Subharmonic techniques in multiscale analysis: Lecture 3

W. Schlag (University of Chicago)

Maiori, September 2013

Cartan estimate

How do we control the large negative values of a sub-harmonic function? Cartan's estimate reduces it to $\|\mu\| \log |z|$.

Theorem

Fix $0 < \varepsilon \le 1$. *Let*

$$u(z) = \int_{\mathbb{C}} \log |z - \zeta| \, d\mu(\zeta) \tag{1}$$

for some positive finite measure μ. *For any* $0 < H < 1$ *there* $\exists \{D(z_j, r_j)\}_{j=1}^{\infty}$ *disjoint so that*

$$\sum_j r_j^{\varepsilon} \le H^{\varepsilon} \tag{2}$$

$$u(z) > -\|\mu\| \left[\varepsilon^{-1} + \log \frac{1}{H}\right] \quad \forall\, z \in \mathbb{C} \setminus \bigcup_{j=1}^{\infty} D(z_j, 5r_j). \tag{3}$$

Cartan theorem

- For $P(z) = \prod_{j=1}^{n}(z - z_j)$ one has $|P(z)| \geq (H/e)^n$ outside disks D_j with $\sum_j r_j \leq 5H$. Due to maximum principle can assume that each disk contains a zero.
- Typically can set $\varepsilon = 1$. However, sending $\varepsilon \to 0$ we get that $\dim\{u = -\infty\} = 0$, where dim refers to Hausdorff dimension.

Proof of Cartan: Fix $\varepsilon > 0$. For any $p > 0$ we say that z is p-good if

$$\mu(D(z,r)) \leq p\, r^{\varepsilon} \qquad \forall\, r > 0.$$

We now cover the set of bad points by disks with radius determined from the property of being a bad disk. Then we pass to a Vitali sub-cover.

Proof of Cartan

Vitali covering theorem: there are disjoint disks $\{D(z_j, r_j)\}_{j=1}^{\infty}$ (possible empty) with the property that

$$\mathcal{B}_{\varepsilon,p} := \{z \in \mathbb{C} \mid z \text{ is } p\text{-bad}\} \subset \bigcup_{j=1}^{\infty} D(z_j, 5r_j)$$

and

$$\sum r_j^{\varepsilon} \leq \frac{1}{p}\|\mu\|.$$

Setting $p = H^{-\varepsilon}\|\mu\|$, this latter inequality is exactly (2). Furthermore, if $z \notin \mathcal{B}_{\varepsilon,p}$, then

$$u(z) \geq \int_{|z-\zeta|\leq 1} \log|z-\zeta|\, d\mu(\zeta) = -\int_0^1 \frac{\mu(D(z,r))}{r}\, dr$$

Now we use the definition of *a good point* to bound this from below.

Indeed, if $z \notin \mathcal{B}_{\varepsilon,p}$, then

$$-\int_0^1 \frac{\mu(D(z,r))}{r}\,dr \geq -\int_0^H pr^{\varepsilon}\frac{dr}{r} - \int_H^1 \|\mu\|\,\frac{dr}{r}$$
$$= -\|\mu\|(\varepsilon^{-1} + \log\frac{1}{H}),$$

as claimed. □

We now consider two very different examples:

- $\mu = n\delta_0$. Then
$$u(z) = n\log|z-1|$$
$$|\{x \in \mathbb{T} : \frac{1}{n}u(e(x)) < -\lambda\}| \leq \exp(-\lambda)$$
- $\mu = \sum_{j=1}^{n} \delta_{\zeta_j}$ where ζ_j are n^{th} roots of unity. Then
$$u(z) = \log|z^n - 1|$$
$$|\{x \in \mathbb{T} : \frac{1}{n}u(e(x)) < -\lambda\}| \leq \exp(-n\lambda)$$

Applying Cartan

Both of these subharmonic functions have Riesz mass n, and mean zero over the torus.

Conclusion:

- For $\mu = n\delta_0$ the Cartan estimate is essentially optimal
- But not for $\mu = \sum_{j=1}^{n} \delta_{\zeta_j}$ where ζ_j are n^{th} roots of unity. Here Cartan is extremely wasteful.

However, *the second example is much closer to what we need*. Clearly, any kind of large deviation estimate that we have seen up to now **does not follow directly** from Cartan. But as we shall see, we can still apply it provided we localize to a smaller portion of the Riesz mass.

It might be natural to try to establish *structure* of the Riesz mass, or for the distribution of zeros. This can be done by the avalanche principle, but it can be difficult and requires $L > 0$ (see the work of Goldstein-S, GAFA 2008, Annals 2009). Instead, we shall now *use the dynamics to* **localize to small portions of the Riesz measure**. This cuts down the mass.

Derivation of LDT from Cartan

Diophantine condition:

$$\|n\omega\| \geq \frac{c}{n^a} \quad \forall n \geq 1 \tag{4}$$

where $c = c(\omega) > 0$, $a > 1$, and $\|\cdot\|$ is the distance to the nearest integer. Riesz representation:

$$u(z) = \log \|M_n(z, E)\| = \int \log |z - \zeta| \, \mu(d\zeta) + h(z)$$

on some rectangle $R \supset [0, 1]$. Then $0 \leq u(z) \lesssim n$, and $\mu(R) + \|h\|_{L^\infty(R)} \lesssim n$. Fix a small $\delta > 0$ and take n large. Then there is a disk $D_0 = D(x_0, n^{-2\delta})$ with the property that $\mu(D_0) \lesssim n^{1-2\delta}$. Write

$$\begin{aligned} u(z) &= u_1(z) + u_2(z) \\ &= \int_{D_0} \log |z - \zeta| \, \mu(d\zeta) + \int_{\mathbb{C}\setminus D_0} \log |z - \zeta| \, \mu(d\zeta) \end{aligned} \tag{5}$$

