Qualifying Exam Syllabus

Hong Suh

April 9th, 1-4pm, 762 Evans

Committee: Fraydoun Rezakhanlou (Advisor), Lawrence Craig Evans (Chair), Alan Hammond, Venkatachalam Anantharam.

1 Major topic: Probability

References: Rick Durrett, Probability: Theory and Examples, fourth edition.

- *Law of large numbers (Chapter 2).* Independence, weak law of large numbers, Borel-Cantelli lemma, strong law of large numbers, Kolmogorov's 0-1 law, large deviations
- *Central limit theorems (Chapter 3).* Weak convergence, characteristic functions, central limit theorems
- *Martingales (Chapter 5).* Conditional expectation, martingales, upcrossing lemma, martingale convergence theorem, Doob's inequality, uniform integrability, backwards martingales, Hewitt-Savage 0-1 law, optional stopping theorems
- *Markov chains (Chapter 6).* Recurrence and transience, stationary measures, asymptotic behavior
- *Brownian motion (Chapter 8).* Gaussian process, Brownian motion, Blumenthal's 0-1 law, strong Markov property, Ito's formula

2 Major topic: Partial Differential Equations

References: Lawrence C. Evans, Partial Differential Equations, second edition.

- *Four important linear PDE (Chapter 2).* Transport equation, Laplace's equation, Heat equation
- *Nonlinear first-order PDE (Chapter 3).* Characteristics, Hamilton-Jacobi equations, scalar conservation laws
- *Sobolev spaces (Chapter 5).* Approximation by smooth functions, extension operator, trace operator, Gagliardo-Nirenberg-Sobolev inequality, Morrey's inequality, Rellich-Kondrachov theorem, Poincaré's inequality
- *Second-order elliptic equations (Chapter 6).* Lax-Milgram theorem, Fredholm alternative, elliptic regularity, maximum principles, eigenvalues and eigenfunctions
- *Linear evolution equations (Chapter 7).* Second-order parabolic equations, semigroup theory

- *Calculus of variations (Chapter 8).* Euler-Lagrange equations, regularity theory, constraints, mountain pass theorem
- *Hamilton-Jacobi equations (Chapter 10).* Viscosity solutions, existence and uniqueness, control theory

3 Minor topic: Dynamical Systems

References: Fraydoun Rezakhanlou, Lectures on Dynamical Systems (part II).

- *Ergodicity.* invariant measures, von Neumann ergodic theorem, Birkhoff ergodic theorem, Kingman's subadditive ergodic theorem, mixing
- *Entropy.* Topological entropy, metric entropy, Shannon-McMillan-Breiman theorem
- *Lyapunov exponents.* Oseledets theorem, Ruelle's inequality, Pesin's entropy formula

4 Qualifying Exam Transcript

My committee was Craig Evans (CE), Alan Hammond (AH), Fraydoun Rezakhanlou (FR), and Venkatachalam Anantharam (VA). It went roughly like this:

CE: Feel free to choose whatever topic you want to start with.

HS: I'll choose PDE.

CE: Why does everyone choose PDE first? Ha. State Lax-Milgram.

I state it

CE: Ok, now apply Lax-Milgram to the PDE $-\triangle u = f$ on U, where $f \in L^2$, and $u = 0$ on ∂U.

HS: For this, the bilinear form is symmetric and in fact you don't need Lax-Milgram to get existence and uniqueness here.

CE: Just use it anyway – write down the bilinear form.

HS: It goes like $B[u,v] = \int Du \cdot Dv \, dx$. The equation we want to solve is $B[u,v] = (f,v)$.

CE: Ok, let's skip the bound $|B[u,v]| \leq \alpha \|u\|_{H^1} \|v\|_{H^1}$. How do we get the lower bound $B[u,u] \geq \beta \|u\|_{H^1}^2$?

HS: We have the bound $c\|u\|_{L^2}^2 \leq \|Du\|_{L^2}^2$ since $u \in H_0^1$, by Poincare's inequality. This gives the lower bound.

CE: State Poincare's inequality.

