Status Report for the Princeton University Junior Research Seminar / Undergraduate Mathematics Laboratory

Steven J. Miller

February 18, 2005

Contents

1 Introduction

The purpose of the Junior Research Seminar / Undergraduate Mathematics Laboratory at Princeton is to form a research team of undergraduates, graduate students and faculty to investigate interesting unsolved conjectures theoretically and experimentally. The UML is sponsored by a VIGRE grant of the National Science Foundation.

In addition to the standard lecture-homework classes, we wanted a class where the undergraduates would work on hot conjectures and see what kinds of problems mathematicians study. In the sciences and engineering, undergraduates are often exposed to state of the art projects through experimental labs; we wanted to bring a similar experience to the math majors. We are competing with other departments (computer science, financial math) for majors; we feel interesting classes such as this help keep people, if not math majors, than at least math minors and people who appreciate and understand what we do.

The undergraduates often have enough theory to understand the basic framework and proofs of simple cases. Building on this, they then numerically test the conjectures. The undergraduates learn a good deal of theory, they learn about coding, simulations and optimization (very marketable skills), and they get to see what is out there. The graduate students and the faculty get a potent calculating force for numerical investigations. Many of the problems the students have investigated arise from graduate dissertations or current faculty research. It has been very easy finding graduate students and faculty excited about working on this course; at the end of a semester or year, instead of having a folder of solution keys to calculus, a graduate student should be able to co-author an experimental math paper with the undergrads.

This program is exportable; in fact, Professors Peter Sarnak and Steven J. Miller ran a similar course at NYU during Fall 2002. Building on the success of the Princeton program, Brian Conrey, David Farmer, Chris Hughes, Steven J. Miller and Michael Rubinstein ran a Vertically Integrated Summer Program in Computational Mathematics at AIM during Summer 2003; participants from Princeton included one Junior Faculty member and two undergraduates. Professor Miller then ran a similar program at The Ohio State University (both during the academic year 2003-2004 and during the summer).

The class requires at least modest computer support, as well as an enthusiastic faculty / graduate student team. The challenge is finding interesting problems which (1) the undergraduates know enough theory to understand and (2) are amenable to the undergraduates' coding abilities.

For (1), for three years we arranged a parallel lecture series to the class to give a quick introduction to the needed material. For (2), we try to have an extensive list of problems of varying coding difficulty. This is not a course in programming, but a course for mathematicians using computers as a tool. Fortunately, there are numerous problems in number theory and graph theory where the mathematical prerequisites are light and the coding is straightforward.

A class like this is significantly more work than a normal class for the faculty and graduate students, but it is also a fun class to teach, and the benefits are worth the extra time. Further, it is an outstanding opportunity prepare graduate students for advising students and running research groups.

Below is a shortened version of a description of the classes. For more information and to view the undergraduates' reports, please go to

http://www.math.princeton.edu/mathlab/index.html

Some of the results from the Random Matrix Theory and Ramanujan Graph problems were presented by Steven J. Miller at the Computational Number Theory Workshop at the Foundations of Computational Mathematics Conference in Minneapolis (August, 2002). For the slides of the talk, go to

http://www.math.princeton.edu/mathlab/finalreport/minn.pdf

For an expanded version of the lecture, go to

http://www.math.princeton.edu/mathlab/finalreport/cnt_sjm.pdf

Steven J. Miller also presented a report on the Undergraduate Mathematics Lab at a National Science Foundation conference on computation (Washington, D.C., September 2002).

Steven J. Miller and two students of the course (Leo Goldmakher and Atul Pokharel) gave two presentations at the International Conference on Statistics in Hawaii (June 4 to June 7, 2003). The presentations included results from Elliptic Curves, as well as an analysis of the design and implementation of such a course (from both the faculty and student points of view). For slides of the talk, go to

http://www.math.princeton.edu/mathlab/talks/HawaiiPresentation.pdf

Our presentation was mentioned prominently in a summary article of the conference; the article is available at

http://www.math.princeton.edu/mathlab/finalreport/HawaiiStatConf2003_ReportOnPton.pdf

2 $2000-2001$

2.1 Class Structure

For the first third of the semester Professor Peter Sarnak and Steven Miller lectured twice a week. The first was a background talk, and the second would cover the topic in more depth and end with a list of unsolved conjectures that are amenable to experimental testing. Topics covered were

1. Hardy-Littlewood Circle Method
2. Random Matrix Theory
3. Ramanujan Graphs
4. Elliptic Curves

By the end of the first third of the semester, each junior had chosen a problem to investigate.

During the second third of the term the juniors split their time between researching their problem and writing programs to test unsolved conjectures. Additional lectures on the above topics were now alternated with the juniors giving brief reports of their findings. Even though most students were doing very different research projects, by attending the same lectures and listening to periodic reports students learned about all the topics. Moreover, while not doing identical topics, different students often faced similar optimization / coding problems. Using class time to constantly discuss work resulted in brainstorming sessions between the Juniors and the faculty advisors.

In the final third of the semester the juniors finished their investigations, wrote up their reports, and gave longer presentations to the class.

2.2 Problems Investigated

The juniors investigated the following topics:

2.2.1 Random Matrix Theory

The local scaled spacing distribution between the eigenvalues of a random large symmetric matrix whose entries are chosen to be I.I.D. Gaussian is a well known celebrated result from random matrix theory. It is believed (and is a central open problem) that the same universal laws hold for the eigenvalue spacings if one chooses the matrix entries as I.I.D., but not necessarily Gaussian. There is not that much evidence for such a universality conjecture and numerical experiments would be illuminating (there are some done for matrices of size 20x20 going back to the 50's and 60's, but one can go a lot further with today's eigenvalue algorithms and computers). Two juniors investigated 400x400 matrices with elements chosen from the Poisson, Cauchy, and Uniform distributions, as well as the eigenvalue distributions of sparse and band matrices. In all cases, excellent agreement with theory was observed.

