

April Log

May 3, 2021

1 4/1

1.1 Goals

1. Review Complex - Crush exam

1.2 To Do

1. Schedule thesis
2. Submit MCS honors form
3. Commit to Georgia Tech - Done

1.3 Complex Review

Theorem 1. *Let \hat{f} be such that*

$$|\hat{f}(\xi)| \leq Ae^{-2\pi a|\xi|}$$

for some constants $A, a > 0$.

Then f is the restriction to \mathbb{R} of a function holomorphic in the strip $S_b = \{z \in \mathbb{C} : |\operatorname{Im}(z)| < b\}$ for $0 < b < a$.

Proof. We have the limiting approximation

$$f_n(z) = \int_{-n}^n \hat{f}(\xi) e^{2\pi i \xi z} d\xi$$

Can be extended to a definition in the strip S_b by our assumption on boundedness of \hat{f} . Entireness comes from fact we are integrating a holomorphic integrand which is a limit of holomorphic functions(proved via Morera).

We have uniform convergence to f and thus f is holomorphic (as uniform limit of holomorphic functions). \square

Remark 1. Note this implies f and \hat{f} cannot both have compact support unless $f = 0$. Since if f vanishes in an open interval then since it's holomorphic it must be 0 everywhere.

Theorem 2. *Let f be continuous with moderate decrease on \mathbb{R} . Then f has an extension to the complex plane that is entire with $|f(z)| \leq Ae^{2\pi M|z|}$ for some $A > 0$ if and only if \hat{f} is supported in the interval $[-M, M]$.*

Proof. Suppose \hat{f} is supported in $[-M, M]$. Then

$$f(x) = \int_{-M}^M \hat{f}(\xi) e^{2\pi i x \xi} d\xi \leq \|\hat{f}\|_{\infty} \int_{-M}^M |e^{2\pi i x \xi}| d\xi \leq \|\hat{f}\|_{\infty} 2M e^{2\pi M|x|}$$

where we bound $|e^{2\pi i x \xi}|$ with $e^{2\pi M|x|}$. Further f is entire since it is an integral of holomorphic integrand.

Now suppose $|f(x)| \leq Ae^{2\pi M|x|}$ and entire. Two observations:

1. The above argument gives us the stronger bound of $|f(z)| \leq e^{2\pi M|y|}$ instead of in terms of $|z|$
2. But this is not enough as we need some control over decay at ∞

Step 1. Use the stronger bound $|f(z)| \leq \frac{e^{2\pi M|y|}}{1+x^2}$ to show $\hat{f}(z) = 0$ when $|z| > M$

Step 2. Use slightly stronger bound $|f(z)| \leq e^{2\pi M|y|}$. \square

Remark 2. Note this means \hat{f} cannot be extended when f is exponentially bounded and entire since otherwise it is 0.

Theme 1. A key idea here is to shift the fourier transform by some imaginary offset iy and note this should still agree with real case via contour argument, as sides go to ∞ .

List of important tools:

- Phragmen-Lindelof: Allows us to uniformly bound holomorphic functions over unbounded sections assuming uniform boundedness on the boundary and exponential boundedness inside

Proof Chains

- Maximum-modulus \implies Phragmen Lindelof \implies Paley-Wiener

Remark 3. Often limiting arguments are examples of a reduction of more structure to less, like in the case of Phragmen Lindelof where we use $e^{-\epsilon\pi|z|^2}$ as a subduing function.

Let the limiting structure be the reducer, and the target structure be the reduced. Then the form of the reducer is guided by the reduced. In this case we need to recover the identity in the limit, and we need some sufficient decay.

Problems:

Prop 1.

$$\sum_{n=-\infty}^{\infty} \frac{1}{(\tau+n)^k} = \frac{(2\pi i)^{k-1}}{(k-1)!} \sum_{m=1}^{\infty} m^{k-1} e^{2\pi i m \tau}$$

Proof. First inclination was to take $f(x) = \frac{1}{(\tau+x)^k}$ but it works better if we take $f(z) = z^{k-1} e^{2\pi i \tau z}$

<https://math.stackexchange.com/questions/2927148/the-fourier-transform-of-fx-xiy-k-for-y0-k2>

□

Remark 4. Difficulties: Picking right function to take fourier transform of. Didn't know how to evaluate fourier transform.