The disks in the proof

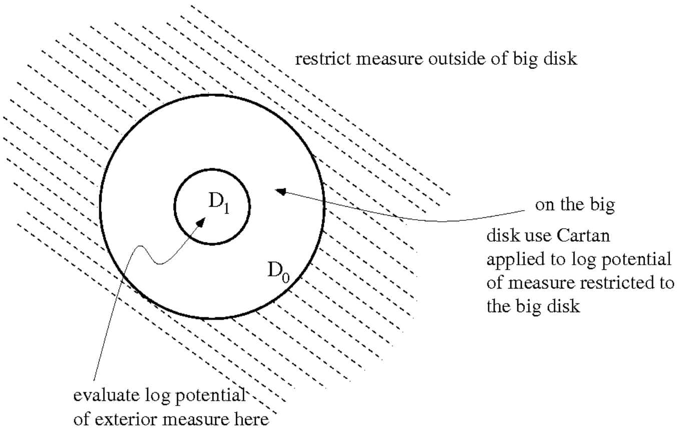

Figure: The two disks D_0 and D_1

Set $D_1 = D(x_0, n^{-3\delta})$. Then by differentiation

$$|u_2(z) - u_2(z')| \lesssim \|\mu\|(n^{-2\delta})^{-1} n^{-3\delta} = n^{1-\delta} \quad \forall z, z' \in D_1$$

Cartan's theorem applied to $u_1(z)$ yields disks $\{D(z_j, r_j)\}_j$ with $\sum_j r_j \lesssim \exp(-2n^\delta)$ and so that

$$u_1(z) \gtrsim -n^{1-\delta} \qquad \forall z \in \mathbb{C} \setminus \bigcup D(z_j, r_j)$$

Since also $u_1 \leq 0$ on D_1 as well as $|h(z) - h(z')| \lesssim n|z - z'|$, it follows that

$$|u(z) - u(z')| \lesssim n^{1-\delta} \quad \forall z, z' \in D_1 \setminus \bigcup D(z_j, r_j) \tag{6}$$

We now use the shift dynamics to move an arbitrary point from $\mathbb{T}$ into the small disk D_1. This can be done by shifting fewer than some inverse power of the radius of the disk, and the expense controlled by almost invariance under shift – as in the Fourier series based proof.

LDT from Cartan

From the Diophantine property, $\forall\, x, x' \in \mathbb{T}\ \exists\, k, k' \lesssim n^{4\delta}$ s.t.

$$x + k\omega, x' + k'\omega \in D_1 \qquad \text{mod } \mathbb{Z}$$

In order to avoid the Cartan disks $\bigcup_j D(z_j, r_j)$ we need to remove a set $\mathcal{B} \subset \mathbb{T}$ of measure $\lesssim \exp(-n^{\delta})$. Then from the almost invariance, for any $x, x' \in \mathbb{T} \setminus \mathcal{B}$,

$$|u(x) - u(x')| \lesssim n^{4\delta} + n^{1-\delta} \lesssim n^{1-\delta}$$

Therefore, the function does not deviate from its *mean* by more than this amount.

Conclusion :

$$|\{x \in \mathbb{T} : |\log \|M_n(x, E)\| - nL_n(E)| > n^{1-\delta}\}| < \exp(-n^{\delta})$$

which is a weak form of the LDT (with basically any Diophantine condition). But it is enough for Anderson Localization, for example.

Cartan for analytic functions

Theorem

Let f *be analytic on* $|z| \leq 10$*, and satisfies* $|f(z)| \leq M$*,* $f(0) = 1$*. There exists an absolute constant* C *such that for any* $0 < \eta < 1$ *one has*

$$\log|f(z)| \geq -H(\eta)\log M, \quad H(\eta) = \log\left(\frac{C}{\eta}\right)$$

for all $z \in D(0,1) \setminus \bigcup_j D_j$ *where* $\sum_j r_j \leq \eta$*.*

This follows from the previous Cartan, but first need to factor out the zeros from the function. See B. Ya. Levin's book, *Lectures on entire functions*, AMS 1996, page 79.

Cartan in higher dimensions

How much of this survives for other types of dynamics, e.g. on the torus $\mathbb{T}^d$? The same strategy applies to ergodic multi-dim shifts. First, we define Cartan sets inductively:

Definition

Let $0 < H < 1$. We say that $\mathcal{B} \in \mathrm{Car}_1(H)$ if $\mathcal{B} \subset \bigcup_j D(z_j, r_j)$ with

$$\sum_j r_j \leq C_0\, H.$$

If $d \geq 2$ and $\mathcal{B} \subset \mathbb{C}^d$ we say that $\mathcal{B} \in \mathrm{Car}_d(H)$ if there exists $\mathcal{B}_0 \in \mathrm{Car}_{d-1}(H)$ so that

$$\mathcal{B} = \{(z_1, z_2, \ldots, z_d) : (z_2, \ldots, z_d) \in \mathcal{B}_0 \ \text{ or } \ z_1 \in \mathcal{B}(z_2, \ldots, z_d) \ \text{ for some } \ \mathcal{B}(z_2, \ldots, z_d) \in \mathrm{Car}_1(H)$$