HS: *I start to state it, and become nervous about* $p = \infty$ Can I do it for $p = 2$? *I get an affirmative answer from CE.* It is precisely what I stated above: if $u \in H_0^1(U)$, then $\|u\|_{L^2} \leq C\|Du\|_{L^2}$.

CE: Now prove it.

HS: We can use GNS to prove it. If $u \in C_c^1(\mathbb{R}^n)$... (*CE stops me*)

CE: This will work. But I want you to do it by contradiction.

HS: Ok. Suppose not. Then there exists a sequence $u_k \in H_0^1(U)$ with $\|u_k\|_{L^2} \geq k\|Du_k\|_{L^2}$. We may rescale the u_k so that it has L^2 norm one. Therefore $\|Du_k\|_{L^2} \leq 1/k$. Since $2^* < 2$, ... (*CE corrects me, "no, it's the other way'*) oh, $2 < 2^*$, so we may apply Rellich compactness to get that since $\|u_k\|_{L^2} = 1$ and $\|Du_k\|_{L^2} \leq 1/k$, their H^1 norm is bounded so we have a convergent subsequence u_{k_j} which converges in L^2 to u. Since $\|Du_k\|_{L^2}$ are bounded, they must converge weakly so $Du_k \rightharpoonup Du$. Since $\|Du_k\|_{L^2} \leq 1/k$, we have $\|Du\|_{L^2} = 0$ and so u is constant a.e. in U. Since $\|u_k\|_{L^2} = 1$, we have $\|u\|_{L^2} = 1$. (*I fumble a bit here*) I know I get a contradiction with the boundary condition, which is $Tu = 0$...

CE: What does T mean here?

HS: It is the trace operator.

CE: What do we know about it?

HS: We know it is a bounded linear operator from $L^2(U)$ to $L^2(\partial U)$ with $\|Tu\|_{L^2(\partial U)} \leq C\|u\|_{L^2(U)}$...

CE: No, that's wrong. It is not L^2 on the right hand side.

HS: Ah, it is H^1.

CE: Yes.

HS: But this doesn't quite help with what we want because the inequality goes the wrong way...

CE: Yes. What is the other thing we know about the trace operator?

HS: We know that if $u \in C^1$, then $Tu = u$ on ∂U. *fumble* Actually just $u \in C$. *CE agrees*

CE: Yes... use this fact.

HS: Well, these are not continuous...

CE: It is continuous! u is just constant. Almost everywhere of course.

HS: Ah, then if u is continuous and constant, and $Tu = u$ on ∂U, then $u = 0$ on ∂U so u must be constantly zero.

CE: Yes, and we get a contradiction. Ok, let's move on. You have Hille-Yosida on here. State the Hille-Yosida theorem.

HS: Hille-Yosida gives necessary and sufficient conditions for a linear operator on Banach space to be the generator of a semigroup. (*Sometimes I state the general principle in the hopes that they let me off the hook*) The theorem goes like this: let $A : X \to X$ be a closed linear operator, and suppose A is densely defined.

CE: So the A doesn't go from X to X, it has a smaller domain.

HS: Ah yes (*I change X to $D(A)$*) then $(0, \infty) \subseteq \rho(A)$ and $\|R_\lambda\| \leq 1/\lambda$, if and only if A generates a contractive semigroup.

CE: What is R_λ?

HS: It is $(\lambda I - A)^{-1}$, and $\rho(A)$ is where R_λ makes sense.

CE: Ok that is fine. Suppose we have a Markov process with compact state space, say X. Is there an associated semigroup?

HS: If we have $E_x g(M(t))$ where M is the Markov chain, then $u(x,t) = E_x g(M(t))$ solves some parabolic PDE...

CE: Write $u(x,t)$ in terms of semigroups.

HS: We can write $u(x,t) = S(t)u$, where u is the initial condition.

CE: You have g on the left over there, and u on the right.

HS: Ah yes it is $u(x,t) = S(t)g$.

FR: You are missing x.

HS: Yes, $u(x,t) = (S(t)g)(x)$.

CE: Ok, that's fine. Consider the nonlinear PDE $-\triangle u = f(x,u)$ in U and $u = 0$ in ∂U. Is this variational?

HS: I know a similar simple version is variational, without the x, so my guess is that this is variational.