2.2.2 Ramanujan Graphs

Let X_N be a random cubic graph on N vertices, i.e., 3 edges emanate from each vertex. Let A be the adjacency matrix for X_N, i.e., $a_{vw} = 1$ if v is joined to w and 0 otherwise. A is symmetric and its eigenvalues all lie in $[-3, 3]$. The largest eigenvalue is equal to 3. Denote by $\lambda_1(X)$ the next to largest eigenvalue of A. In the construction of highly connected sparse graphs ("expanders"), it is desirable to have $\lambda_1(X)$ as small as possible. It is known that for any sequence of X_N's

$$\liminf_{N\to\infty} \lambda_1(X) \geq 2\sqrt{2}$$

The question is whether, with probability 1, $\lambda_1(X) \to 2\sqrt{2}$ as $N \to \infty$. If the answer to the above question is yes, then the random graph would be as good (for this purpose) as the arithmetically constructed Ramanujan graph.

Two juniors investigated large collections of random graphs with N vertices (N up to 1000). One junior investigated by generating large numbers of graphs randomly; the other by performing a random walk on the space of graphs. As the number of vertices increased, the average of the λ_1's seem to approach $2\sqrt{2}$, and the standard deviation seems to decrease.

There are known constructions for families of Ramanujan graphs using Number Theory for $(p^r + 1)$-regular graphs. The juniors also investigated 7-regular graphs (the first number with no known construction), and observed similar results.

2.2.3 Hardy-Littlewood Varieties

Let V be a variety (for example a hypersurface) defined by integral polynomial equations in $(x_1, \ldots, x_n)$. Let $V(Z)$ be the integral points in V and let $N_V(T)$ be the number of integral solutions x with $||x|| \leq T$. We say V is a "Hardy-Littlewood Variety" if the asymptotic behavior of $N_V(T)$ as $T \to \infty$ is given by a product of local densities which count the number of solutions in V over the real numbers and the p-adic numbers (the exact form of this product of local densities comes from the Hardy-Littlewood method). The known Hardy-Littlewood varieties are ones shown to be so by their method, which requires that n be very large compared to the degrees of the defining equations. However, it appears (from a number of points of view) that to be a Hardy-Littlewood variety is not so restrictive. A junior examined this by looking at a nonsingular cubic form F in six variables, and observed the results predicted by the main arcs.

2.2.4 Prime Spacings

It is a folklore belief that the local spacings between consecutive primes follow the laws of random numbers. A junior investigated both spacings between adjacent primes and spacings between primes in arithmetic progressions for p about 10^{15}, observing excellent agreement with predictions.

2.2.5 Elliptic Curves

An L-function may be associated to an elliptic curve, and one can investigate whether or not the Generalized Riemann Hypothesis holds (namely, all non-trivial zeros should have real part equal to $\frac{1}{2}$). Let m be the order of vanishing of the L-function at the central point $s = \frac{1}{2}$. The Birch and Swinnterton-Dyer conjecture asserts that m equals the order of the Mordell-Weil group. If both conjectures are true, a contour integral of a certain transform of the logarithmic derivative of the L-function (after normalization) should have mean m and error term of a certain size. A junior investigated this for a variety of Elliptic Curves where the rank of the Mordell-Weil group was known. For small rank (rank 0 and 1), good agreement was observed.

2.2.6 $\{n^2\alpha\}$

It is known that for any irrational α, the fractional parts of $n^2\alpha$ for $n \in \mathbf{N}$ are uniformly distributed. However, for most α it is not known whether the distribution of the fractional parts is Poissonian; while it is known that in some special cases the distribution is not Poissonian, it is conjectured that these numbers form a set of measure zero. A junior numerical observed Poissonian distributions for specific α (such as π, e, and $\sqrt{2}$) by examining the consecutive spacing measures, as well as investigating the consecutive spacing measure for an irrational number known not to yield a Poissonian distribution.

3 $2001 - 2002$

3.1 Class Structure

For the first half of the Fall semester, Professor Andrew Wiles and Steven J. Miller lectured twice a week on background material and basic elliptic curve theory. After all the problems were introduced, the juniors chose their problems. The class continued to meet in small groups, where advanced material for the relevant problems was given in more detail.

Unlike the previous year, all problems shared a common theme, namely an attempt to investigate the observed excess rank phenomenon in families of elliptic curves.

Consider one-parameter families of elliptic curves. By construction, we can force points of infinite order on the curves. Let us assume we have a family where we have forced r points on each curve E_t. By the Birch and Swinnerton-Dyer Conjecture, the geometric rank of the group of rational points should equal the order of vanishing of the associated L-function at the critical point. Assuming the Sign Conjecture, we might expect half the curves to have rank r, and half to have rank $r+1$. (The functional equation being odd / even forces another zero). Previous investigations (see, for example, Fermigier) for many curves with small values for t have shown that this is not the case; ie, that for small t, for many families about 32% have rank r, 48% have rank $r+1$, 18% have rank $r+2$, and 2% have rank $r+3$. Is this excess rank above the forced rank a general phenomenon that persists in the limit, or the result of small numbers?

3.2 Investigations

To numerically investigate the above, the juniors split into several groups. The idea was to study very thoroughly lots of properties of a small number of families, and look for correlations.