Cartan sets in higher dimensions

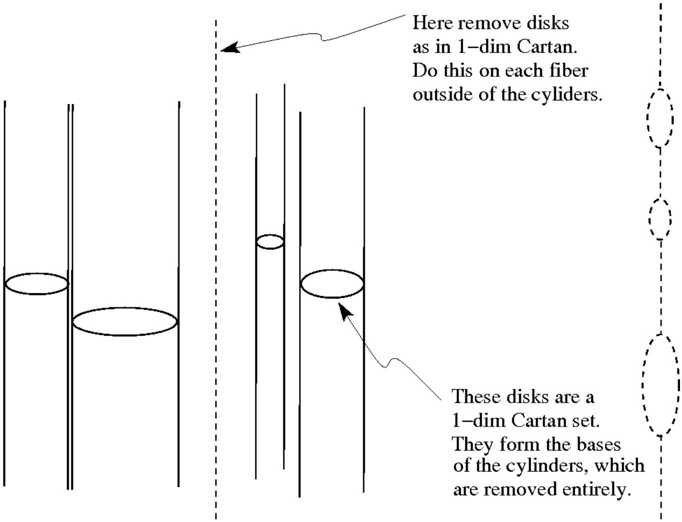

Figure: An illustration of a Cartan-2 set

Local regularity for pluri-subharmonic functions

Theorem (Goldstein-S. Annals 2001)

Suppose $u : D(0,2)^d \to [-1,1]$ *continuous,* $u \le 1$, $\sup_{x \in D(0,1)^d} u(x) = 0$, *and s-h in each variable, i.e.,* $z_1 \mapsto u(z_1, z_2, \ldots, z_d)$ *is s-h* $\forall\, (z_2, \ldots, z_d) \in D(0,2)^{d-1}$ *etc. Given* $r \in (0,1)$ *there exists* $\Pi = D(x_1^{(0)}, r) \times \ldots \times D(x_d^{(0)}, r) \subset \mathbb{C}^d$ *polydisk with* $x_1^{(0)}, \ldots, x_d^{(0)} \in [-1,1]$ *and a Cartan set* $\mathcal{B} \in \mathrm{Car}_d(H)$ *with* $H = \exp(-r^{-\beta})$ *s.t.*

$$|u(z_1, \ldots, z_d) - u(z_1', \ldots, z_d')| < C\, r^{\beta} \quad \forall\, (z_1, \ldots, z_d), (z_1', \ldots, z_d') \in \Pi \setminus \mathcal{B}$$

Constants $\beta, C > 0$ *depend only on the dimension* d.

In fact, for small r most points $x_1^{(0)}, \ldots, x_d^{(0)} \in [-1,1]$ can be chosen as the center of Π. Proof is based on Cartan's theorem, and elementary – but a bit tricky – analysis (Hardy-Littlewood maximal function).

LDT in higher dimension

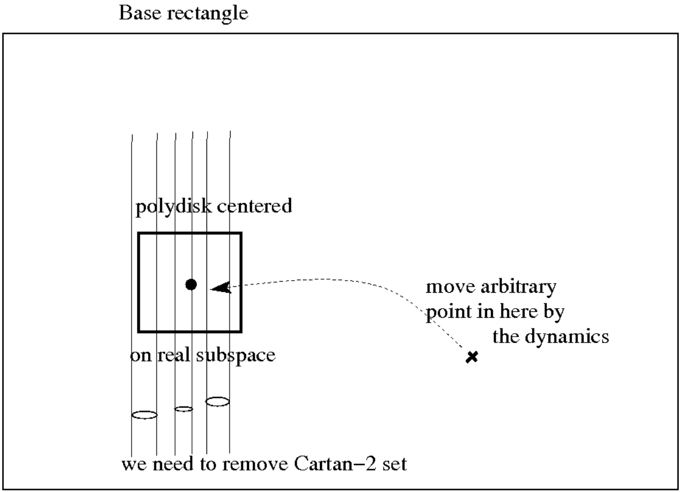

Figure: Moving a point into the polydisk of small variation

LDT for shifts on higher-dimensional tori

Definition

Diophantine condition: $\boldsymbol{\omega} \in \mathbb{T}^d$ *satisfies*

$$\|\boldsymbol{\omega} \cdot \mathbf{k}\| \geq \frac{C(\varepsilon_1)}{|\mathbf{k}|^{d+\varepsilon_1}} \quad \forall\, \mathbf{k} \in \mathbb{Z}^d$$

A.e. $\boldsymbol{\omega}$ *satisfies this for any* $\varepsilon_1 > 0$.

Set $r = n^{-\tau}$, use shift by $\boldsymbol{\omega}$ to move any point into a polydisk Π from above for $u(\boldsymbol{\theta}) = n^{-1} \log \|M_n(\boldsymbol{\theta}, E)\|$. Here M_n is the propagator relative to the shift dynamics on $\mathbb{T}^d$ defined by $\boldsymbol{\omega}$.
Conclusion: there exists $\sigma > 0$ so that

$$|\{\boldsymbol{\theta} \in \mathbb{T}^d : |n^{-1} \log \|M_n(\boldsymbol{\theta}, E)\| - L_n(E)| > n^{-\sigma}\}| < \exp(-n^{\sigma})$$

for large n.
This is sufficient, with other ingredients such as semi algebraic set machinery, to prove Anderson localization.

Shift dynamics on $\mathbb{T}^d$, $d \geq 2$

Weaker regularity results on integrated density of states using the Avalanche Principle/LDT approach. Gives only modulus of continuity $\exp(-|\log|E-E'||^b)$ for $b>0$ small. On the other hand, expect *better regularity for more frequencies*, since more "randomness" in the system means we are closer to i.i.d. potentials.

So we are far from understanding multi-dim shift potentials. Is there a sharp LDT for multi-dim shifts, analogous to what we have in one dimension? In other words, for a.e. $\omega \in \mathbb{T}^d$ do we have for any $\kappa > 0$ some $c = c(\kappa, \omega) > 0$ with

$$|\{\boldsymbol{\theta} \in \mathbb{T}^d : |n^{-1}\log\|M_n(\boldsymbol{\theta}, E)\| - L_n(E)| > \kappa\}| < \exp(-c\,n)$$

for all $n \geq 1$?

Note: Fourier transform can also be used in higher dimensions, but exponential bounds are tricky, weaker power-type estimates easy. Upgrade these to exponential estimates by the **splitting lemma**.

