequences of growing bodies.

Conj (Götze '98) For $m_{\mathcal{X}_n}$ - a.e. $L \subset \mathbb{R}^n$, for $B = S = B(0,1)$ euclidean ball, $\forall \varepsilon > 0$

$$\#(L \cap rS) = \mathrm{Vol}(rS) + O\left(r^{\frac{n}{2}-\frac{1}{2}+\varepsilon}\right)$$
$$= \mathrm{Vol}(rS) + O\left(\mathrm{Vol}(rS)^{\frac{1}{2}-\frac{1}{2n}+\varepsilon}\right)$$

where implicit const. depends on L, S, ε.

Let $\{\Omega_r : r>0\}$ be a collection of Borel sets. We say it is an <u>unbounded ordered family</u> if: (i) $r_1 \le r_2 \Rightarrow \Omega_{r_1} \subset \Omega_{r_2}$

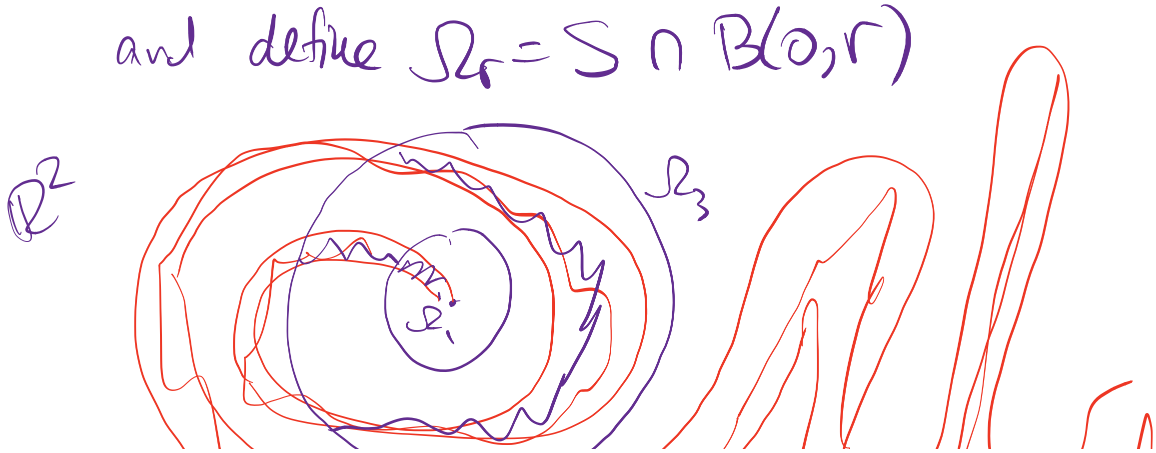
(ii) $\mathrm{Vol}(\Omega_r) < \infty$ and $\mathrm{Vol}(\Omega_r) \xrightarrow[r\to\infty]{} \infty$.

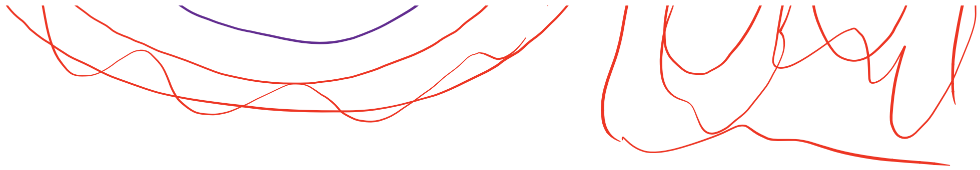
(iii) $\forall V>0\ \exists r$ s.t. $\mathrm{Vol}(\Omega_r) = V$.

<u>Examples</u> * $\Omega_r = B(0,r)$

* For any $S \subset \mathbb{R}^n$ $\mathrm{Vol}(S) = \infty$,

and define $\Omega_r = S \cap B(0,r)$





<u>Thm</u> (Schmidt '60) Let $\{\Omega_r : r>0\}$ be an unbounded ordered family in $\mathbb{R}^n$. Then for every $\varepsilon>0$, for $m_{\mathcal{X}_n}$-a.e. $L\in\mathcal{X}_n$,

$$(*)\quad \#(L\cap\Omega_r) = \mathrm{Vol}(\Omega_r) + O\left(\mathrm{Vol}(\Omega_r)^{\frac{1}{2}+\varepsilon}\right)$$

(implicit const depends on L and on ε).
$\{\Omega_r\}$

<u>Remark</u> Schmidt's result does not attain the bound $\frac{1}{2}-\frac{1}{2n}+\varepsilon$ in Conj, but it's the best result known.