Tricks: Picking right function, evaluating FT using integration by parts

Prop 2. Show if $\text{Im}(\tau) > 0$

$$\frac{(2\pi i)}{1!} \sum_{m=1}^{\infty} m e^{2\pi i m \tau} = \frac{\pi^2}{\sin(\pi \tau)^2}$$

Proof. Nothing too tricky. Must evaluate the sum via exponential series which can be done since $e^{2\pi i m \tau} \rightarrow 0$ as $m \rightarrow \infty$ since $\text{Im}(\tau) > 0$. If the sum is f is the primitive F . We can evaluate F then differentiate. Differentiation can be justified via DCT and the exponential convergence. □

Remark 5. Difficulties: none

Tricks: Using exponential power series closed formula and integrating to simplify sum.

Goal: Do all SS problems in chapter 4

1.

Proof. Key point is that $\int_{-\infty}^t f(x) dx = 0$ for every t hence must be 0. □

Question 1. When can a functions fourier transform be extended to be entire?

Answer 1. Sufficiently when we have sufficient decrease to bound the integral and apply DCT (like with gaussian).

Necessary?

2 4/2

2.1 Goals

1. Finish DRL
2. Start and do as much of evolution as possible
3. Work on thesis
4. Practice Jazz
5. Clean list

2.2 To Do

1. Schedule thesis
2. Submit MCS honors form

2.3 Exam Review

1.

Prop 3. Fix $z_1, \dots, z_n \in \mathbb{S}^1$. Wanted to show we can find z on unit circle such that

$$f(z) = \prod_{i=1}^n |z - z_i| = 1$$

Proof. I tried an elementary method arguing the roots of unity are the worst case.

Slicker to notice $f(0) = 1$ (I did not notice this) and then realize it must attain its maximum on boundary which is strictly greater than 1. And done by continuity since $f(z_j) = 0$ (I did realize this). \square

Prop 4. If f entire with $|f(z)| \leq e^{-ax^2+by^2}$ we want to show

$$|\hat{f}(\xi)| \leq e^{-a'\xi^2} \quad \xi \in \mathbb{R}$$

Proof. To show this should have contour integrated along $x - iy$ off the real line but I didn't do this (thought of it right at end). Instead tried some weak fourier inversion stuff and that's it. \square

2.4 DRL

OH:

Remark 6. Q-learning is basically q-value iteration but over an infinite input space. So instead of updating everything in an order we do so randomly based on the actions we have selected. This is why we have target and new network

Normal update is done in passes, where pass k computes k th values and we update by taking max over reward plus discount plus max over next actions of next state.

Tricks: use bitmasks for dones to avoid for loops

3 4/3

3.1 Goals

1. Start and do as much of evolution as possible
2. Work on thesis
3. Practice Jazz
4. Clean list

3.2 To Do

1. Schedule thesis
2. Look for internships

4/5

3.3 Goals

1. Grind thesis
2. Finish DRL
3. Complex homework
4. Clean list

3.4 To Do

1. Submit MCS form

3.5 DRL

Policy Gradient vs DQN:

One does gradient descent on an objective and the other an iterative value update.

<https://flyingfelix.github.io/2017/10/12/dqn-vs-pg.html>

Policy gradients seek to directly optimize over policy space (instead of learning q-values and constructing the policy via a max). The drawback of policy gradient is the high variance of estimating gradients.

Question 2. Why do policy gradient methods have high variance?

Answer 2. Just because gradient can look wildly different depending on directions of trajectories.