BMO

The following function space is both natural and of fundamental importance in this context. Recall the space of functions of *bounded mean oscillation (BMO)* on $\mathbb{T}$

$$\sup_{J \subset \mathbb{T}} \frac{1}{|J|} \int_J |f - f_J| \, dx = \|f\|_{\mathrm{BMO}} \tag{7}$$

Can be done on any interval, dimension etc. Typical example:

$$u(x) = \log |e(x) - \zeta| \in \mathrm{BMO}(\mathbb{T}) \quad \text{uniformly in} \quad \zeta \in \mathbb{C}$$

Discrete logarithm:

$$u(x) = \sum_{n=0}^{\infty} \chi_{[0<x<2^{-n}]} \in \mathrm{BMO}([0,1])$$

This example displays a characteristic of BMO: on the x-axis we have a geometric scale, but in y-axis we have an arithmetic scale.

BMO

In fact, this comes directly from the definition of BMO.

$$|u_J - u_{J^*}| \leq \frac{|J^*|}{|J|} \|u\|_{\mathrm{BMO}}$$

if $J \subset J^* \subset \mathbb{T}$. So geometric scale in x, arithmetic scale in u. The natural conclusion of these ideas is the following fundamental lemma.

John-Nirenberg inequality: For all $u \in \mathrm{BMO}$, $\lambda > 0$

$$|\{x \in \mathbb{T} \,:\, |u(x) - \langle u \rangle| > \lambda\}| \leq C \exp\big(-c\lambda/\|u\|_{\mathrm{BMO}}\big) \qquad (8)$$

where c, C are absolute constants. Thus $u \in L^p$ for all finite p, and we can use any finite power in (7). The proof of (8) is based on a Calderón-Zygmund decomposition at arithmetic sequence of heights.

BMO

Consider a subharmonic function as in the Riesz representation:

$$u(x) = \int \log|e(x) - \zeta|\, \mu(d\zeta) + h(x)$$

where size of μ, h are bounded by some constant B (Riesz mass). Then clearly $\|u\|_{\mathrm{BMO}(\mathbb{T})} \lesssim B$. But BMO is more precise: let $\{\zeta_j\}_{j=1}^N \subset \mathbb{T}$ and $P(z) = \prod_{j=1}^N (z - \zeta_j)$. Then

$$\sup_{z \in \mathbb{T}} |P(z)| < e^{\tau} \Rightarrow \|\log|P(e(\cdot))|\|_{BMO} \simeq D_N(\{\zeta_j\}) \lesssim \sqrt{N\tau}$$

where

$$D_N(\{\zeta_j\}) = \sup_{I \subset \mathbb{T}} \left| \#\{j \,:\, \zeta_j \in I \mod 1\} - N|I| \right|$$

is the usual discrepancy. This is a classical result of *Erdös-Turan*.

Hilbert transform and BMO

BMO is central to harmonic analysis: Hilbert transform takes $L^\infty \to \mathrm{BMO}$, but NOT $L^\infty \to L^\infty$:

$$(Hf)(x) = \mathrm{PV} \int_{\mathbb{R}} \frac{f(y)}{x-y}\, dy \quad \text{line}$$

$$(Hf)(\theta) = \mathrm{PV} \int_{\mathbb{T}} \cot(\pi(\theta - \varphi)) f(\varphi)\, d\varphi \quad \text{circle}$$

The Hilbert transform is also related to subharmonic functions:

$$u(x) = \int_{\mathbb{R}} \log|x-y|\, f(y)\, dy$$

$$u'(x) = (Hf)(x)$$

These properties are key to the Erdös-Turan inequality, see also the splitting lemma later on. The dual of BMO is the real Hardy space H^1. The Hilbert transform is bounded $H^1 \to L^1$. For all this see, for example, Volume 1 of my book with Camil Muscalu.

Suppose

$$H(z) = -\Delta_{\mathbb{Z}^2} + V(z) \tag{9}$$

where $V(z) : \mathbb{Z}^2 \times \mathbb{C} \to \mathbb{C}$ analytic in z on some neighborhood of $\mathbb{T}$, Hermitian for $z \in \mathbb{R}$. Cube $\Lambda \subset \mathbb{Z}^2$ of side length N, collection $\Lambda_\alpha \subset \Lambda$ of smaller cubes of side length n, $N = n^C$. Let $\mathcal{C} := \bigcup_\alpha \Lambda_\alpha$ and assume that for some $z_0 \in \mathbb{T}$ the operator $\|(H_{\Lambda\setminus\mathcal{C}}(x_0) - E)^{-1}\| < \exp(n^b)$. Write, with the restriction operator $R_\mathcal{C}$,

$$\begin{aligned} H_\Lambda(z) - E &= \begin{pmatrix} R_\mathcal{C}(H(z)-E)R_\mathcal{C} & R_\mathcal{C}(H(z)-E)R_{\Lambda\setminus\mathcal{C}} \\ R_{\Lambda\setminus\mathcal{C}}(H(z)-E)R_\mathcal{C} & R_{\Lambda\setminus\mathcal{C}}(H(z)-E)R_{\Lambda\setminus\mathcal{C}} \end{pmatrix} \\ &= \begin{pmatrix} A(z) & B(z) \\ C(z) & D(z) \end{pmatrix} \end{aligned}$$

We have $D(x_0)$ invertible, and stays so in the disk $D_0 := D(x_0, \exp(-n^b/2))$ with the same bound.