CE: We can switch to the simpler one for now. Write it down. (*I write it*) What is the Lagrangian for this?

HS: It is $Lu = \ldots$ (*I realize this makes no sense*) I mean $L(p,z,x) = \frac{1}{2}|p|^2 - F(z)$, where $F' = f$.

CE: Yes that looks right, either plus or minus... (*I think it is correct the way I have it*) Now what about for the original problem?

HS: I suppose we can just add x, to get $L(p,z,x) = \frac{1}{2}|p|^2 - F(x,z)$. Where $F' = f$ in the z variable.

CE: Write down what F is explicitly.

HS: It would be $F(x,z) = \int_0^z f(x,w)\, dw$.

CE: That works. Can you state what a viscosity solution to, say, a Hamilton-Jacobi PDE is? (*I state it*) Tell me how it is related to the following idea: Say we have an ODE $\dot{x}(s) = f(x(s), \alpha(s))$ and $x(t) = x$, ... *He states the whole setup for control theory in Evans ch 10* What can we say about $u(x,t) = \min_{\alpha \in \mathcal{A}} C_{x,t}[\alpha]$?

HS: It satisfies the optimality conditions, or dynamic programming principle, $u(x,t) = \inf_{\alpha \in \mathcal{A}} \{ \int_t^{t+h} r(x(s), \alpha(s))\, ds + u(x(t+h), t+h) \}$.

CE: Can you derive a PDE from here?

HS: I think so. Consider $\frac{u(x,t+h) - u(x,t)}{h}$, and we write $u(x,t)$ with the dynamic programming principle as above...

CE: This is fine, but let's take the derivative along the curve. How about you take $u(x,t)$ from the left hand side and put it on the right, inside the inf?

HS: Ok, then we get $0 = \inf \int_t^{t+h} r(x(s), \alpha(s))\, ds + u(x(t+h), t+h) - u(x,t)$ so the last two terms look like $(\frac{d}{dt} u(x(t), t))h$...

CE: Up to an error term. Why don't you divide by h on both sides?

HS: Ok, we get $0 = \inf \frac{1}{h} \int_t^{t+h} r(x(s), \alpha(s))\, ds + \frac{d}{dt} u(x(t), t)$ and the last term we differentiate and get $Du(x(t), t) \cdot \dot{x}(t) + u_t(x(t), t)$ and since $x(t) = x$ we get $Du(x,t) \cdot \dot{x}(t) + u_t(x,t)$. The first term we just get $r(x(t), \alpha(t))$.

CE: Up to an error term of course.

HS: Yes, so heuristically this works. We can also replace $\dot{x}(t) = f(x, \alpha(t))$. Then we have a PDE.

CE: Yes, and the $\alpha(t)$ we can switch to $a \in A$ since it is the value of α at fixed time. Ok. This is assuming everything is smooth. How do we make this rigorous? (*I look confused, I am panicking as I think he is asking me to prove that this is actually a viscosity solution*) In what sense does this solve the PDE? (*I am relieved*)

HS: Ah, it is a viscosity solution to the PDE.

CE: Ok good. You're fine on PDE. Should we move on? (*I wish I got a break but I got no break here*)

AH: (*Explains the problem, I don't really get it... this goes on for about 3-5 minutes*) Let $x \in [0, 1]$ be a real number and write x in terms of its decimal expansion. (*I already know I'm about to do something I've never done*) Consider the event on which the following happens infinitely often in n: there are equal numbers of each digit in the digits up to and including the nth digit. Is this event Lebesgue measurable, and if so, what is the Lebesgue measure of this event?

HS: Ok... so something like this? *I write some gibberish*

AH: Yes, try to write it precisely. Write out the decimal expansion.

HS: Write $x = .x_1x_2\ldots$. We have $A = \{x \in [0,1] : \#\{i \le n : x_i = j\} = \#\{i \le n : x_i = k\},\ \text{for all } j,k = 0,\ldots,9,\ \text{ infinitely often in } n\}$.

AH: Write it in a more useful way.