<https://www.quora.com/Why-does-the-policy-gradient-method-have-a-high-variance>

3.6 Complex Analysis

4 4/6

4.1 Modeling Evolution

Zhang 2020: Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China

5 4/7

5.1 Goals

1. Finish complex
2. Review DRL
3. Grind thesis

6 4/10

6.1 Goals

1. Grind Thesis
2. Start DRL
3. Work on modeling evolution
4. Start complex

6.2 Modeling Evolution Project

See project notes

6.3 Complex Analysis

Use of the Scharz Lemma: establishes all functions with $g(0) = 0$ and automorphism of \mathbb{D} as rotations in proof of characterization of automorphisms on \mathbb{D}

7 4/11

7.1 Goals

1. Finish Thesis

8 4/13

8.1 Goals

1. Finish Thesis
2. Finish complex
3. Email people
4. Start DRL
5. Propel homework

8.2 To Do

1. Netbox interview thing
2. Email tkocz and other masters people(gilchrist, ago, odonnel, gu)

8.3 Complex Analysis

Green's theorem.

<https://tutorial.math.lamar.edu/classes/calciiii/GreensTheorem.aspx>

For any area A

$$A = \int_C x dy = - \int_C y dx = \frac{1}{2} \int_C x dy - y dx$$

if we suppose $Q_x - P_y = 1$.

$f : U \rightarrow \mathbb{C}$ holomorphic is *Univalent* or *Schlicht* if it is injective. Univalent := holomorphic and injective

Discussion of Koebe 1/4 theorem and Gronwall area inequality.

<http://wwwf.imperial.ac.uk/~dcheragh/Teaching/2016-F-GCA/2016-F-GCA-Ch6.pdf>

$$\int_{f(A)} g(x) dx = \int_A g(f(x)) f'(x) dx +$$

9 4/14

9.1 Goals

1. Finish Thesis
2. Finish complex
3. start drl

9.2 To Do

1. Netbox interview thing
2. Email tkocz and other masters people(gilchrist, gu)

9.3 Relaxation

1. Set
2. Reading fantasy

10 4/15

10.1 Goals

1. Finish Thesis
2. start drl
3. Fix propel

10.2 To Do

1. Netbox interview thing
2. Email tkocz and other masters people(gilchrist, gu)

10.3 Relaxation

1. Set
2. Maia

11 4/16

11.1 Goals

1. Finish Thesis
2. start drl
3. Fix propel

11.2 To Do

1. Netbox interview thing
2. Email tkocz and other masters people(gilchrist, gu)

11.3 Relaxation

1. Set
2. Maia

12 4/17

12.1 Goals

1. Finish Thesis
2. start drl
3. study complex
4. do engage essays

12.2 To Do

1. Netbox interview thing
2. Email tkocz and other masters people(gilchrist, gu)

12.3 Relaxation

1. Set
2. Read model theory

13 4/18

13.1 Goals

1. Finish Thesis
2. start drl
3. Study complex

13.2 To Do

1. Netbox interview thing
2. Email tkocz and other masters people(gilchrist, gu)

13.3 Relaxation

1. Set
2. to be determined

13.4 Complex Analysis Review

Simple connectedness is a necessary condition for Riemann mapping theorem since F holomorphic implies F continuous and continuous functions preserve simple connectedness.

13.4.1 Riemann Mapping Theorem

Theorem 3. *Let $\Omega \subseteq \mathbb{C}$ be open and simply connected. Then there exists bijective and holomorphic $\phi : \Omega \rightarrow \mathbb{D}$ i.e. they are conformally equivalent. Further for arbitrary $z_0 \in \Omega$ we can find unique ϕ_{z_0} such that $\phi_{z_0}(z_0) = 0$ with $|\phi'_{z_0}(z_0)| > 0$.*

Ingredients:

Lemma 1 (Montel's Theorem). \mathcal{F} unif. bdd. family of holo functions on Ω . Then

1. \mathcal{F} is equicontinuous on every compact subset K
2. \mathcal{F} is normal

So in effect families of holo functions that are uniformly bounded are automatically equicontinuous! This is what is needed to extract function from $\mathcal{F} = \{f : \Omega \rightarrow \mathbb{C} \text{ holo, injective, } f(z_0) = 0\}$ using regularity.

