Honors 1: Undergraduate Math Lab[1]

Peter Sarnak[2], Steven J. Miller[3], Alex Barnett[4]

Courant Institute of Mathematical Sciences
New York University
New York, NY

November 19, 2002

[1]Homepage: http://www.math.nyu.edu/~millerj/
[2]E-mail: sarnak@math.princeton.edu
[3]E-mail: millerj@cims.nyu.edu or sjmiller@math.princeton.edu
[4]E-mail: barnett@nmr.mgh.harvard.edu

Abstract

The purpose of the Undergraduate Mathematics Laboratory 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.

The undergrads often have enough theory to understand the basic framework and proofs of simple cases. Building on this, they then numerically test the conjectures. The undergrads 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. A similar course has been run at Princeton ($2000 - 2002$). Many of the problems investigated by the Princeton students arose from graduate dissertations or current faculty research. It has been very easy £nding 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.

Below are the notes from the NYU Fall 2002 class.

Problem List

1. **Primality Testing:**

   In a major theoretical breakthrough, Manindra Agarwal, Nitin Saxena and Neeraj Kayal discovered a deterministic polynomial time algorithm to determine if a number is prime or composite. Previous algorithms are known to be polynomial only under well believed conjectures (GRH), or are probabalistic. Some aspects of current primality testing algorithms will be explored, possibly the distribution of the least primitive root mod p.

2. **Ramanujan Graphs:**

   The construction of graphs that are highly connected but have few edges have many important applications, especially in building networks. To each graph G we can associate a matrix A, where A_{ij} is the number of edges from vertex i to vertex j. Many properties of G are controlled by the size of the second largest eigenvalue of G. One project will be to investigate the distribution of the normalized second largest eigenvalues.

3. **Randomness of Arithmetic Maps:**

   For a prime p, consider the map Inv_p which sends x to its inverse mod p, $x \mapsto \overline{x}$. One project will be to compare this map to random maps mod p. For example, let $L(p)$ be the length of the longest increasing subsequence. If p is congruent to 3 mod 4, the inverse map is a £xed-point-free signed involution, and the length of $L(p)$ can be compared to that from random £xed-point-free signed involutions (studied by Rains and others). Let $S(m,n;c)$ be the Kloosterman sum,

   $$S(m,n;c) \ = \ \sum_{x mod p} e^{2\pi i \frac{mx+n\overline{x}}{c}} \tag{1}$$

   An additional project will be to investigate $\sum_c \frac{|S(1,1;c)|^2}{c^2}$, which is related to number variance of the Upper Half Plane mod SL$(2,\mathbb{Z})$. Finally, let $\sqrt{p} \ \leq \ x \ \leq \ 2\sqrt{p}$. Arrange in increasing order the $\sqrt{p}$ numbers $\overline{x}$, and compare their spacings to Poissonian behavior.

Contents

Chapter 1

Introduction, Primality Testing, Algebra Review

We introduce basic number theory concepts and primality algorithms. Lecture by Peter Sarnak; notes by Steven J. Miller.

1.1 Primality Testing

Given n, is n prime or composite? How dif£cult is this? How long does it take? Brute force: try factors up to $\sqrt{n}$, so can do in $\sqrt{n}$ steps.

$P = NP$ problem. Deep central problem in theoretical Computer Science. P is problems solvable in polynomial number of steps (in terms of input); if equals NP, a lot of problems are solvable quickly.

Telling when n is prime: isn't supposed to be hard, but until two weeks ago, wasn't known to be a P problem.

Notation: $A(x) = O\Big(B(x)\Big)$ if there exists a $C > 0$ (which can be computed; if it cannot be computed, we say so) such that $|A(x)| \leq CB(x)$.

Example: One could show every suf£ciently large odd number is the sum of three primes. However, we didn't know how large suf£ciently large was! IE, we couldn't go through the calculation and make explicit a number N_0 such that if n is odd and greater than N_0, then n is the sum of three primes. (Note: this has been removed, and we know have another proof giving an explicit N_0).