HS: We can write it like $A = \cap_{N=1}^{\infty} \cup_{n\ge N} \{\#\{i \le n : x_i = j\} = \#\{i \le n : x_i = k\},\ \text{for all } j,k = 0,\ldots,9\}$. The sets inside, write A_n, are basically unions and intersections of half-open intervals of length 10^{-n}, so this should be measurable.

AH: Okaaaay. (*Sounds dubious, but perhaps he just sounds like this.*) Now for the measure of the set. There is a simpler way to write the last thing.

HS: Ah I don't know.

AH: Firstly, this cannot happen if n is not a multiple of 10. (*I correct it*) Each digit appears equally often. Then what can we say about the number of times each digit appears?

I fumble for an uncomfortable amount of time, detect frustration in AH's voice, start vomiting internally

AH: *AH basically repeats what he has been saying*

HS: Ah! This means each digit appears exactly n times. (*I write that more formally somehow*)

AH: Yes. Now what can we say about the measure of this? What do you think that it would be? Polynomial in n, exponential in n, etc?

HS: I think it should be exponential (*WRONG*)

AH: Compute it for a simpler version, say binary expansion.

HS: (*I think for a while*) Then each digit appearing exactly n times at time $2n$ is the event that we have n successes among $2n$ trials. So this is a $\text{bin}(2n, 1/2)$. If $X_{2n} \sim \text{bin}(2n, 1/2)$, the measure $m(A_n)$ of A_n is given by $P(X_{2n} = n) = \binom{2n}{n}2^{-2n}$.

AH: What do you guess that this number is?

HS: I would guess exponential? (*WRONG. I was stuck on the exponential for some reason*)

AH: Try computing it.

HS: We can use Stirling's formula to compute it, but I can't remember it.

AH: Yes we could do that, but try a different way. What is the mean and variance of this X_{2n}? You already mentioned that its mean is n.

HS: We should be able to do a CLT argument... *I don't really know how to proceed*

AH: Try computing $P(X_{2n} = n + \sqrt{n})$.

HS: Ah... CLT regime... but it's discrete... ah... *my head is buzzing like a bee in heat and I cannot think*

AH: Just think about the pdf of Gaussian. (*Or something like that*)

HS: Ok, it should be comparable to $P(X_{2n} = n)$.

AH: Yes, now what can you say?

HS: (*finally I kinda get it*) There are about $\sqrt{n}$ possible values between n and $n + \sqrt{n}$ and each of their probabilities are comparable. and $P(n \le X_{2n} \le n + \sqrt{n})$ is bounded below by a constant. So each one should be about $1/\sqrt{n}$ so $P(X_{2n} = n) \approx 1/\sqrt{n}$.

AH: Yes. Now what can we say about the decimal case?

HS: I would guess that it is the same, so $1/\sqrt{n}$.

AH: No... try the base 3 case.

HS: (*I ignore his idea which I shouldn't have done*) For base 4 we should get square of what we got before (*FALSE FALSE FALSE*) so probably for base 10 we have $(1/\sqrt{n})^5$.

Later Milind and I computed that it is $(1/\sqrt{n})^9$*, using Stirling's formula. I realized the next morning that we can heuristically think of it this way: for base 3, let* X_{3n} *be the number of 0s at time* $3n$*, and let* Y_{3n} *be the number of 1s. The number of 2s are determined by* X_{3n}, Y_{3n}*. And the probability we are looking at is* $P(X_{3n} = n,\ Y_{3n} = n)$*. But* X_{3n} *and* Y_{3n} *should be roughly independent in large* n*, because these events are both centered around the mean. Then the probabilities split and we get* $1/n$ *for base 3 case. For decimal case, we get a splitting of* 9 *probabilities, each around* $1/\sqrt{n}$*, so we get* $(1/\sqrt{n})^9$.

AH: No... (*He seems to give up on me*) it's not that, but I won't pursue this longer. Assuming it is that, what is the final answer?

HS: Since this is summable, by Borel Cantelli, the probability this happens infinitely often is zero. (*internally screaming because he gave up on me*)

AH: Yes. I think we should move on. Suppose we have a Markov chain. What are the conditions for it to converge to a stationary distribution? (*or something like that*)

HS: If we have p a transition probability and p is irreducible and aperiodic, then $p^n(x, y) \to \pi(y)$ for each $x \in S$, if π is a stationary distribution, if it exists.