Proof. Pick $z, w \in K \subseteq \Omega$ close. $|z - w| < r$. Then using cauchy formula

$$f(z) - f(w) = \frac{1}{2\pi i} \int_{D_{2r}(w)} f(\xi) \left(\frac{1}{\xi - z} - \frac{1}{\xi - w} \right) d\xi$$

So since f uniformly bounded simply need to bound $\frac{1}{\xi - z} - \frac{1}{\xi - w}$. Further

$$\left| \frac{1}{\xi - z} - \frac{1}{\xi - w} \right| \leq \left| \frac{w - z}{(w - \xi)(z - \xi)} \right| \leq \frac{|w - z|}{2r^2}$$

which is enough to close our integral estimate.

Cauchy formula + uniform bound and some estimates on difference of fractions allow us to get equicontinuity

Then AA effectively gives us our second statement. We prove this for completeness. We extract the sequence g_m via a diagonalization argument. Choose $\{\omega_j\}$ dense in Ω and countable. Then we choose our functions to converge on this set for each point by taking subsequences of subsequences. Then we choose our sequence to be the diagonal of this matrix construction. It remains to argue uniform convergence. Note we pick using bolzano weierstrass and boundedness. Continuity + convergence on countable dense set controls whole function. Uniform continuity comes from picking covering set via compactness of size δ for ϵ and getting convergence among each of the circles and picking over the largest one(maxing).

Diagonal picking + total boundedness gives uniform convergence from equicontinuity □

Key new ingredient: Cauchy's integral formula.

Lemma 2. Uniform limits of holomorphic injective functions are injective or constant.

Proof. Suppose $g = \lim_n g_n$ not injective. Then $z \neq w, g(z) = g(w)$. Want to show constant.

MY ATTEMPT:

Derivative could be 0, or arbitrary difference could be 0.

Set $d_n(\xi) = g_n(\xi) - g_n(z)$. So $d_n(w) \rightarrow 0$ uniformly. AFSOC not, then this is an isolated 0.

My END:

$$\frac{1}{2\pi i} \int_{\partial D_\epsilon(w)} \frac{d'(\xi)}{d(\xi)} d\xi = 1 \text{ WHY??}$$

But $\int_{\partial D_\epsilon(w)} \frac{d'_n(\xi)}{d_n(\xi)} d\xi = 0$ because holomorphic. Uniform convergence should give convergence of integral but insufficient. □

Key new ingredient integration of holomorphic functions.

Proof of . Uniqueness: straightforward. If two such maps do exist, F and G , for z_0 we let $H(z) = \frac{F(z)}{H(z)}$ and note $H(z_0) = 1$. Note then the singularity is removable and otherwise H holomorphic and is bounded. Hence it is constant and 1.

ALTERNATIVELY note $H = F \circ G^{-1}$ is an automorphism of \mathbb{D} . We know $H' = F'(G^{-1}) \circ G^{-1'}$. Also $H(0) = 0$ implying its a scaling i.e. $H(z) = cz$ for some $c = H'(0) = 1$.

So uniqueness follows by considering the composition which must be an automorphism of \mathbb{D} fixing 0. Then since it must be a scaling we simply compute the derivative to see the scaling is actually identity.

Existence: $\mathcal{F} = \{f : \Omega \rightarrow \mathbb{C} \text{ holomorphic, injective, } f(z_0) = 0\}$ is compact.

Step 1: Reduce to Ω to $0 \in \Omega' \subseteq \mathbb{D}$.

Step 2: Extract f with maximal $|f'(0)|$ and via uniform convergence show in \mathcal{F} .

Step 3: Show this f onto. We do this by tweaking f so that $f'(0) > s$ the supremal $f'(0)$ if there exists α it does not hit in \mathbb{D} . \square

Note: square root defined via log.

13.5 Chapter 8 Problems

Resources:

<https://math.berkeley.edu/~murphy/185-Solutions7.pdf>

Problem 1: Crux is using Rouché to argue for sufficiently close w that we at least two zeroes via lowest degree term of F power series being order two or more.

14 4/19

14.1 Goals

1. Work DRL 2. Do Propel 3. Do engage

14.2 Relaxing

1. Read PDE and data 2. Fantasy

14.3 Jazz

For 16 bar solo of same chord count 4 bars at a time

Good way to hold form together is to draw on melody (particularly when needing help).