For the proof we will some lemmas and notation.

By (iii), for each $N\in\mathbb{N}$ $\exists\, r_N$ s.t. $S_N=\Omega_{r_N}$ satisfies $\mathrm{Vol}(S_N)=N$.

Let $\rho_N=\mathbb{1}_{S_N}$, for $N_1<N_2$ let

$${}_{N_1}\rho_{N_2} = \rho_{N_2} - \rho_{N_1} \underset{\uparrow}{=} \mathbb{1}_{S_{N_2} \setminus S_{N_1}}$$

sets are nested.

Lemma 1 (dyadic decomposition) Let $T \in \mathbb{N}$ and let $K_T = \{(N_1, N_2) \in \mathbb{N} \times \mathbb{N} : 0 \le N_1 < N_2 \le 2^T$, $\exists t \in \mathbb{N} \cup \{0\}$, $2^t \mid N_1$, $N_2 = N_1 + 2^t\}$

Then $\displaystyle\sum_{(N_1,N_2) \in K_T} \mathrm{Var}(\widehat{{}_{N_1}\rho_{N_2}}) \le 2(T+1)2^T$.

Example $T=4$ $t=4$ $N_1 = 0$ $N_2 = 2^T$

$t=3$ $N_1 = 0, N_2 = 8$ $N_1 = 8, N_2 = 16$

$t=2$ $(N_1, N_2) \in \{(0,4), (4,8), (8,12), (12,16)\}$

$\vdots$

$t=0$ $\{(k, k+1) : k \in \{0, \dots, 15\}\}$.

Pf $\mathrm{Var}(\widehat{{}_{N_1}\rho_{N_2}}) \le 2 \|{}_{N_1}\rho_{N_2}\|_\infty \int_{\mathcal{X}_n} {}_{N_1}\rho_{N_2} \, dm_{\mathcal{X}_n}$

$= 2\,\mathrm{Vol}(S_{N_2} \setminus S_{N_1}) = 2(N_2 - N_1)$ (1)

(note $\int_{N_1}^{N_2}$ not cts, can derive from Rogers 2^{nd} moment bound and an approximation argument).

Each value of $N_2 - N_1 = 2^t$, for $t \in \{0, \dots, T\}$, occurs 2^{T-t} times, and therefore

$$(2) \quad \sum_{(N_1,N_2)\in K_T} (N_2 - N_1) = \sum_{0\le t\le T} 2^t 2^{T-t} = (T+1)2^T.$$

Combining (1) and (2) gives the conclusion.

Lemma 2 Fix $\varepsilon > 0$, let C be as in Rogers 2^{nd} moment bound. For all $T \in \mathbb{N}$, there is $\mathrm{Bad}_T \subset \mathcal{X}_n$ with

$$m_{\mathcal{X}_n}(\mathrm{Bad}_T) \le \frac{C}{((\log 2)T - 1)^{1+\varepsilon}}$$

s.t. for every every $N \le 2^T$, and every $L \notin Bad_T$,

(3) $$\left(\widehat{\rho}_N(L) - N\right)^2 \le T(T+1)2^T(\log(2)T-1)^{1+\varepsilon}$$

Pf of Thm assuming Lemma 2

$$\sum_{T \in \mathbb{N}} \frac{1}{(\log(2)T-1)^{1+\varepsilon}} < \infty$$

By Borel-Cantelli, for m_χ a.e. L, $\exists T_0 = T_0(L)$ s.t. $\forall T \ge T_0$, $L \notin Bad_T$.