Theorem 1.1.1 (Agrawal, Kayal, Saxena). *There is a procedure which runs in at most $O\Big(\log^{12} n\Big)$ steps determines whether n is prime. (Might be a little more* *than* $\log^{12} n$*, ie, might be something like* $\log^{12} n(\log\log n)^A$*).*

There were algorithms that were known and faster, but only known to work all the time assuming certain well believed hypotheses (Riemann Hypothesis, RH).

(Go to http://www.math.nyu.edu/$\sim$millerj/problemlist/problems.htm for a copy of their paper).

Need a feel for numbers. When n is big, $\log n$ (to any power) is much less than n. For practical applications, the size of the constant is important, as a constant of size 10^{100} would make an algorithm useless for our real world applications (ie, for the ranges we can reach). In AKS, the constants are tractable.

Technical Point: In AKS, they quote a theorem from number theory (they treat this as a black box: someone from this class will hopefully investigate this result further). *There are many primes* p *for which* $p-1$ *has a large prime factor* q*,* $q > p^{\frac{2}{3}}$. Related to Sophie Germain primes: primes p where $p-1 = 2q$, q prime. (She showed that for primes like this, $x^p + y^p = z^p$, you can solve Fermat's Last Theorem for such primes. It is not known if there are in£nitely many primes like this). AKS does not need to know there are in£nitely many Sophie Germain primes; fortunately all they need is that there are suf£ciently many primes p with $p-1$ with large prime factors.

Similar to Twin Primes: primes p_1, p_2 with $p_1 - p_2 = 2$. We don't know if there are in£nitely many twin primes, but we do have heuristics (Hardy-Littlewood) predicting how many twin primes there are (and we observe exactly that many). Sophie Germain primes are more subtle, but should be able to get heuristics. For twin primes and related quantities, see David Schmidt's report on Prime Investigations (Princeton Undergraduate Math Lab, $2000-2001$). Anyway, this would be a good project.

1.2 Arithmetic Modulo p

Number Theory: the study of whole numbers. $\mathbb{Z}$ is the integers, look at $\mathbb{Z}/n\mathbb{Z}$ $= \{0, 1, 2, \dots, n-1\}$. This is a £nite group (under addition); in fact, it is a £nite ring (can also multiply, have inverses for the non-zero elements only if n is prime).

Notation: $x \equiv y \bmod n$ means $x - y$ is a multiple of n.

Try and solve in $\mathbb{Z}$ the equation $2x+1 = 2y$. The left hand side is odd, the right hand side is even. Thus, there are no solutions. Really, just did arithmetic mod 2 or in $\mathbb{Z}/2\mathbb{Z}$. Harder: $x^2+y^2+z^2 = 8n+7$. This never has a solution. Look modulo 8. The RHS is 7 modulo 8. What are the squares mod 8? $1^1 = 1, 2^2 = 4, 3^2 = 1$, $4^2 = 0$, repeats. See there is no way to add three squares and get 7.

Idea: First, try and solve the equation modulo different primes. If you cannot solve it for some prime, then you cannot solve it over the integers.

1.2.1 Algebra Review

$\mathbb{Z}/n\mathbb{Z}$: do arithmetic over this ring. $\left(\mathbb{Z}/n\mathbb{Z}\right)^*$ are the invertible (multiplicatively) elements in the ring $\mathbb{Z}/n\mathbb{Z}$, ie, x is in $\left(\mathbb{Z}/n\mathbb{Z}\right)^*$ if there is a y such that $xy \equiv 1 \mod n$. Note: if $\gcd(x, n) > 1$ (ie, x and n have a common prime divisor p), then you cannot invert x (there is no y with $xy \equiv 1 \mod n$). Why? $xy \equiv 1 \mod n$ means $xy = 1 + \lambda n$ for some integer λ. But if $p|x$ and $p|n$, then $p|1$ which is absurd. Exercise: if $\gcd(x, n) = 1$, there is an inverse (Euclidean Algorithm).

The cardinality (number of elements in the set) of $\left(\mathbb{Z}/n\mathbb{Z}\right)^*$ is the number of $x \in \{0, 1, 2, \ldots, n-1\}$ such that $\gcd(x, n) = 1$. We denote the number of such x by $\phi(x)$, the Euler totient function. (Good Reference: H. Davenport: The Higher Arithmetic). Note that if p is prime, $\phi(p) = p - 1$. This implies that $\left|\left(\mathbb{Z}/p\mathbb{Z}\right)^*\right| = p - 1 = \mathbb{Z}/p\mathbb{Z} - \{0\}$. IE, we have a £eld if n is a prime, as every non-zero element is invertible.