Idea: abcd descending and ascending chord to a and back

Listen to kind of blue album

15 4/20

15.1 Goals

1. DRL
2. Engage essays
3. Complex

15.2 Relaxation

1. PDEandML
2. Set

15.3 DRL

Advantage of simulated dynamics model is that it is much less expensive/costly to simulate runs in head than in environment.

Optimal control:

<https://ai.stackexchange.com/questions/11375/what-is-the-difference-between-reinforcement-learning-and-optimal-control>

16 4/21

16.1 Goals

1. Grind DRL
2. Work complex
3. Clear head

16.2 To Do

1. Email tkocz - done
2. Do netbox thing

16.3 Relaxation

1. Set
2. Read PDE and Data

16.4 DRL Homework

17 4/22

17.1 Goals

1. Grind DRL
2. Work complex
3. Model evolution

17.2 Tkocz notes

1. Justify expectatio integral formula in Haagerup.
2. Make sure to get all full-stops

17.3 DRL Notes

Idea behind MPC is to predict T steps in advance using optimizer. Then at the next step we iteratively update plan after stepping through environment. Plan formation and adaptation.

18 4/24

18.1 DRL

Homework 4:

Questions:

1. How is MPC used in Random policy selection

19 4/25

19.1 Goals

1. Finish DRL
2. Make reserach presentation
3. Modeling evolution

19.2 To Do

1. Netbox stuff
2. Email Gilchrist
3. Email to set up thesis

19.3 Relaxation

1. Maia
2. QFT stuff
3. Read war and peace

19.4 DRL

Apparently eval shorthand for sess.run as default

20 4/26

20.1 Goals

1. Finish DRL
2. Make reserach presentation
3. Modeling evolution
4. Jazz

20.2 To Do

1. Netbox stuff

2. Email Gilchrist
3. Email to set up thesis
4. Email exam

20.3 Relaxation

1. Maia
2. QFT stuff
3. Read Perdido street station

20.4 DRL

How to print in tensorflow computation? For example how to print losses?

Use pdb to debug tensorflow values

20.5 Jazz

Want to use roots as much as possible. Use voice leading between top note and melody note.

Should play more open, for example move C to right hand. Thirds and sevenths are in right hand, not left hand. Sometimes sevenths can be doubled. This is particularly true for solo piano playing.

Samba is heavy Brazilian rhythm, faster version of slow Tangerine. Most well written songs can be written in a variety of rhythmic settings. Good way to practice is to mess these up.

Sometimes sharped nine vs flat nine can be interchanged via good voicing. recall good voicing means minimal movement

21 4/27

21.1 Goals

1. need to write scientific analysis for propel
2. Finish engage
3. Finish drl
4. Make research presentation
5. Work on evolution

21.2 To Do

1. Netbox
2. Email people about thesis

21.3 Relaxation

1. Set
2. Maia

22 4/28

22.1 Goals

1. need to write scientific analysis for propel
2. Finish engage
3. Finish drl
4. Make research presentation
5. Work on evolution

22.2 To Do

1. Netbox
2. Email people about thesis

22.3 Relaxation

1. Set
2. Read perdido station

22.4 Thesis Presentation

Basic Research:

<https://blogs.ams.org/mathmentoringnetwork/2014/08/04/math-talk-preparing-your-conference-presentation/>

<https://control.com/technical-articles/what-is-model-predictive-control-mpc/>

<https://www.quora.com/What-are-the-pros-and-cons-of-model-based-and-model-free-reinforcement-learning>

22.5 Thesis Presentation Tex Bits

<https://www.codecogs.com/latex/eqneditor.php>

<https://math.asu.edu/resources/computer-resources/how-make-slides-latex>

$$(\mathbb{E}|\sum_{i=1}^n a_i \epsilon_i|^p)^{1/p} \leq C_{p,q} (\mathbb{E}|\sum_{i=1}^n a_i \epsilon_i|^2)^{1/2}$$

Thesis Remarks:

In general consider taking out some text.

Slide 2: Trivially true with $C_{p,q} = 1$ when $p < q$ via monotonicity of moments. The khintchine inequalities show equivalence of moments. Emphasize $C_{p,q}$ denotes the best possible constant.