Good and bad cubes

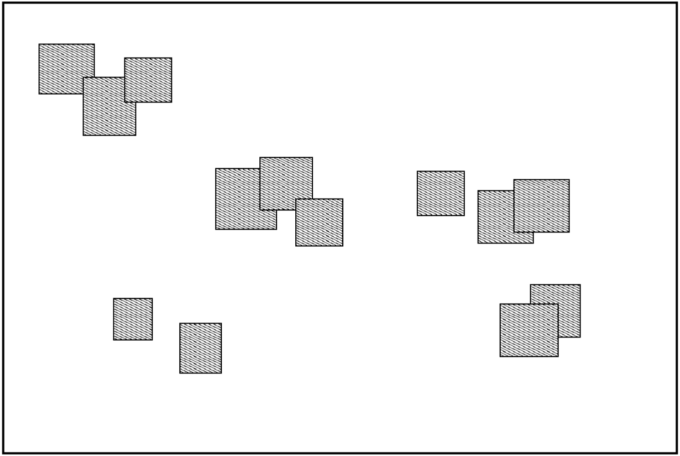

Figure: The geometry of the resolvent identity

Matrix-valued Cartan theorem, introduction

Feshbach formula: On D_0 matrix $H_\Lambda(z) - E$ is invertible iff

$$M(z) := A(z) - B(z)D(z)^{-1}C(z)$$

is invertible, and

$$\begin{aligned} \|M(z)^{-1}\| &\lesssim \|(H_\Lambda(z) - E)^{-1}\| \\ &\lesssim (1 + \|D(z)^{-1}\|^2)(1 + \|M(z)^{-1}\|) \end{aligned} \tag{10}$$

Clearly, we need a *nondegeneracy condition*. What if $M(z)$ is constant, noninvertible? The point is that we can use the information from the previous smaller scale n. This only gives very weak bounds in terms of the measure estimates, but they are enough to guarantee nondegeneracy. In fact, from the measure estimate at that scale, we can make sure that **all small cubes** in Λ are **good** for, say 50% of the points in D_0.
Now we use the *first inequality* in (10) to conclude the following.

From the smaller scale n we know that $\|M(z)^{-1}\| < e^{n^b}$ at say, 50% of points of $D_0 \cap \mathbb{R}$. Use self-adjointness to get a lower bound $|\det M(z)| \geq \|M(z)^{-1}\|^{-\#\mathcal{C}}$ at those points.
Jensen controls the Riesz mass: maximum - value at one point.
Therefore: Riesz mass of the subharmonic function $\log|\det M(z)|$ is at most $n^b \#\mathcal{C}$. So if this is at most N^b, $b < 1$, then we can apply Cartan (or John-Nirenberg) to conclude that with $\rho = \exp(-n^b/2)$ there is the measure estimate

$$|\{x \in (x_0 - \rho, x_0 + \rho) \, : \, \log\|(H_\Lambda(x) - E)^{-1}\| > \lambda\}| \lesssim \rho \exp(-\lambda/(\#\mathcal{C} \cdot n^b)) \tag{11}$$

which is the desired estimate at large scale N. NOTE: Only the range $\lambda < N^{1-\varepsilon}$ is useful for resolvent identity, since the exponential decay of the Green function can only be on the scale of the *side-length* of the big cube, i.e., N.

Matrix-valued Cartan

Theorem

Let $H(z)$ be analytic $N \times N$-matrix valued function, defined on disk $D_0 = D(x_0, \rho)$, Hermitian on real axis. With $B_1, B_2 \geq 1$,

- $\|H(z)\| \leq B_1$ *in* D_0
- *For each* $x \in D_0 \cap \mathbb{R}$ *there exists* $\Lambda \subset [1, N]$ *s.t.* $|\Lambda| < M$,

$$\|(R_{\Lambda^c} H(x) R_{\Lambda^c})^{-1}\| < B_2 \tag{12}$$

- *We have the non degeneracy condition*

$$|\{x \in D_0 \cap \mathbb{R} \ : \ \|H(x)^{-1}\| > B_3\}| \leq \frac{\rho}{100 B_1 B_2} \tag{13}$$

Then there is the measure estimate

$$\begin{aligned} &|\{x \in D(x_0, \rho/2) \cap \mathbb{R} \ : \ \log \|H(x)^{-1}\| > \lambda\}| \\ &\lesssim \rho \exp\Big(-c \frac{\lambda}{M \log(M + B_1 + B_2 + B_3)}\Big) \end{aligned} \tag{14}$$

Proof of matrix valued Cartan

First, we show that (12) is stable:

$$R_{\Lambda^c} H(z) R_{\Lambda^c} = R_{\Lambda^c} H(x_1) R_{\Lambda^c} + R_{\Lambda^c}(H(z) - H(x_1)) R_{\Lambda^c} \qquad (15)$$

By Cauchy estimate: $\|H(z) - H(x_1)\| \leq 2B_1 \rho^{-1}|z - x_1|$ for $|z - x_1| < \rho/2$. Apply Neuman series to (15):

$$\|(R_{\Lambda^c} H(z) R_{\Lambda^c})^{-1}\| \leq 2B_2 \quad \forall\, |z - x_1| < \rho/(2B_1 B_2) =: \rho^*$$

The point here is that we are not changing Λ. We apply Feshbach formula as before. Need to invert

$$D(z) := R_\Lambda H(z) R_\Lambda - R_\Lambda H(z) R_{\Lambda^c}(R_{\Lambda^c} H(z) R_{\Lambda^c})^{-1} R_{\Lambda^c} H(z) R_\Lambda$$

which satisfies $\|D(z)\| \lesssim B_1^2 B_2$. By non degeneracy condition (13) inside of $D(x_1, \rho^*) \cap \mathbb{R}$ have many points x where $\|H(x)^{-1}\| < B_3$ and thus also $\|D(x)^{-1}\| \lesssim B_3$. We now recall that $H(x)$ Hermitian for real x, and thus $D(x)$ is also Hermitian. This implies that for those x

$$|\det D(x)| \geq \|D(x)^{-1}\|^{-M}$$

Proof of matrix valued Cartan

Therefore, in $D(x_1, \rho^*) \cap \mathbb{R}$ have many points where

$$\log|\det D(x)| \geq -M \log B_3$$

We also have the upper bound on all of $D(x_1, \rho^*)$

$$u(z) := \log|\det D(z)| \leq M \log(M + B_1 + B_2)$$

Now apply Jensen/Cartan estimate to $u(x_1 + \rho^*\zeta)$ to conclude that

$$\begin{aligned} &|\{x \in D(x_1, \rho^*) \cap \mathbb{R} \,:\, \log|\det D(x)| < -\lambda\}| \\ &\lesssim \rho^* \exp\Big(-c\frac{\lambda}{M\log(M + B_1 + B_2 + B_3)}\Big) \end{aligned}$$