$\left(\mathbb{Z}/n\mathbb{Z}\right)^*$, for any n, is a £nite Abelian group (we have inverses under multiplication, and order of multiplication doesn't matter). Finite Abelian Groups is a trivial subject: Structure Theorem for Finite Abelian Groups: product of cyclic groups.

For $n = p$ a prime, $\left(\mathbb{Z}/n\mathbb{Z}\right)^*$ is a cyclic group of order $p - 1$.

If G is a group (have identity, closed under some binary operation, have inverses with respect to the binary operation, operation is associative), we say the order of $x \in G$, $\text{ord}(x)$, is the least positive power m such that $x^m = e$, where $e \in G$ is the identity of the group. In a £nite group, every element has £nite order (proof: use the pidgeonhole principle).

Theorem 1.2.1 (Lagrange). *ord(x) | ord(G).*

Corollary 1.2.2 (Fermat's Little Theorem). *For any prime p, if $\gcd(a, p) = 1$, then $a^{p-1} \equiv 1$ mod p.*

1.2.2 Using Fermat's Little Theorem

To check and see if a number n is prime, why not check if $a^{n-1} \equiv 1 \bmod n$ if $\gcd(a,n) = 1$. It is very easy to quickly compute $\gcd(a,n)$ (use Euclid's Algorithm). Sketch: without loss of generality, let $a < n$. Write $n = b_0 a + b_1$, and now the gcd of a and n is the same as that of a and b_1 (note $0 \leq b_1 < a$).

How long does it take to raise a to the $n-1$ power? Use repeated squares and base 2 expansion. Example $100 = 64 + 32 + 4$, or $1 \cdot 2^6 + 1 \cdot 2^5 + 1 \cdot 2^2$. Thus, do a^2, $a^2 \cdot a^2$, $a^4 \cdot a^4$, $a^8 \cdot a^8$, $a^{16} \cdot a^{16}$, $a^{32} \cdot a^{32}$. Then $a^{100} = a^{64} \cdot a^{32} \cdot a^4$.

If, by choosing an a, you £nd $a^{n-1} \not\equiv 1 \bmod n$, you have a certi£cate for compositeness, but you have no idea what the factors of n are!

When a machine uses Fermat's Little Theorem, it randomly chooses a £xed number of a's between 1 and $n-1$. Problems: how many times should you run these tests to be very con£dent of the result; are there any numbers which always pass this test, yet are composite?

About eight years ago it was proved there are in£nitely many Carmichael numbers (numbers n such that $a^{n-1} \equiv 1 \bmod n$ for all a, but n is composite).

1.2.3 Quadratic Reciprocity

p, q odd primes. We de£ne $\left(\frac{a}{p}\right)$ to be 1 if a is a non-zero square mod p, 0 if $a = 0$, and -1 otherwise (ie, if a is not a square mod p). Note a is a square mod p if there exists an $x \in \{0, 1, \dots, p-1\}$ such that $a \equiv x^2 \bmod p$. For p an odd prime, half the non-zero numbers are squares, half are not.

Exercise: $\left(\frac{a}{p}\right) = a^{\frac{p-1}{2}} \bmod p$ for odd p. Note the above squared is $a^{p-1} \equiv 1$.

Theorem 1.2.3 (Quadratic Reciprocity). $\left(\frac{q}{p}\right) = \left(\frac{p}{q}\right) \cdot (-1)^{\frac{p-1}{2}\frac{q-1}{2}}$, *$p$, q odd primes.*

Gauss gave at least four proofs of this deep result. If either p or q is equivalent to 1 mod 4, then one has $\left(\frac{q}{p}\right) = \left(\frac{p}{q}\right)$, ie, I'm a square root modulo you if you are a square root modulo me.

Carmichael numbers were behaving like primes. We want to get rid of them. Instead of testing $a^{n-1} \equiv 1 \bmod n$, test and see if $a^{\frac{n-1}{2}} \equiv \left(\frac{a}{n}\right) \bmod n$. Similar to the Euclidean Algorithm, can computer $\left(\frac{a}{n}\right)$ in $logn$ steps by constant applications of Quadratic Reciprocity.