AH: What is the definition of stationary distribution? (*I state it*) Sometimes one may not exist; what is a sufficient condition for one to exist?

HS: We are in the irreducible case – if for some state x, we have $E_x T_x < \infty$ then a stationary distribution exists.

AH: Any idea how to write it down?

HS: Yes. Pick this state x, and define stationary measure $\mu_x(y) = E\sum_{n=0}^{T_x-1} 1_{X_n=y}$. (*I realize now that I forgot to write x on the subscript, I should've written E_x*) Sum over y to get $\mu_x(S) = E\sum_{n=0}^{T_x-1}\sum_{y\in S} 1_{X_n=y} = E\sum_{n=0}^{T_x-1} = E_x T_x < \infty$. Divide by $E_x T_x$ to get the distribution.

AH: Yes. Let's move on to Brownian motion. Suppose we have on the continuous functions on $[0,1]$ the uniform norm, and consider an open set. Suppose it contains a function which starts 0 at time 0. Can you show that Brownian motion has a positive probability of staying in that set?

HS: I first consider an open ball, say $B_\delta(f) = \{g \in C[0,1] : \|f-g\|_\infty < \delta\}$ with $f(0) = 0$. This is basically a δ tube around the graph of the function f. (*I draw some graphs – it is clearer with graphs*)

FR: Are there any conditions on f?

AH: Just continuous should suffice.

HS: Now I draw some boxes here – I draw boxes to cover the graph of f while staying in the δ tube. So our probability that B_t is in $B_\delta(f)$ is bounded below the probability that B_t stays within these boxes, say $A_1, \ldots, A_n$.

AH: How do we know we have finitely many boxes?

HS: By compactness of $[0,1]$, we can find a finite number of boxes. (*AH is satisfied*) Now look at the first box, A_1. We want the probability that BM stays within this box with width $[0,t]$... so first look at $P(\sup_{s\le t} B_s \le b)$ which by reflection principle is equal to $(1/2)P(B_t \le b)$... No, reflection principle doesn't go that way... but we can write $P(\sup_{s\le t} B_s \le b) = 1 - P(\sup_{s\le t} B_s \ge b) = 1 - (1/2)P(B_t \ge b)$... wait, no, it is 2, not 1/2. (*committee mumbles in agreement*) so we have $P(\sup_{s\le t} B_s \le b) = 1 - 2P(B_t \ge b)$.

AH: Yes, now you need that $P(B_t \ge b)$ to be less than 1/2.

HS: Yes, we get that since $b > 0$, or else the box doesn't cover the graph of f.

AH: There are some subtle issues you may run into here, you need to take care of the inf and so on. (*FR comments on the future steps as well*)... does anyone else want to ask questions? (*FR signals that he wants to ask questions*)

CE: Let's take a break. (*people agree. THANK GOD*)

HS: I pace down the halls, visit Milind and tell him I had a hard time with AH's questions. I come back.

FR: State the definition of Lyapunov exponents and their significance.

HS: Oh I thought you were going to ask further probability questions!

FR: No it's ok, but perhaps I may come back to it later... (*He does not*)

HS: Lyapunov exponents measure how much orbits diverge over time, or converge. (*I should've mentioned direction but doesn't matter*) To define them we need to state Oseledets theorem. (*I state it*)

FR: Good. How does this connect to what you said, like how much orbits diverge? What $A(x)$ should we be thinking of?

HS: We should be thinking $A(x) = (dT)_x$, and $A_n(x) = (dT^n)_x$.

FR: Good. What happens in the case in which the ergodic measure here is on a periodic set?

HS: When it is on a fixed point, it is just linear algebra.

FR: Ok, let's try it on this simpler case. Suppose x is a fixed point and state Oseledets theorem there.

HS: In this case $A(x)$ has no dependence on x since it just goes from $T_xX \to T_xX$ and now the lyapunov exponents are simply given by the log of the eigenvalues of this matrix. The splitting is given by the eigenspaces. Should I do the whole thing with Jordan normal form?

FR: No need. Now try the periodic case.

HS: (*I say some high-level bullshit*)

FR: Write it out.