But if $x \notin$ set on the left-hand side then by Cramer's rule

$$\log\|D(x)^{-1}\| \leq CM \log(M + B_1 + B_2) + \lambda \lesssim \lambda$$

Now cover the original interval by shorter ones of length ρ^*. □

Skew shift

Let $T : \mathbb{T}^2 \to \mathbb{T}^2$, $T(x,y) := (x+\omega, x+y) \mod \mathbb{Z}^2$ (higher-dimensional versions analogous). Then for $n \geq 1$, $T^n(x,y) = (x+n\omega, nx+y+n(n-1)\omega/2)$. Define potential $v_n(x,y) := V(T^n(x,y))$ where V analytic on $\mathbb{T}^2$.

Big problem: extend $z \mapsto u(z,y) := n^{-1} \log \|M_n(z,y)\|$ to strip $|\operatorname{Im} z| < 1$. **But this costs a big loss in the Riesz mass:** $\mu(\mathbb{C}) \lesssim n$. To see this another way, suppose we take product

$$u(z) = \prod_{j=1}^{n} e(jz)$$

Then $|u(z)| \leq e^{n^2/2}$ for $|\operatorname{Im} z| \leq 1$, and $n^{-1} \log |u(z)| \lesssim n$.

LDT does not follow from any methods presented so far, they cannot absorb this kind of mass. Again we expect better properties due to "more randomness" but it becomes harder to analyze. In Bourgain, Goldstein, S., CMP 2001 this difficulty was overcome by means of a certain analytical device, the **splitting lemma**. LDT was established there for large disorders, also using the AP.

The splitting lemma

Lemma (Bourgain-Goldstein-S., CMP 2002, Bourgain's book)

Let u *be subharmonic on some annulus*

$$\mathcal{A}_\rho = \{1 - \rho < |z| < 1 + \rho\}, \qquad \rho > 0$$

small. Suppose $u \leq B$ *on* $\mathcal{A}_\rho$ *(where* $B \geq 1$*) and* $\sup_{\mathbb{T}} u = 0$*. Assume that* $u = u_0 + u_1$ *on* $\mathbb{T}$ *where* $\|u_0\|_{L^\infty(\mathbb{T})} \leq \varepsilon_0$ *and* $\|u_1\|_{L^1(\mathbb{T})} \leq \varepsilon_1$*. Then*

$$\|u\|_{\mathrm{BMO}(\mathbb{T})} \leq C(\varepsilon_0 + \sqrt{B\varepsilon_1}) \tag{16}$$

where $C = C(\rho)$*.*

Note: From Riesz representation alone we get $\|u\|_{\mathrm{BMO}(\mathbb{T})} \leq CB$. The John-Nirenberg inequality then yields something similar to Cartan, but the latter is more precise (structure of the disks can be very useful, since we often also need to bound their number – complexity bounds!).

LDT via the splitting lemma, a road-map

We shall now elaborate an "abstract scheme" of how one might obtain LDT inductively and for large disorder.

Two main ingredients:

(A) the avalanche principle

(B) the quantitative Birkhoff ergodic theorem (which is the general estimate on the average of shifts of subharmonic functions)

These yield a preliminary LDT, where the size of the deviation set is only polynomially small (relative to the big scale), which is essentially the deviation set from the previous scale. Hence the splitting lemma acts as a "boosting" argument, in which we allow the size of the deviation to increase slightly, with the benefit of obtaining a much better estimate on the size of the deviation set.

Applying the splitting lemma

Take some ergodic transformation $T : \mathbb{T} \to \mathbb{T}$ and non constant analytic function V near $\mathbb{T}$. We let $M_n(x, E)$ be the propagator matrices for the equation

$$\psi_{n+1} + \psi_{n-1} + \lambda V(T^n x)\psi_n = E\psi_n$$

At some initial scale $n_0 \gg 1$ we take λ so large that we have a bound of the form, with $n = n_0$

$$|\{x : |n^{-1} \log \|M_n(x, E)\| - L_n(E)| > n^{-\sigma}\}| \leq \exp(-n^\sigma) \quad (17)$$

This follows from $|\{x \in \mathbb{T} \ : \ |V(x) - v_0| \leq t\}| \leq Ct^b$ for all v_0, t and some constants $C, b > 0$ depending on V. Freeze such a large λ. Next, suppose that for all $N \geq 1$

$$z \mapsto N^{-1} \log \|M_N(z, E)\| \quad (18)$$

is s-h on some fixed complex neighborhood of $\mathbb{T}$ with Riesz mass $O(N^A)$, $A \geq 0$ fixed.

Applying the splitting lemma

Now we show how to prove (17) at some much larger scale $N = n_1$ without increasing λ. Idea is to apply the avalanche principle by writing M_N as product of shifted M_{n_0}. Modulo technicalities involving $L_{n_0}(E), L_{2n_0}(E)$ we can do this just from LDT (17) provided $\log N \ll n_0$.
Thus, off a bad set $\mathcal{B} \subset \mathbb{T}$ of measure $O(N^2 e^{-n_0^\sigma})$ we have

$$\Big| N^{-1} \log \|M_N(x,E)\| + N^{-1} \sum_{k=1}^{N} n_0^{-1} \log \|M_{n_0}(T^k x, E)\| - N^{-1} \sum_{k=1}^{N} (2n_0)^{-1} \log \|M_{2n_0}(T^k x, E)\| \Big| \leq C(Ne^{-Ln_0} + n_0 N^{-1})$$

The advantage here is that we are averaging a s-h function $n_0^{-1} \log \|M_{n_0}(T^k x, E)\|$ of relatively small Riesz mass over a long orbit of the dynamics. This is the point were one needs to exploit *quantitative ergodic properties* of the dynamics.