Key Test: Will test that $a^{\frac{n-1}{2}} \equiv \left(\frac{a}{n}\right) \bmod n$ for $1 \leq a \leq C \log^2 n$. If fails for some a, the number is composite. If it passes, by the Riemann Hypothesis (RH), then it is true for all a. Will then show this is a valid test for primality (ie, unlike

Carmichael numbers, if this is satis£ed for all a up to $C \log^2 n$, then n is prime. This will take about $\log^4 n$ steps).

Algebra Books: Herstein (Topics in Algebra); Birkhoff-Mclean (Algebra), Lang (Undergraduate Algebra).

1.3 Lecture Next Week

Steve will talk about reciprocity, £nite £elds.

1.4 WWW Resources

http://mathworld.wolfram.com/ is a good place to look up unfamiliar terms.

Chapter 2

Notation, Euclid's Algorithm, Lagrange's Theorem, Riemann Zeta Function

We will review notation, Euclid's Algorithm, Lagrange's Theorem, and prove there are in£nitely many primes (three ways: following Euclid, by studying the Riemann Zeta Function $\zeta(s)$ as $s \to 1$, and by analyzing $\zeta(2)$. Lecture by Steven J. Miller; notes by Alex Barnett and Steven J. Miller.

2.1 Notation

- $a|b$: a divides b, *i.e.* the remainder after integer division $\frac{b}{a}$ is 0.
- (a, b) : Greatest Common Divisor (GCD) of a and b. Sometimes written $\gcd(a, b)$.
- $x \equiv y(\mod n)$: An equality once both sides of the equation have been taken modulo n. Equivalently, there exists an integer a such that $x = y + an$.
- wlog : 'without loss of generality'. For example, if we have two numbers x and y, it is often convenient to know which is larger and which is smaller. Without loss of generality, we can say $x \leq y$, as the case $x \geq y$ is handled identically (after permuting the variables).
- s.t. : such that.

- $\forall$: for all.
- $\exists$: there exists.
- big O notation : $A(x) = O(B(x))$, read "$A(x)$ is of order $B(x)$", is shorthand for, there is a $C > 0$ (which we can explicitly calculate), s.t. $|A(x)| \leq C\,B(x), \forall x$.
- $|S|$ or $\#S$: number of elements in the set S.
- $\#\{$*condition*$\}$: number of objects satisfying the *condition*.
- p : unless otherwise stated, a prime number.
- $\mathbb{Z}$: the set of integers.
- $\mathbb{Z}/n\mathbb{Z}$: the additive group of integers mod n.
- $(\mathbb{Z}/n\mathbb{Z})^*$: the multiplicative group of invertible elements mod n.
- $\mathbb{Q}$: the set of rational numbers. $\mathbb{Q} = \{x : x = \frac{p}{q}, p, q \in \mathbb{Z}, q \neq 0\}$.
- $\mathbb{R}$: the set of real numbers.
- $\mathbb{C}$: the set of complex numbers.
- $\left(\frac{a}{p}\right)$: Legendre symbol of a and p, de£ned as

$$\left(\frac{a}{p}\right) = \begin{cases} 0, & \text{if } p|a, \text{that is, } a = 0(\mod p) \\ 1, & \text{if } \exists x \text{ s.t. } x^2 = a(\mod p) \\ -1, & \text{if the above does not exist.} \end{cases} \quad (2.1)$$

  The symbol tests the question, "Does a have a square root in the *£eld* of arithmetic modulo p?"
- 'weak bound' : an inequality constraining some quantity which does a very poor job of getting close to the true size of the quantity. That is, a not very useful bound.

2.2 Euclid's algorithm for GCD

Tells you if two positive integers x and y have a GCD greater than 1 by £nding it. Therefore it's a 'constructive proof'. It is also 'deterministic' (involves no random choices). A fast procedure, *i.e.* takes only O(log y) steps. Remember that the number of digits $\propto \log y$. wlog we take $y > x$.

Each step is the 'black box' integer division routine we'll call D, which given the pair x, y returns the pair of integers b, r s.t. $r < x$ and satisfying

$$y = bx + r. \tag{2.2}$$

Note that this step is polynomial in the number of digits (probably $\sim (\log y)^2$ — anyone?).