Let $N \ge N_0 = 2^{T_0}$, and let T s.t. $2^{T-1} \le N < 2^T$. By Lemma 2,

(4) $$\left(\widehat{\rho}_N(L) - N\right)^2 \le T(T+1)2^T(\log(2)T-1)^{1+\varepsilon} = O\left(N(\log N)^{3+\varepsilon}\right)$$

For any $\eta > 0$, let N be such that $N \in \mathbb{N}$,

$Vol(\Omega_r) \in [N, N+1)$. Let $r_1 \leq r < r_2$ s.t.

$Vol(\Omega_{r_1}) = N$, $Vol(\Omega_{r_2}) = N+1$

$S_N = \Omega_{r_1} \subset \Omega_r \subset \Omega_{r_2} = S_{N+1}$. Then

(5) $\#(\Omega_{r_1} \cap L) - (N+1) \leq$

$\leq \#(\Omega_r \cap L) - Vol(\Omega_r) \leq \#(S_{N+1} \cap L) - N$

Taking square root in (4), both sides of (5) are $O(N^{\frac{1}{2}+\varepsilon})$.

So $\#\Omega_{r_1} \cap L - Vol(\Omega_r) \leq C N^{\frac{1}{2}+\varepsilon}$

and the opposite inequality is similar.

This proves thm, modulo lemma 2.

Pf of Lemma 2 Define

$$Bad_T = \left\{ L \in \mathcal{X}_n : \sum_{(N_1,N_2) \in K_T} \left({}_{N_1}\hat{S}_{N_2}(L) - (N_2 - N_1) \right)^2 > (T+1) 2^T ((\log 2) T + 1)^{1+\varepsilon} \right\}.$$

Let $$X = \sum_{(N_1,N_2)\in K_T} \left(\widehat{{}_{N_1}\rho_{N_2}}(L) - (N_2 - N_1) \right)^2,$$

so $$\mathbb{E}(X) = \sum_{(N_1,N_2)\in K_T} \mathrm{Var}\left(\widehat{{}_{N_1}\rho_{N_2}} \right)$$

Let $$a = \frac{1}{c}\left(\log(2)T - 1 \right)^{1+\varepsilon}$$

Then $$\mathrm{Bad}_T = \left\{ L \in \mathcal{X}_n : X > C(T+1)2^T \cdot a \right\}$$

$$\underset{\text{Lemma 1}}{\subset} \left\{ L \in \mathcal{X}_n : X > a\,\mathbb{E}(X) \right\}$$

By the Markov inequality,

$$m_{\mathcal{X}}(\mathrm{Bad}_T) < \frac{1}{a} = \frac{c}{(\log(2)T-1)^{1+\varepsilon}}$$

To prove (3), assume $N \leq 2^T$ and $L \notin \mathrm{Bad}_T$.

Express $[0,N)$ as a union of intervals $[N_1, N_2)$ where (N_1, N_2), that is

$$[0,N) = \bigsqcup_{(N_1,N_2)\in I_N} [N_1, N_2), \quad \text{where } I_N \subset K_T,\ \# I_N \leq T.$$

This possible by writing N in dyadic expansion, for example (taking $T=4$ as in previous example) $13 = 8+4+1$,

so $[0,13) = [0,8) \sqcup [8,12) \sqcup [12,13)$.

Then $\widehat{\rho}_N = \sum_{(N_1,N_2) \in \mathcal{I}_N} \widehat{{}_{N_1}\rho_{N_2}}$,

$$\widehat{\rho}_N(L) - N = \sum_{(N_1,N_2)\in \mathcal{I}_N} \left(\widehat{{}_{N_1}\rho_{N_2}}(L) - (N_2 - N_1) \right)$$

$$\leq \left(\sum_{(N_1,N_2)\in \mathcal{I}_N} \left(\widehat{{}_{N_1}\rho_{N_2}}(L) - (N_2-N_1) \right)^2 \right) \sqrt{T}$$

Cauchy-Schwarz
$$\sum_{i=1}^m a_i = \sum_{i=1}^m a_i \cdot 1 \leq \left(\sum_{i=1}^m a_i^2 \right)^{\frac{1}{2}} \cdot \sqrt{m}$$

$$\leq \sqrt{T} \left((T+1)\, 2^T \left(\log(2) T - 1 \right)^{1+\varepsilon} \right)^{\frac{1}{2}}$$

def of Bad_T

Now square both sides to obtain (3).