Properties studied were the analytic rank (the order of vanishing of the L-function at the critical point), the geometric rank (the order of the Mordell-Weil group), the number of points of low naive and low canonical height, the distribution of the signs of the functional equation, the distribution of $a_t(p)$ (Sato-Tate Conjecture), and the distribution of the first zero above the critical point ($s = \frac{1}{2}$).

For some projects (in particular, the sign investigations) the programs ran very quickly; for others (points of low height, first zero above the critical point) the computation time is significantly longer.

During the Spring semester the juniors and Steve Miller continued to refine the programs. Upon completion, a detailed analysis will be done of the above quantities. The goal is to have a publication for a journal such as Experimental Mathematics, where certain qualitative phenomenon can be reported.

For completeness, a more thorough description of each problem is provided below.

3.2.1 The Sign Conjecture

We can analytically continue the L-function of an elliptic curve, and obtain a functional equation $\Lambda(s, E) = \epsilon_E \Lambda(2-s, E)$, where $\epsilon_E = \pm 1$. The Sign Conjecture for a one-parameter family of elliptic curves asserts that (in the limit) $\epsilon_E = 1$ and $\epsilon_E = -1$ are equally likely. If the family is sufficiently general, we expect the Sign Conjecture to hold.

One student investigated the distribution of signs for 10^7 curves in several such families.

3.2.2 Points of Low Height

For any elliptic curve, we can study the rational solutions $P = (x, y)$. Let $x = \frac{p}{q}$ in lowest terms. The naive height is defined as $h(P) = \max(|p|, |q|)$. Associated to the naive height is the more technical canonical height, which has several useful properties. In particular, the canonical height allows us to determine if rational points are independent.

Juniors calculated all rational solutions with low height for many elliptic curves for two purposes: first, to be used to determine the geometric rank (see below); second, to investigate correlations between the number of points of low height and excess rank. From the Gross-Zagier formula, we know there can be relations between rank and height of points.

3.2.3 Analytic and Geometric Rank

In its simplest form, the Birch and Swinnerton-Dyer Conjecture states the geometric rank of an elliptic curve should equal the analytic rank. Using the canonical height and the lists of observed points of low height, juniors obtained lower bounds for the geometric rank; using standard elliptic curve algorithms (although coding them efficiently was non-trivial), the juniors wrote code to approximate the analytic rank.

All of Fermigier's examples are for one-parameter families where $A_{\mathcal{E}}(p) = \frac{1}{p} \sum_{t(p)} a_t(p)$ is constant. These are natural families to consider, as it is easy to force rational independent points on the curve and create families with rank over $\mathbf{Q}(t)$; however, these may not have the same behavior as a generic one-parameter family. Therefore, more general families will also be investigated.

3.2.4 Sato-Tate

By Hasse, $|a_p| \leq 2\sqrt{p}$; we can therefore write a_p as $2\cos\theta_p$. For curves without complex multiplication, it is conjectured that the angles θ_p are uniformly distributed with respect to the Sato-Tate density, ie, in the limit the percent of curves with angles in $[a, b]$ is $\frac{1}{2\pi} \int_a^b \sin^2 t dt$.

If, however, the elliptic curve has non-zero rank, a lower order bias term should appear in the Sato-Tate conjecture. While this bias will be too small to observe, it is hoped that some of the consequences (such as a predisposition of the $a_t(p)$'s to be negative) will be observable.

3.2.5 Lowest Zero

We expect the zeros of an elliptic curve to follow the GUE distribution. To see this may be beyond the computational ability of the UML (to see the convergence may require more computations than we can do on our machines; many such investigations use large amounts of time on supercomputers). We can, however, look at the first few zeros near the critical point. If we are (again) looking at a family, we can study the distribution of the lowest (or second lowest) zero. We may want to normalize with respect to the logarithm of the conductor.

By looking at functions related to $L(s, E)$, we can construct a real valued function on the line $Re(s) = \frac{1}{2}$ which has the same zeros as $L(s, E)$. By looking for sign changes in this function, we can find the zeros; by standard methods of Complex Analysis, we can do contour integrals to find the number of zeros in a given region. If the number of sign changes equal the contour integral, then we have proved the Generalized Riemann Hypothesis in a region. Note the contour integral is almost certainly not an integer. We will have to approximate it, and then knowing it is an integer will allow us to move from an approximate counting of the number of zeros to an exact counting.

4 $2002 - 2003$

Enrollment increased from 8 students in $2000 - 2001$ to 11 students in $2001 - 2002$ to 17 students in $2002 - 2003$. To accomodate this increase, the staff of the Junior Research Seminar was increased from 2 to 5. Each semester there were two professors, two graduate students, and one undergraduate computer TA (whose job was to answer basic computer questions).

4.1 Fall 2002

Instructors were Professors Steven J. Miller and Ramin Takloo-Bighash, graduate TAs were Harald Helfgott and Florin Spinu, and the undergraudate TA was Salman Butt.

The class structure was similar to previous years, except there was an extensive series of lectures by Ramin Takloo-Bighash leading up to the proof of Roth's Theorem over $\mathbb{Q}$.

Additionally, all course notes were typed up and made available on the web. The course notes are the nucleus for a textbook written by Miller and Takloo-Bighash.

Topics investigated were

4.1.1 Digits of Continued Fractions

For almost all numbers (in the sense of measure) in $(0,1)$, the probability that the n^{th} digit in the Continued Fraction expansion is $k \in \mathbb{N}$ equals $\log_2\left(1+\frac{1}{k(k+2)}\right)$ as n tends to infinity (the Gauss-Kuzmin Theorem). The continued fraction of x is finite if and only if x is rational; it is periodic if and only if x is a quadratic irrational. Several different experiments were run to try and determine if special sets were exceptions to the Gauss-Kuzmin rule.