ALGORITHM:

Start with the pair y, x and perform D to get b_1, r_1.
Perform D on x, r_1 to get b_2, r_2.
Perform D on r_1, r_2 to get b_3, r_3.

$$\vdots$$

Perform D on r_{n-2}, r_{n-1} to get b_n, r_n.

Stop when r_n is either

- 0, in which case r_{n-1} is the GCD, or
- 1, in which case the GCD is 1, that is, x and y are relatively prime.

The procedure works because the D step gives an r which inherits *all* common divisors of y and x. This is easy to see by writing D as $r = y - bx$. Therefore all the adjacent pairs in the sequence $y, x, r_1, r_2 \cdots r_{n-1}$ share the same GCDs. The sequence is also descending $y > x > r_1 > r_2 > \cdots > r_{n-1}$, so must reach the case that $r_{n-1} | r_{n-2}$, in which case r_{n-1} is the GCD and $r_n = 0$, or that a remainder of 1 is reached, which implies no common divisors. We have a worst-case scenario that each remainder is smaller than the previous by a constant factor $c < 1$ (I believe this is the inverse of the Golden Ratio $(\sqrt{5} - 1)/2 \approx 0.618 \cdots$ — anyone?), giving geometric (exponential) shrinkage of the r's. Therefore the worst case is that the answer is reached in $n = O(\log\ y)$ steps, and the whole algorithm is therefore polynomial in $\log y$.

2.3 Lagrange's Theorem

2.3.1 Basic group theory

Group G is a set of elements g_i satisfying the four conditions below, relative to some binary operation. We often use multiplicative notation ($g_1 g_2$) or additive notation ($g_1 + g_2$) to represent the binary operation. For de£niteness, we use multiplicative notation below; however, one could replace xy with $b(x, y)$ below.

If the elements of G satisfy the following four properties, then G is a group.

1. $\exists e \in G$ s.t. $\forall g \in G : eg = ge = g$. (Identity.) We often write $e = 1$ for multiplicative groups, and $e = 0$ for additive groups.

2. $\forall x, y, z \in G : (xy)z = x(yz)$. (Associativity.)

3. $\forall x \in G, \exists y \in G$ s.t. $xy = yx = e$. (Inverse.) We write $y = x^{-1}$ for multiplication, $y = -x$ for addition.

4. $\forall x, y \in G : xy \in G$. (Closure.)

If commutation holds ($\forall x, y \in G$, $xy = yx$), we say the group is Abelian. Non-abelian groups exist and are important. For example, consider the group of $N \times N$ matrices with real entries and non-zero determinant. Prove this is a group under matrix multiplication, and show this group is not commutative.

H is a *subgroup* of G if it is a group and its elements form a subset of those of G. The identity of H is the same as the identity of G. Once you've shown the elements of H are closed (ie, under the binary operation, $b(x, y) \in H$ if $x, y \in H$), then associativity in H follows from closure in H and associativity in G.

For the application to Fermat's Little Theorem you will need to know that the set $\{1, x, x^2, \cdots x^{n-1}\}$ where n is the lowest positive integer s.t. $x^n = 1$, called the *cyclic group*, is indeed a subgroup of any group G containing x, as well as n divides the order of G.

For a nice introduction to group theory see: M. Tinkham, *Group Theory and Quantum Mechanics*, (McGraw-Hill, 1964) or S. Lang, *Undergraduate Algebra*.

2.3.2 Lagrange's Theorem

The theorem states that if H is a subgroup of G then $|H|$ divides $|G|$.

First show that the set hH, *i.e.* all the elements of H premultiplied by one element, is just H rearranged (Cayley's theorem). By closure hH falls within H.

We only need to show that hh_i can never equal hh_j for two different elements $i \neq j$. If it were true, since a unique h^{-1} exists we could premultiply the equation $hh_i = hh_j$ by h^{-1} to give $h_i = h_j$, which is false. Therefore $hh_i \neq hh_j$, and we have guaranteed a 1-to-1 mapping from H to hH, so $hH = H$.