Applying the splitting lemma

We now therefore assume that we have *(not circular!)*

$$\left|\{x : |N^{-1}\sum_{k=1}^{N} n_0^{-1}\log\|M_{n_0}(T^k x, E)\| - L_{n_0}(E)| > N^{-2\sigma}\}\right| \leq N^{-A-1}$$

and similarly for $2n_0$.

Collecting everything we see that we can take in the splitting lemma

$$\varepsilon_0 := Ne^{-Ln_0} + n_0 N^{-1} + N^{-2\sigma},\ \varepsilon_1 := N^2 e^{-n_0^\sigma} + N^{-A-1}$$

This implies that provided $N = e^{n_0^\delta}$, $0 < \delta < \sigma$

$$\left\| N^{-1}\log\|M_N(x,E)\| \right\|_{\mathrm{BMO}(\mathbb{T})} \lesssim Ne^{-Ln_0} + n_0 N^{-1} + N^{-2\sigma}$$
$$+ N^{-\frac{1}{2}} + N^{(2+A)/2} e^{-n_0^\sigma/2} \lesssim N^{-2\sigma}$$

whence (17) follows from J-N at scale N.

The splitting lemma in $\mathbb{T}^2$

This strategy is employed in [B-G-S, 2002] for the skew shift, but we need a 2-dim splitting lemma, see [B], p 22.

Lemma

$u(x,y) : \mathbb{T}^2 \to \mathbb{R}$ *extends pluri-s-h on* $\mathcal{A}_\rho \times \mathcal{A}_\rho$, *with* $u \le B$ *and* $\sup_{\mathbb{T}^2} u = 0$. *Assume*

$$|\{(x,y) \in \mathbb{T}^2 : |u(x,y) - \langle u \rangle| > \varepsilon_0\}| < \varepsilon_1$$

Then

$$|\{(x,y) \in \mathbb{T}^2 : |u(x,y) - \langle u \rangle| > \varepsilon_0^{\frac{1}{4}}\}| < \exp\big(- c\delta^{-1}\big)$$

$$\delta = \varepsilon_0^{\frac{1}{4}} + (B/\varepsilon_0)^{\frac{1}{2}} \varepsilon_1^{\frac{1}{4}}$$

This basically follows by slicing and the 1-dimensional result.

Proof of the splitting lemma in one dimension

Riesz representation: $\mu \geq 0$, $\mu(\mathbb{C}) \leq B$, h harmonic extension,

$$u(x) = \int_{\mathcal{A}_\rho} \log|e(x) - \zeta|\,\mu(d\zeta) + h(x) =: v(x) + h(x) \tag{19}$$

Then, with $\mathcal{H}$ being the Hilbert transform,

$$\partial_x v(x) = \mathcal{H}\nu,$$
$$\frac{d\nu}{dx}(x) = 2\pi \int \frac{|\zeta|\cos(2\pi(x-y))}{|e(x-y) - |\zeta||^2}\,\mu(d\zeta), \;\; \zeta = |\zeta| e(y)$$

Let P_ε be the mollifying operator with kernel $\varepsilon^{-1}\varphi(t/\varepsilon)$, $\varepsilon > 0$ to be chosen. φ a smooth mass-normalized bump function. Then

$$u = v - P_\varepsilon v + h - P_\varepsilon h + P_\varepsilon u_0 + P_\varepsilon u_1$$
$$\|h - P_\varepsilon h\|_\infty \leq B\varepsilon, \;\; \|P_\varepsilon u_0\|_\infty \leq \varepsilon_0, \;\; \|P_\varepsilon u_1\|_\infty \leq \varepsilon_1/\varepsilon$$

Now $v - P_\varepsilon v = \partial_x^{-1}\mathcal{H}[\nu - P_\varepsilon \nu] = \mathcal{H}\partial_x^{-1}[\nu - P_\varepsilon \nu]$.

Proof of splitting lemma

Then

$$\|v - P_\varepsilon v\|_{\mathrm{BMO}} \le \|\partial_x^{-1}[\nu - P_\varepsilon \nu]\|_\infty \le \sup_J |\langle \nu - P_\varepsilon \nu, \chi_J\rangle| \lesssim \sup_{\mathrm{diam}(I)=\varepsilon} \nu(I)$$

Finally, with τ_J a smooth bump adapted to J, $\mathrm{diam}(J) = \varepsilon$

$$\begin{aligned}\nu(J) &\lesssim \langle \nu, \tau_J\rangle = |\langle \partial_x^{-1}\mathcal{H}\nu, \partial_x \mathcal{H}\tau_J\rangle| = |\langle v, \partial_x \mathcal{H}\tau_J\rangle| \\ &\lesssim |\langle u_0, \partial_x \mathcal{H}\tau_J\rangle| + |\langle u_1, \partial_x \mathcal{H}\tau_J\rangle| + |\langle h, \partial_x \mathcal{H}\tau_J\rangle| \\ &\lesssim \|u_0\|_\infty \|\mathcal{H}\partial_x \tau_J\|_1 + \|u_1\|_1 \|\partial_x \mathcal{H}\tau_J\|_\infty + \|\tau_J\|_1 \|\partial_x \mathcal{H} h\|_\infty \\ &\lesssim \varepsilon_0 + \varepsilon_1/\varepsilon + B\varepsilon\end{aligned}$$

Conclusion:

$$\|u\|_{\mathrm{BMO}} \lesssim \varepsilon_0 + \varepsilon_1/\varepsilon + B\varepsilon \lesssim \varepsilon_0 + \sqrt{B\varepsilon_1}$$

by setting $\varepsilon = \sqrt{B/\varepsilon_1}$.