For example, all algebraic numbers form a set of measure zero; thus, the Gauss-Kuzmin theorem does not yield any information about the distribution of their digits. Algebraic numbers of degree 3 and higher were tested. Detailed statistical analyses were performed, indicating terrific agreement with Kuzmin's Theorem.

Additionally, special values of special functions (Riemann-Zeta function, Gamma function) were tested; the results were consistent with the Gauss-Kuzmin distribution.

4.1.2 Closed Form Expansions of Continued Fractions

Closed form expressions for the continued fractions of many quadratic irrationals are derived. Consider a finite difference equation satisfying

- $G_{n+1} = a_n G_n + b_n G_{n-1}$.
- $a_n = m$, where $m \in \mathbb{N}$ and $b_n = 1$ for all n;

note $m = 1$ gives the Fibonacci numbers.

Let G_n denote the n^{th} term of the sequence, and let τ denote the $\lim_{n\to\infty} \frac{G_{n+1}}{G_n}$. We derive formulas for $\frac{G_n}{G_{n-k}}$; in particular, these allow us to determine closed form continued fraction expansions of τ^k for any positive integer k.

These formulas were first conjectured by examining results of numerical investigations of such ratios. In fact, the results of the simulations were essential in "finding" the proofs.

4.1.3 Periodicity in Continued Fractions

These investigations dealt with the length of the period of the continued fraction expansion of $\sqrt{n}$ as a function of n, establishing a (new?) result concerning the average period length for $k < \sqrt{n} < k+1$, where k is an integer, and, following numerical experiments, tried to formulate the best possible bounds for this average length and for the maximum length of the period of the continued fraction expansion of $\sqrt{n}$, with $\lfloor\sqrt{n} = k$.

4.1.4 Poissonian and Non-Poissonian Behavior in $\{n^k\alpha\}$

The ordered spacings between $\{n^2\alpha\}$, n ranging from 1 to N_m, are known to have non-Poissonian behavior for certain N_m. An analysis was made of the length around these N_m one must travel before recovering Poissonian behavior. Additionally, the ordered spacings were studied for n ranging from N to $N + M$ for M much smaller than N, both for $\{n^2\alpha\}$ and $\{n^3\alpha\}$. For such small ranges, nothing is known theoretically about how the spacings should behave.

4.1.5 Rational Relations of Continued Fractions

Many continued fractions are almost periodic, where instead of having k digits repeat, we have a block of length k, with each entry a linear function of the block number; we call these Linearly Periodic numbers. For several interesting numbers x, the continued fraction expansions of $x + \frac{p}{q}$ are calculated, and an analysis of the dependence of the block length and functions on $\frac{p}{q}$ is performed.

4.1.6 Lone Runner Problem

The Lone Runner Conjecture says that if there are $k + 1$ runners travelling around a circular track with constant and distinct velocities, there will be a time when one runner is at least a distance of $\frac{1}{k+1}$ from the others runners, i.e. he is "lonely." Numerical and theoretical investigations were performed.

4.2 Spring 2003

Staff: Professors Yakov Sinai, Steven J. Miller; graduate students Alexander Bufetov and Lior Silberman; undergraduate TA: Atul Pokharel.

Topics investigated include the Goldbach Conjecture, equidistribution of roots of polynomials mod p, the dynamical piston, the $3x + 1$ problem, Random Graphs, low lying zeros of elliptic curves.

Reports are available on-line.

4.3 Web Database of Results

In addition to providing undergraduates with research experience in investigating unsolved conjectures, the UML aims to create a database of such research to assist future investigations. Many of previous semesters' projects naturally lead to future investigations. These future research topics have been noted, and the documentation of the current researcher (programs, papers) will facilitate such future work.

5 Class Resources

Fine 408 is reserved solely for undergraduate use. There are six computers, a blackboard, couch, refrigerator, and work space.

A variety of computational programs and packages are available, including C, Java, Magma, Maple, Mathematica, Matlab, and Pari. An extensive library of sample computer programs (in the various langauges) has been assembled.

6 Sample of Results

Statements of the problems, programs written, and final reports are available online at http://www.math.princeton.edu/ $\sim$ mathlab/index.html.

For definiteness, we give a sample of results obtained in the Undergraduate Mathematics Laboratory.

6.1 Random Matrix Theory

The results below are joint with Rebecca Lehman and Yi-Kai Liu.

Consider $N \times N$ symmetric matrices with entries i.i.d.r.v. chosen from a fixed probability distribution P.

GOE Conjecture: As $N \to \infty$, the probability density of the distance between two consecutive (normalized) eigenvalues approaches $\frac{\pi^2}{4}\frac{d^2\Psi}{dt^2}$ (the GOE distr). $\Psi(t)$ is (up to constants) the Fredholm determinant of the operator $f \longrightarrow \int_{-t}^{t} K * f$, with kernel $K = \frac{1}{2\pi}\left(\frac{\sin(\xi-\eta)}{\xi-\eta} + \frac{\sin(\xi+\eta)}{\xi+\eta}\right)$.

This is only known if the entries are chosen from the Gaussian. The consecutive spacings are well approximated by Axe^{-Bx^2}.

Semi-Circle Law: Assume P has mean 0, variance 1, other moments finite, and let $\frac{\lambda_j}{2\sqrt{N}}$ be the normalized eigenvectors.