He liked that I say khintchine proved they exist. Example where these do not exist? When the moment is not finite kinda answers this.

Slide 3: Write down $C_{p,q}$ is supremum over a, n for ratio of norms. Frame as optimization problem. Just say type and cotype are invariants of banach spaces, and averaging over random signs is a way to construt them. INclude stuff on Grothendieck and other thing as bullet list

Khintchine proved these to prove law of iteratd logarithm. This was his motivation(in 1923). Rediscovered by Hardy and Littlewood. Hardy first proved upper bound by $\sqrt{n \log(n)}$ refined by Khintchine(he knew about Hardy) by $\sqrt{n \log(\log(n))}$. This tighter version later showed by Hardy(who did not know about Khintchine???)

Also projection not intersection

Slide 4: Instead of citing numbers puts names and dates

Slide 5: Just define usg for random variables, not random vectors

Slide 6: Have display somewhere saying constants are sharp because of central limit theorem. Say expectations converge via expection argument with CLT

Slide 8: Say moment generating function. Tkocz would let go L' class. Write why random sings are type L

Slide 9: Bullet these. Or write as remarks/observations. Good to structure thigs.

Slide 11: Get rid of type L'. Run argument in finite case and then just pass to limit. Write out explicitly what the elementary symmetric functions are. Explicitly say we are comparing terms. Just most importantly define elementary symmetric functions

Slide 13: Write proof of lemma. Just say atom from product is a characteristic function. Just by taking IFT.

Slide 14: Make point this is khintchine inequality for $C_{p,2}$ since this is the variance. Fix variance. Try to minimize and maximize the p th moment. Looking for extremals under this constraint. Variance is the second moment

Slide 16: Move before slide 15 and write example of sequence majorization. Fix typo. Also to mention proof key is to modify two coefficients at a time and use independence and show nonnegativity in a difference this way.

Slide 19: Bullet

Slide 20: Instead of writing out explicitly just write in terms of Gaussian

Slide 20: Proof relies on Gaussian swapping argument and using independence

Slide 21: Proof relies on convexity argument

Slide 22: Haagerup argument. Throughout formula for c_1 . Equality in $n = 1$ case. Would be nice to display fourier analytic integral representation

Slide 23: No need to display references

Note: No need to write our names on new theorems. Slides 10 and slide 14

22.6 DRL HW 4 MPC

How is it possible random actions with ground truth and no MPC do better than MPC with a simulated dynamic? Does random even use MPC? Is CEM with learned dynamic MPC bad because learning bad? Maybe because the policy is MPC and random moves are done to generate episodes(instead of CEM fitness)?

22.7 Modeling Evolution Notes

23 5/1

23.1 Goals

- Modeling evolution grind
- Finish presentation

24 5/2**24.1 Thesis Notes**

Slide Point Targets

1. Say what I'm presenting
2. Go over summary and make clear these are new results
3. Mention triviality when $p < q$ and via monotonicity of moments. State existence which shows all moments equivalent.
4. Averaging over random signs way to construct type and cotype
5. Makes sure to mention most comparisons only known to second norm.
6. NA
7. Mention used in polynomials to collect terms
8. NA
9. Make sure to emphasize sharp constants due to CLT
10. NA
11. Walk through why random signs type L
12. Mention useful to then show Khintchine
13. Emphasize Hadamard factorization
14. Emphasize sharpness of constants via CLT
15. NA
16. NA
17. NA
18. Emphasize still khintchine because variance fixed and optimizing p th moment.
19. Emphasize mass intuition (spreading v.s. congregating). Gaussian is spread out case.
20. NA
21. Connect back to mass metaphor

22. NA

23. NA

24. NA

25. NA

26. NA

should say second moment easy to compute

remove rotation invariance

Remove reference slide 7

remove type L'

a_j

Fix slide 21 grammar.

Should say I'm going to present new results on type L random variables and Symmetric discrete uniform random variables. I say this in the outline

Highlight random signs definition

Use box for definitions instead of highlighting

Metaphor for spreading and congregating mass useful

Sydney wants more pictures