Next we show that the sets g_iH and g_jH must either be completely disjoint, or identical. Assume there is some element in both. Then $g_ih_1 = g_jh_2$. Multiplying on the right by $h_i^{-1} \in H$ (since H is a subgroup) gives $g_i = g_jh_2h_1^{-1}$. As H is a subgroup, $\exists h_3 \in H$ such that $h = h_2h_1^{-1}$. Thus $g_i = g_jh_3$. Therefore, as $h_3H = H$, $g_iH = g_jh_3H = g_jH$, and we see if the two sets have one element in common, they are identical. We call a set gH a *coset* (actually, a left coset) of H.

Clearly

$$G = \bigcup_{g \in G} gH \tag{2.3}$$

Why do we have an equality? As $g \in G$ and $H \subset G$, every set on the right is contained in G. Further, as $e \in H$, given $g \in G$, $g \in gH$. Thus, G is a subset of the right side, proving equality.

There are only £nitely many elements in G. As we go through all g in G, we see if the set gH equals one of the sets already in our list (recall we've shown two cosets are either identical or disjoint). If the set equals something already on our list, we do not include it; if it is new, we do. Continuing this process, we obtain

$$G = \bigcup_{i=1}^{k} g_iH \tag{2.4}$$

for some £nite k. If $H = \{e\}$, k is the number of elements of G; in general, however, k will be smaller.

Each set g_iH has $|H|$ elements. Thus, $|G| = k|H|$, proving $|H|$ divides $|G|$.

2.4 Introduction to Riemann zeta function

2.4.1 Prelude: Euclid's proof of in£nity of primes

Given the set of primes $p_1 \cdots p_n$ you can always construct the number $\prod_{i=1}^{n} p_i + 1$ which is indivisible by any of the given $p_1 \cdots p_n$. Therefore this number must be divisible only by primes greater than p_n, or must be prime itself. Therefore there exists a prime greater than p_n. An analysis of this proof gives a very weak lower bound on the number of primes less than x. The worst case scenario is that $\prod_{i=1}^{n} p_i + 1$ is the next prime. Thus, if we had $n-1$ primes up to $x = \prod_{i=1}^{n-1} p_i + 1$, we would have n primes up to

$$\prod_{i=1}^{n-1} p_i \cdot \Big(\prod_{j=1}^{n} p_j + 1 \Big) + 1 \tag{2.5}$$

Thus, having at least $n-1$ primes less than x, we have at least n primes less than (basically) x^2. One can quantify this further. One should get something like there are at least n primes less than 2^n or 4^n.

The zeta function will give us other ways to prove this, and to get a better estimate on the *prime counting function*,

$$\pi(x) \ \equiv \ \#\{p < x\}, \tag{2.6}$$

giving the number of primes below any number x.

2.4.2 De£nition, two forms

The Riemann zeta function $\zeta(s)$ is de£ned, for Re$(s) > 1$, by

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}. \tag{2.7}$$

We prove the useful fact that, for Re$(s) > 1$,

$$\sum_{n=1}^{\infty} \frac{1}{n^s} \ = \ \zeta(s) \ = \ \prod_{\substack{\text{primes} \\ p}} \left(1 - \frac{1}{p^s}\right)^{-1}, \tag{2.8}$$

which we call LHS and RHS. We call the product over primes an Euler Product.

To show equivalence, we use the Fundamental Theorem of Algebra (FTA) that all positive integers can be expressed as a single, unique, product of prime factors. Expanding all reciprocals in the RHS using the geometric series sum formula $(1-x)^{-1} = 1 + x + x^2 + x^3 + \cdots$, gives for the RHS,

$$(1 + 2^{-s} + 2^{-2s} + 2^{-3s} + \cdots)(1 + 3^{-s} + \cdots)(1 + 5^{-s} + \cdots) \cdots .$$

Remarkably, due to the FTA, we can associate 1-to-1 each term (choice of prime factors) on the RHS with each n on the LHS. For instance, $n = 12$ from LHS is accounted for by the RHS term,

$$2^{-2s} \cdot 3^{-s} \cdot 1 \cdot 1 \cdot 1 \cdots = \frac{1}{(2^2 \cdot 3)^s} = \frac{1}{12}.$$

Each combination of RHS terms corresponds uniquely to a single n.