Spectrum of ergodic Schrödinger operators

For self-adjoint operators

$$(H_x\psi)_n = \psi_{n+1} + \psi_{n-1} + v_n(x)\psi_n, \quad n \in \mathbb{Z}$$

with $v_n(x)$ an "ergodic potential", i.e., $v_n(x) = V(T^n x)$ and $T : X \to X$ ergodic transformation on a probability space X, and $V : X \to \mathbb{R}$ measurable. Then there exists *fixed* compact set $K \subset \mathbb{R}$ with $\mathrm{spec}(H_x) = K$ for a.e. $x \in X$. This follows from ergodic theorem and property of the spectral resolution E_x of H_x

$$E_x = S^{-1} \circ E_{Tx} \circ S, \quad S = \text{ right shift}$$

In addition, $\mathrm{spec}_{ac}(H_x), \mathrm{spec}_{ac}(H_x), \mathrm{spec}_{sc}(H_x)$ are also deterministic. Eigenvalues are **NOT** deterministic, but their **closure** is.
Anderson localization means precisely that $\mathrm{spec}_{pp}(H_x) = \mathrm{spec}(H_x)$ and eigenfunctions decay exponentially. *Most famous problem in this area:* Anderson conjecture in three dimensions for the random case.

Spectrum of ergodic Schrödinger operators

For 1-dimensional shift and analytic potential known in many cases that $\mathrm{spec}(H_x)$ is a Cantor set, i.e., the gaps are everywhere dense. Goes back to Marc Kac (ten martini problem etc.). J. Puig, Avila-Jitomirskaya verified this for Harper operator (almost Mathieu) and irrational rotation number. Their arguments are non-constructive and indirect.

Goldstein-S., Annals 2009: similar result for *general analytic potential*, a.e. frequency, assuming $L > 0$. The argument is **constructive, and iterative**, in the spirit of a KAM argument. Control the distance between most eigenvalues for some finite volume

$$|E_j^{(N)}(x,\omega) - E_k^{(N)}(x,\omega)| > e^{-N^\delta}$$

and then identify pre-gaps on finite volume by the *formation of double resonances*. Control non-collapse of these gaps at larger scales by zero counts of determinants.

For two and more frequencies it is not expected to find gaps: prove that all gaps are closed.

Basic mechanism behind gap formation

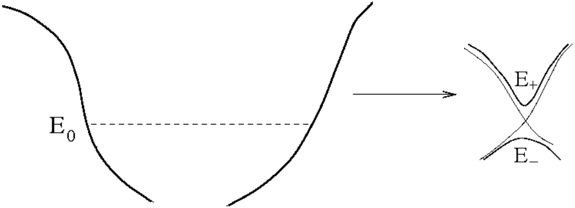

Figure: Crossing of graphs of eigenvalues create gap

$$\det \begin{pmatrix} \lambda_1(x) - E & \varepsilon \\ \varepsilon & \lambda_2(x) - E \end{pmatrix} = 0, \quad \lambda_1(x_0) = \lambda_2(x_0) = E_0$$
$$E_{\pm}(x) = \frac{1}{2}(\lambda_1(x) + \lambda_2(x)) \pm \sqrt{(\lambda_1(x) - \lambda_2(x))^2 + 4\varepsilon^2} \tag{20}$$

This is a reflection of the fact that for the Dirichlet problem eigenvalues are simple. On the level of eigenfunctions the following is going on:

Basic mechanism behind gap formation

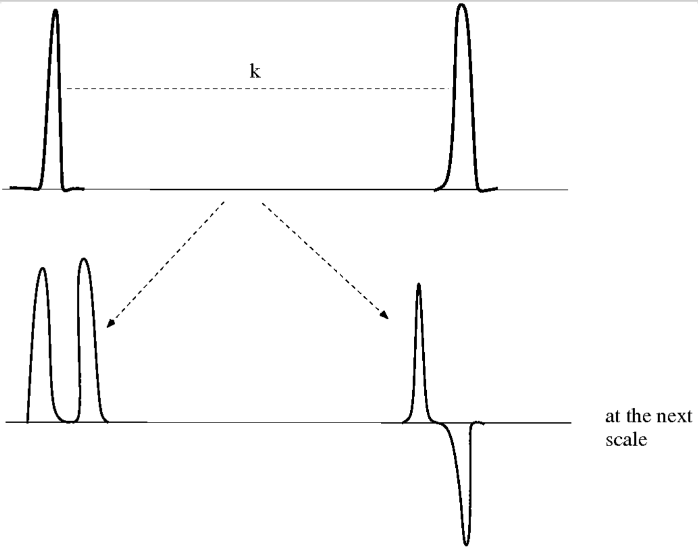

Figure: Crossing of graphs of eigenvalues create two peaks

Summary of Lecture 3

- Cartan estimate provides an exponential estimate on the measure of low-lying level sets in terms of Riesz mass. Does not distinguish between different measures of the same mass.
- BMO provides an alternative to Cartan, more sensitive to structure of the measure.
- One possible advantage of Cartan is that it yields a complexity estimate as well, at least for polynomials but only for the variable relative to which we measure deviations. The latter restriction is too severe.
- We use estimates on semi-algebraic sets instead, which allow us to make complexity statements relative to all variables.
- Matrix-valued Cartan estimate is designed for higher dimensions, resolvent identity. It is a fairly robust tool, also somewhat crude. Does not use the structure of the Riesz mass, only the size – the latter is given by the number of bad sites at the smaller scale.

References

See Lecture 1