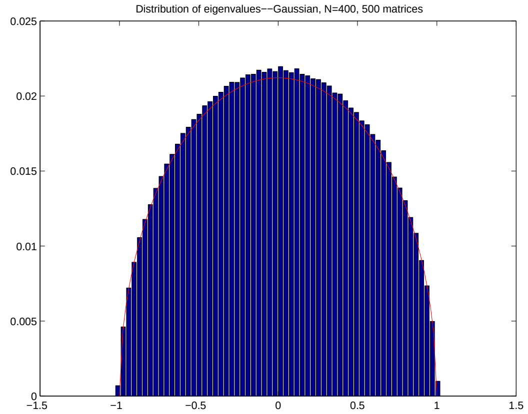
$$\begin{aligned} \text{If} \quad \mu_{A,N}(x) &= \frac{1}{N}\sum_{j=1}^{N}\delta\left(x-\frac{\lambda_j}{2\sqrt{N}}\right) \\ \text{Then} \quad \mu_{A,N}(x) &\to \frac{2}{\pi}\sqrt{1-x^2} \text{ with probability } 1 \end{aligned}$$

The juniors investigated many probability distributions. In every case, they observed the normalized distances between the eigenvalues converging towards the GOE distribution as the size of the matrices increased.

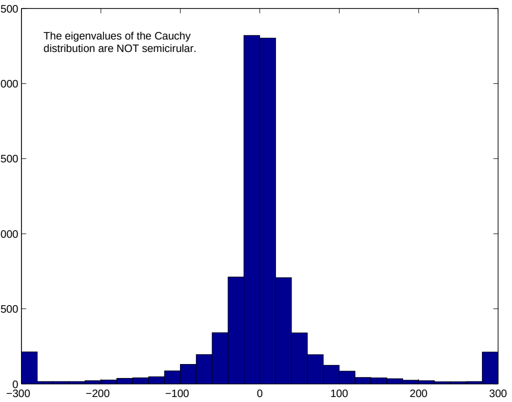
Probability distributions chosen include the uniform on $[-1,1]$, the Cauchy (which has infinite variance and hence does not satisfy the conditions for the Semi-Circle Law), the discrete Poisson, sparse matrices (entries are $+1$ with probability p, -1 with probability p, and 0 with probability $1-2p$) and Gaussian band matrices.

6.1.1 Semi-Circle Law

Already at $N = 400$ we observe good agreement in the Semi-Circle Law for matrices with entries chosen from the Gaussian. Not surprising, the Cauchy Distribution (with infinite variance) does not satisfy the Semi-Circle Law. In particular, we observe with significant probability large eigenvalues. Later we observe, however, that the spacings between the central (ie, staying in the bulk of the spectrum) normalized eigenvalues arising from the Cauchy Distribution, as the size of the matrices increases, tend to the GOE.



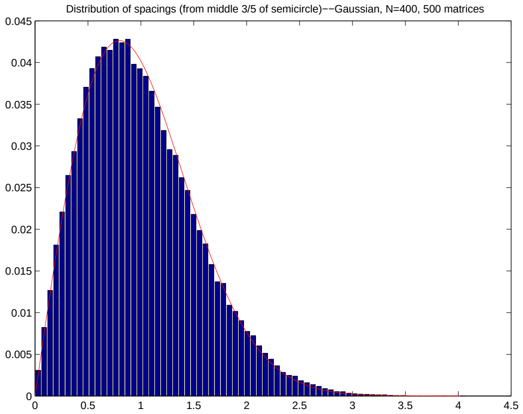
500 Matrices: Gaussian 400×400



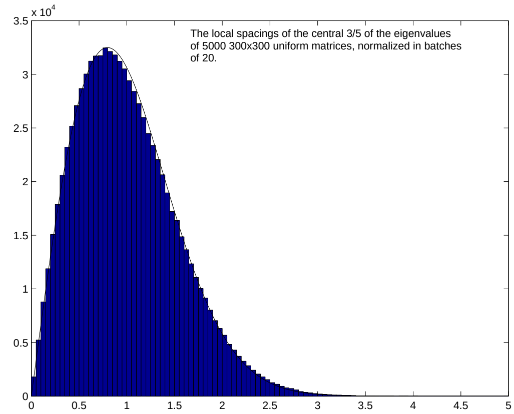
Cauchy: Not-Semicircular (Infinite Variance), $P(t) = \frac{1}{\pi(1+t^2)}$

6.1.2 GOE Conjecture

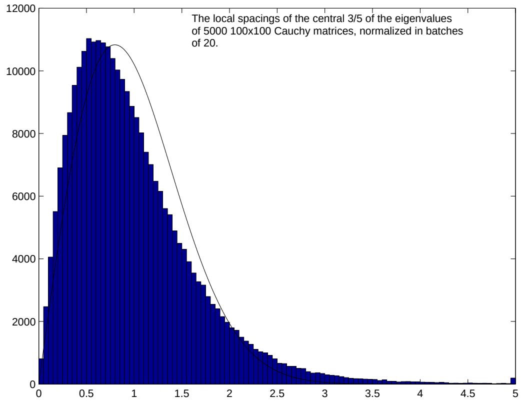
For comparison purposes, below is the distribution of spacings between normalized eigenvalues when the entries are chosen from the Gaussian distribution:



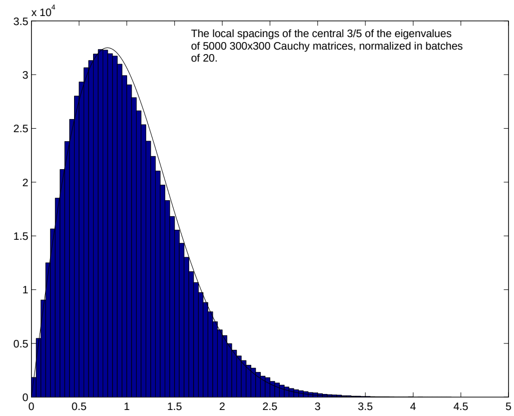
400 Gaussian Matrices, 400×400



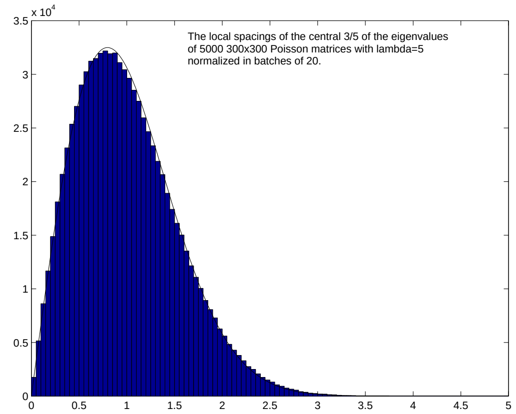
5000 Uniform on $[-1, 1]$ matrices: 300×300



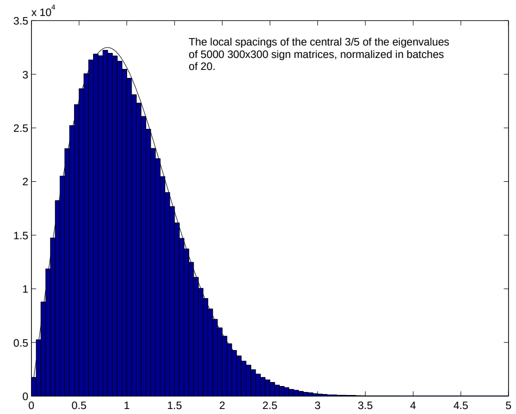
5000 Cauchy matrices: 100×100



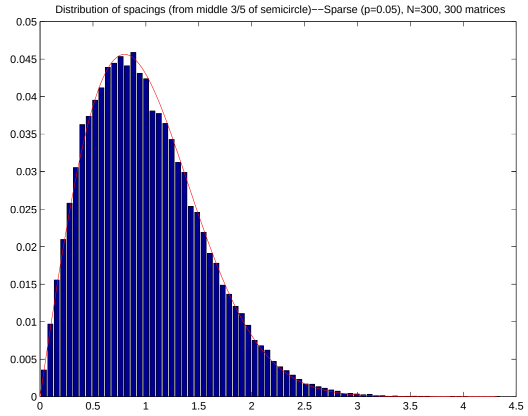
5000 Cauchy matrices: 300×300



5000 Poisson matrices, $P(n) = \frac{\lambda^n}{n!}e^{-\lambda}$, $\lambda = 5$, 300×300



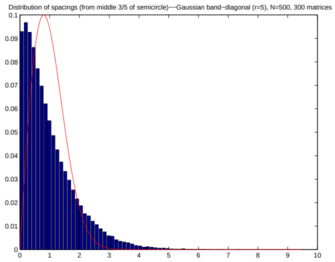
5000 Poisson matrices, $P(n) = \frac{\lambda^n}{n!}e^{-\lambda}$, $\lambda = 20$, 300×300



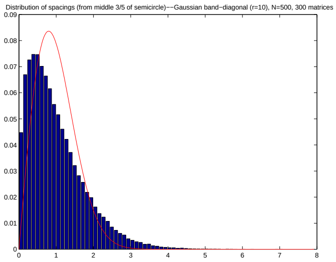
300 Sparse Matrices, $p = .05$ for $+1$, $p = .05$ for -1, $p = .90$ for 0, 300×300

6.1.3 Band Matrices

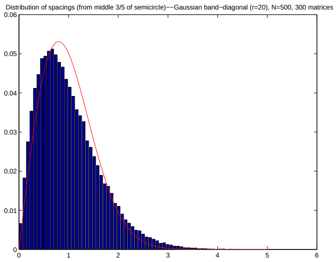
The juniors also investigated band matrices of width r (a matrix with non-zero entries along the first r diagonals above and below the main diagonal). The entries are chosen from the Gaussian distribution.



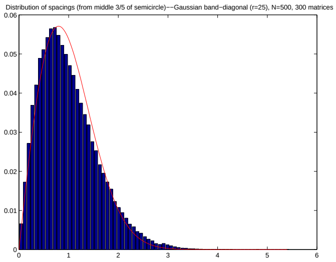
300 Band Matrices, 500×500, $r = 5$



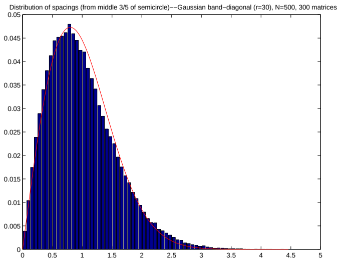
300 Band Matrices, 500×500, $r = 10$



300 Band Matrices, 500×500, $r = 20$



300 Band Matrices, 500×500, $r = 25$



300 Band Matrices, 500×500, $r = 30$

6.2 Ramanujan Graphs

The results below are joint with Peter Richter and Kevin Chang.

Let G_n be a family of k-regular graphs with n vertices, and let the adjacency matrix have eigenvalues λ_j.