HS: For $A_n(x)$ write $A_n(x) = A_{km+r}(x) = A_{km}(T^r(x)) \circ A_r(x)$.

FR: It may be more helpful to write the whole thing out and group them in groups of m.

HS: We can write $A_{km+r}(x) = A(T^{km+r-1}(x)) \circ \cdots \circ A(T^{(k-1)m+r-1}(x)) \circ \cdots$ (*I get stopped by FR*)

FR: Write it the opposite way, starting from the beginning, not the end.

HS: Write $A_{km+r}(x) = \cdots \circ (A^{T^{m-1}(x)} \circ \cdots \circ A(T(x)) \circ A(x))^k$ (*I wrote from right to left here, and added the power of k at the end*) and the remainder goes in the front.

FR: Ok, that is fine. What is the connection between entropy and Lyapunov exponents?

HS: Ruelle's inequality states that if X is a smooth manifold, $T : X \to X$ is a smooth map, then $h_{top}(T) \leq \sum_{i=1}^{k} n_i \ell_i^+$ where n_i are the multiplicities of the Lyapunov exponents. Ah and $\mu \in \mathcal{I}_T^{er}$ since Lyapunov exponents are defined on μ.

FR: You have Shannon-McMillan-Breiman theorem on here. State the theorem.

HS: If $\mu \in \mathcal{I}_T^{er}$, then $-\frac{1}{n} \log \mu(C_{\xi_n}(x)) \to h_\mu(T, \xi)$ in L^1. I suppose I have to define these quantities. Here ξ is a finite μ-partition, which means it is a partition of a measure one set in X. Then $C_\xi(x)$ is the unique element of ξ containing x, which is defined μ-a.e.. Now define $I_\xi(x) = -\log \mu(C_\xi(x))$, and $H_\mu(\xi) = \int I_\xi(x)\mu(dx)$, and $h_\mu(T, \xi) = \lim n^{-1} H(\xi_n)$ where $\xi_n = \xi \vee T^{-1}\xi \vee \cdots \vee T^{-n+1}\xi$.

FR: Yes. Do you remember the proof?

HS: Yes, somewhat... we want to show a sort of ergodic theorem for $n^{-1} I_{\xi_n}(x)$ and write it as a kind of ergodic sum. We need to define conditional entropy for this (*I define it*) and we use the identity $I_{\xi \vee \eta}(x) = I_\xi(x) + I_{\eta|\xi}(x)$. I want to use this on $n^{-1} I_{\xi \vee T^{-1}\xi \vee \cdots \vee T^{-n+1}\xi}(x)$...

FR: It may be better to reverse ξ and η in your lemma.

HS: Ok, then $I_{\xi \vee \eta}(x) = I_\eta(x) + I_{\xi|\eta}(x)$. Now $I_{\xi_n}(x) = I_{\xi \vee T^{-1}\xi \vee \cdots}$ (*FR stops me*)

FR: We want to take the last $n-1$ terms together, not the first $n-1$ terms.

HS: Ah, then $I_{\xi_n}(x) = I_{T^{-1}\xi \vee \cdots \vee T^{-n+1}\xi}(x) + I_{\xi|\xi(1,n-1)}(x)$. Then by invariance of μ, the first term is $I_{\xi(0,n-2)}(T(x))$. Then we can iterate this...

FR: Good, I believe you can keep going. This is not on your syllabus, but sometimes you can get a large deviations principle related to the entropy of the system. (*or something like that*) Do you remember what it is?

HS: I remember you said something in class about this... in the variational statement $h_{top}(T) = \sup_{\mu \in \mathcal{I}_T^{er}} h_\mu(T)$, in nice cases we have a unique maximizer and that serves as the "mean" in the large deviations principle, as in it is the minimum for the rate function.

FR: Do you know an example for which we actually know this happens?

HS: The shift on m symbols.

FR: Good. I think that is enough.

They send me out of the room, I have a pretty good feeling but I got smashed on that first probability question which makes me feel queasy. I sit down on the ground still stressed out of my goddamn mind. CE brings me back in and says I seemed a little bit weaker in the probability section but I was fine, they congratulate me.