2.4.3 $\zeta(s)$'s Behaviour and the In£nitude of Primes

We take the limit of s going to 1 and compare sides. LHS gives

$$\lim_{s \to 1} \zeta(s) \;=\; \lim_{s \to 1} \sum_{n=1}^{\infty} \frac{1}{n^s} = \sum_{n=1}^{\infty} \frac{1}{n}. \tag{2.9}$$

This sum diverges. Why? Crudely, $\sum_{n=1}^{N} n^{-1}$ is close to $\int_1^N \frac{dy}{y}$ which equals $\log N$. The de£nition of 'close' can be tightened up. For instance, you can create upper and lower bounds by approximating the integral by rectangular strips, getting

$$\sum_{n=2}^{N} \frac{1}{n} \leq \int_1^N \frac{dy}{y} \leq \sum_{n=1}^{N-1} \frac{1}{n}. \tag{2.10}$$

As the two sums differ by a bounded amount, we see the sum grows like $\log N$.

As s goes to 1, if there are only £nitely many primes than the product over primes is well behaved (ie, £nite). Therefore, there must be in£nitely many primes!

Further study of the zeta function will lead us to a good estimate for $\pi(x)$.

A second proof follows from the fact that $\zeta(2m) =$ rational $\cdot \pi^{2m}$, for integer m. This is known as a *Special Values* proof, as we are using the value of $\zeta(s)$ at a special value. We need the fact that π^2 is irrational. $\zeta(2) = \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$, which is irrational. Thus, the right hand side (the product over primes) must also be irrational; however, if there are only £nitely many primes, when $s = 2$ the right hand side is rational! Thus, there must be in£nitely many primes.

Please see Steve's notes, and URLs for more information on all of the above.

Chapter 3

Legendre Symbols, Gary Miller's Primality Test and GRH

We review the Legendre Symbol. We discuss Gary Miller's primality test, and show that if the General Riemann Hypothesis (for Dirichlet Characters) is true, then Miller's test correctly determines if a number n is prime or composite, and runs in time $O(\log^4 n)$. Lecture by Peter Sarnak; notes by Steven J. Miller.

3.1 Review of the Legendre Symbol

Recall Fermat's Little Theorem: $a^{p-1} \equiv 1 \bmod p$ if p is prime.

Given n, check $a^{n-1} \equiv 1 \bmod n$ for many a's relatively prime to n. **Exercise:** If ever this is not satis£ed, then n *must* be composite.

There are composite numbers (called Carmichael numbers) which satisfy $a^{n-1} \equiv 1 \bmod n$ for all a, yet are not prime. The £rst Carmichael number is $561 = 3 \cdot 11 \cdot 17$; the third is $1729 = 7 \cdot 13 \cdot 19$. **Exercise:** Prove all Carmichael numbers *must* be square-free.

Aside: 1729 has an interesting history (Ramanujan, Hardy and taxicabs). Hardy visited Ramanujan in the hospital, Hardy remarked that his taxicab's number was particularly uninteresting; Ramanujan remarks it (1729) is the smallest number which can be written in two different ways as the sum of two cubes. $1729 = 1^3 + 12^3 = 9^3 + 10^3$.

Recall the **Legendre symbol** $\left(\frac{a}{p}\right)$ is 0 if $p|a$, 1 if there is an $x \neq 0$ with $x^2 \equiv a \bmod p$, and -1 if there is no solution to $x^2 \equiv a \bmod p$. **Euler's condition** is $\left(\frac{a}{p}\right) \equiv a^{\frac{p-1}{2}} \bmod p$.

Really, the Legendre symbol is a function on $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$. We can extend the Legendre symbol to all integers. We only need to know a mod p, and we de£ne $\left(\frac{a}{p}\right) = \left(\frac{a \bmod p}{p}\right)$.

Initially the Legendre symbol is de£ne only when the bottom is prime. We now extend the de£nition to all n as follows: let $n = p_1 \cdot p_2 \cdots p_t$ be the product of t distinct primes. Then $\left(\frac{a}{n}\right) = \left(\frac{a}{p_1}\right)\left(\frac{a}{p_2}\right)\cdots\left(\frac{a}{p_t}\right)$. Note this is *not* the same as saying that if a is a square (a quadratic residue) mod n, then a is a square mod p_i for each prime divisor.