1. $\lambda_0(G) = k$ for all $G \in G_n$
2. $\lambda_0(G) > \lambda_1(G)$ iff connected
3. $\liminf_{n\to\infty} \lambda_1(G_n) \geq 2\sqrt{k-1}$

We define the expander constant of a graph X with n vertices V as the largest constant $h(X)$ such that $|\partial A| \geq h(X)|A|$ for all subsets A of V with $\#A \leq \frac{n}{2}$. $|\partial A|$ is the boundary of A (the set of vertices v in $V - A$ with an edge from v to a vertex in A.

We have the following facts:

$$\frac{k-\lambda_1}{2} \;\leq\; h(X) \;\leq\; \sqrt{2k(k-\lambda_1)}.$$

For bipartite graphs:

$$\text{diam}(X) \;\leq\; \frac{\log(2n)}{\log\left(\frac{k+\sqrt{k^2-\lambda_1^2}}{\lambda_1}\right)} \tag{1}$$

where $\text{diam}(X)$ is the largest path between two vertices of X.

Graphs with small λ_1 have small diameters and high expander constants.

We say a k-regular graph is Ramanujan if $\lambda_1 \leq 2\sqrt{k-1}$. These give sparse graphs with small diameters and high connectivity, and are useful for network building. There are known constructions (using Number Theory) for $p^r + 1$, p prime.

The juniors generated large numbers of k-regular bipartite graphs and calculated λ_1 (the second largest eigenvalue).

6.2.1 Questions / Conjectures

Consider all 3-regular bipartite graphs with n vertices.

Question 1: As $n \to \infty$, what percent of the graphs are Ramanujan?

Question 2: As $n \to \infty$, does each graph have $\lambda_1 \to 2\sqrt{2}$?

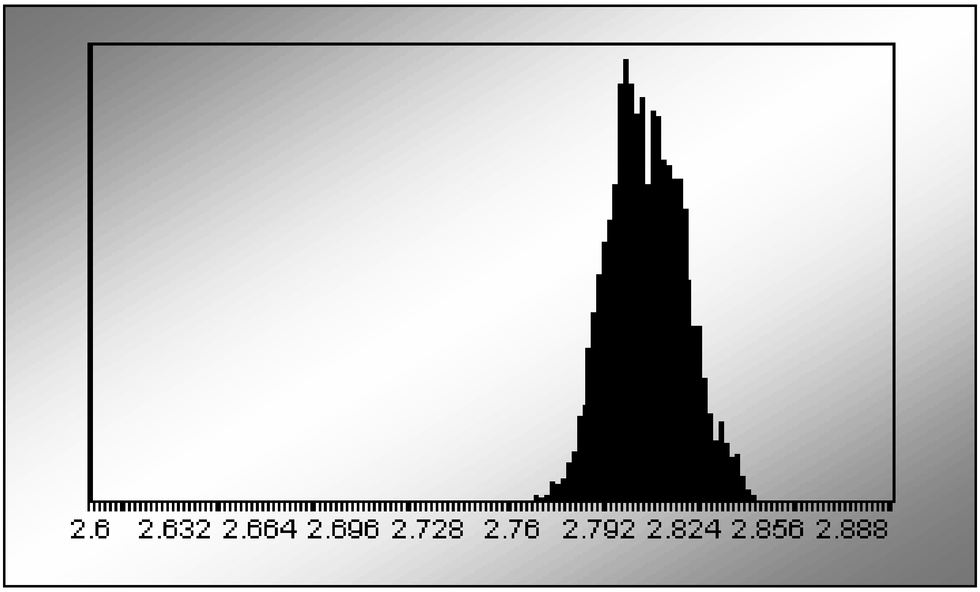
In other words, in the limit is a randomly chosen 3-regular bipartite graph Ramanujan? There are similar questions for n-regular bipartite graphs. Note 7 is smallest number with no known construction.

6.2.2 Results for $k = 3$

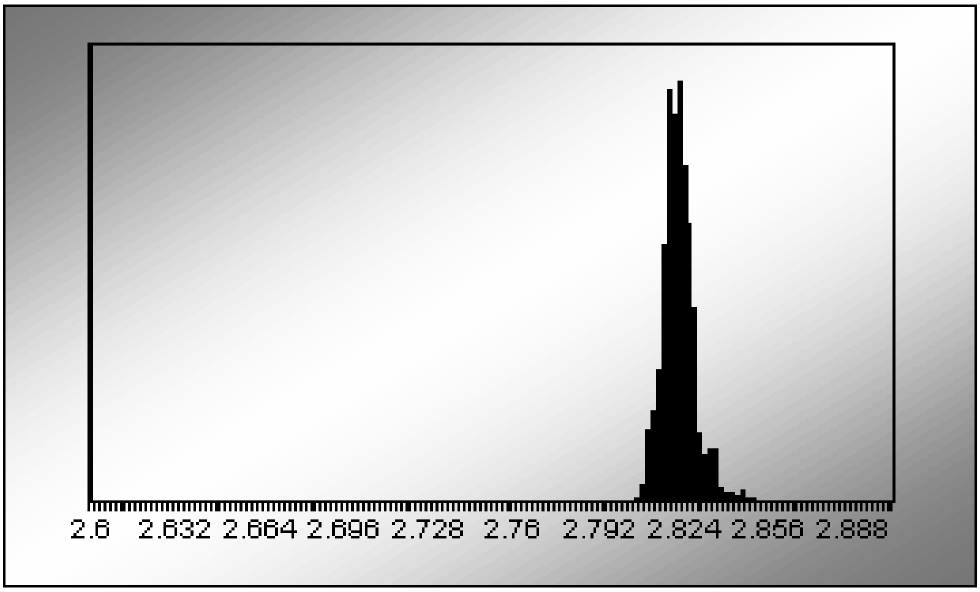
The juniors investigated 5000 randomly chosen 3-regular bipartite graphs for various numbers of vertices.

n	λ_1 mean	st dev	% Ram	λ_1 mean	st dev	% Ram
100	2.8076	0.042	76.14	2.777	0.031	95.28
200	2.8160	0.027	76.36	2.800	0.019	93.06
300	2.8187	0.020	77.38	2.808	0.014	92.84
400	2.8210	0.018	75.20	2.813	0.011	91.22
500	2.8216	0.014	76.62	2.815	0.009	91.40
600	2.8225	0.013	77.54	2.817	0.009	90.90
700	2.8226	0.012	78.46	2.818	0.008	91.00
800	2.8231	0.011	79.68	2.819	0.007	90.58
900	2.8233	0.011	80.34	2.820	0.007	91.06
1000	2.8235	0.009	79.86	2.820	0.006	91.12

First group allows double and triple bonds; second group simple (single bonds only). Note $2\sqrt{3-1} = 2\sqrt{2} = 2.828427$.



λ_1: 5000 simple 3-regular graphs, 300 vertices



λ_1: 5000 simple 3-regular graphs, 1000 vertices

6.2.3 Results: $k = 7$

The juniors investigated 5000 randomly chosen 7-regular bipartite graphs for various numbers of vertices.

n	λ_1 mean	st dev	% Ram	λ_1 mean	st dev	% Ram
100	4.791	0.113	83.74	4.530	0.100	99.90
200	4.833	0.069	82.68	4.709	0.063	99.70
300	4.849	0.053	83.54	4.767	0.048	99.42
400	4.858	0.043	82.90	4.796	0.040	98.92
500	4.865	0.036	82.92	4.815	0.035	98.77
600	4.869	0.032	83.20	4.828	0.031	98.26
700	4.871	0.028	84.02	4.836	0.028	98.20
800	4.874	0.027	83.18			
900	4.875	0.025	82.84			
1000	4.877	0.022	83.92			

First group allows multiple bonds; second group single bonds only. Note $2\sqrt{7-1} = 4.89898$.

6.2.4 Conclusions

From the data, we observe that as the number of vertices increase, λ_1's distribution is tightening around $2\sqrt{k-1}$ for $k = 3$ and 7. The percent of random graphs which are Ramanujan seems stable (with the percent depending on whether or not we allow multiple bonds), though no one wants to conjecture that it remains stable!

7 Future Projects of the UML

In addition to providing undergraduates with research experience in investigating unsolved conjectures, the UML aims to create a database of such research to assist future investigations. Many of previous semesters' projects naturally lead to future investigations. These future research topics have been noted, and the documentation of the current researcher (programs, papers) will facilitate such future work.

7.1 New Textbooks

Professors Steven J. Miller and Ramin Takloo-Bighash have written an undergraduate textbook based on their experiences and the students' results. Tentatively titled *An Invitation to Modern Number Theory*, the book plans to be a self-contained introduction to many hot topics in number theory, and is written in the spirit of the class.

For most of the book, the only pre-requisites are basic calculus, though the relevant theorems are briefly reviewed (MVT, Taylor Series, Change of Variable). Similar to the class, the purpose of the book is to introduce undergraduate math majors to modern topics in number theory.

Topics covered include

1. *Basic Number Theory* (including mod p arithmetic, basic group theory, Quadratic Reciprocity, solving Diophantine equations, algebraic and transcendental numbers);
2. *Continued Fractions* (properties, approximations, constructing transcendental numbers, efficient algorithms for calculating continued fractions);
3. *Roth's Theorem* (applications to solving Diophantine Equations, the proof of Roth's Theorem);
4. *Probabilistic Methods, Fourier Analysis and Equidistribution* (enough background material in basic probability and Fourier Analysis is developed to tackle: (a) the distribution of digits of a continued fraction (Gauss-Kuzmin Theorem); (b) equidistribution of $\{n^k\alpha\}$; Poissonian behavior of $\{n^k\alpha\}$);
5. *Number Theory and Random Matrix Theory* (introduction to L-functions, the Hardy-Littlewood Circle Method and Applications, basic Random Matrix Theory (especially the semi-circle law, GOE behavior, and band matrices), and similarities between ensembles of Random Matrices and families of L-functions.

The book (about 400 pages) will be published by Princeton University Press in Spring 2005, and is designed to be used either for a standard seminar course, or to run a research class. Numerous open problems are described in detail. Further, many results proved or observed by the students are incorporated in the book.

7.2 Vertically Integrated Summer Program in Computational Mathematics

Building on the success of the Undergraduate Research classes, Brian Conrey, David Farmer, Chris Hughes, Steven J. Miller and Michael Rubinstein ran a more intensive summer program in 2003 at the American Institute of Mathematics. We ultimately envision this as a yearly event, with around 10 faculty, 20 graduate students, and 40 undergraduates. There were seven undergraduates (including two undergraduates from Princeton).

The program was similar in spirit to the regular research classes; however, as students' time will not be split with other classes and activities, we saw significantly greater productivity. Further, a core goal was for the faculty to be working towards proofs of various problems. With a much better faculty-student ratio, we feel and saw that this is a realistic goal.

Problems were chosen from Computational Number Theory, especially Random Matrix Theory and zeros of L-functions. For more information, see

http://aimath.org/reu.html

For student reports, see

http://www.math.brown.edu/ $\sim$ sjmiller/AIM/index.htm

8 Contact Information

For questions on the Undergraduate Mathematics Laboratory / Junior Research Seminar, please contact Professor Peter Sarnak (sarnak@math.princeton.edu) or Steven J. Miller (sjmiller@math.brown.edu).