The main result (which allows us to calculate the Legendre symbol quickly and ef£ciently) is the celebrated

Theorem 3.1.1 (Quadratic Reciprocity). *For m, n odd and relatively prime, $\left(\frac{m}{n}\right)\left(\frac{n}{m}\right) = (-1)^{\frac{m-1}{2}\frac{n-1}{2}}$.*

3.2 Gary Miller's Primality Test, 1976

Miller Test: *Given n as input (n must be odd), test whether for $2 \leq a \leq 70 \log^2 n$, $\left(\frac{a}{n}\right) \equiv a^{\frac{n-1}{2}}$ mod n (where, of course, a and n are relatively prime). If this test fails for some a in this range, output* composite*; if it passes the test for all such a, output* Prime.

Note that we can very quickly determine if two number are relatively prime (use the Euclidean Algorithm, which takes $O(\log n)$ steps).

Theorem 3.2.1 (Miller Test Results). *The Miller Test runs in $O(\log^4 n)$ steps. If the output is* composite*, then the number n is composite (ie, the algorithm's result is correct). If we assume GRH (the* General Riemann Hypothesis*, the most important unsolved problem in mathematics), then the output* prime *is also correct.*

Running time: we can compute $\left(\frac{a}{n}\right)$ in $O(\log n)$ steps. By Quadratic Reciprocity, to compute $\left(\frac{a}{n}\right)$ (may assume $-\frac{n}{2} \leq a \leq \frac{n}{2}$), it is enough to compute $\left(\frac{n}{a}\right)$ (with a factor of -1).

Why? We only need to know a mod n, so we may reduce a until $-\frac{n}{2} \leq a \leq \frac{n}{2}$. Note the top (in absolute value) is at most half the size of the bottom (n). We then use quadratic reciprocity to evaluate $\left(\frac{a}{n}\right)$. Up to a factor of -1, $\left(\frac{a}{n}\right) = \left(\frac{n}{a}\right)$. Thus, we may reduce n mod a, so that n mod a lies between $-\frac{a}{2}$ and $\frac{a}{2}$. Again, the top is half the bottom, and the bottom is at most one-quarter of what we started with (n).

We continue this process; we need to do such ¤ippings at most $\log n$ times. Why? Each time the size of the denominator is at most half what it was before. If $2^r = n$, then $r = \log_2 n < \log n$. Thus, after at most $\log n$ passes, the denominator would be about 1. So, a stage or two before would give a denominator around 2 or 4. The point is, in $\log n$ steps, we can reduce to evaluating the Legendre symbol of something where the bottom is of bounded size. Thus, we can evaluate $\left(\frac{a}{n}\right)$ in $O(\log n)$ steps. (Have a lookup table for $n < 10$, et cetera).

We need to evaluate $\left(\frac{a}{n}\right)$ for $C \log^2 n$ choices of a, and for each choice we need to evaluate $a^{\frac{n-1}{2}} \bmod n$ (so we can compare it to $\left(\frac{a}{n}\right)$), which takes $O(\log n)$ steps. Thus, the number of steps is $O(\log^4 n)$.

The Riemann Hypothesis plays a very important catalystic role. It leads us to statements we feel should be true, statements which can often be proved without the full force of Riemann.

Suppose we pass the test for all a with $2 \leq a \leq 70 \log^2 n$ and a relatively prime to n. Gary Miller proved that, if GRH is true, then knowing that n passes the test for all a in this little segment allows us to conclude that n will pass the test *for all* a. Clearly, we know if n passes the test for all a up to $n-1$, then we know n will pass the test for all a relatively prime to n. The power of GRH is that we need only check $\log^2 n$ values of a.

3.2.1 Aside: Finite Abelian Groups

Let A be a £nite Abelian Group, then A is (ie, is isomorphic to) a product of cyclic groups.

Look at $(\mathbb{Z}/n\mathbb{Z}, +)$, the group of integers mod n under clock addition. This is a cyclic group.

The general statement is:

Theorem 3.2.2 (Structure Theorem for Finite Abelian Groups). *Let A be a £-nite Abelian Group. Then there are integers n_1 through n_t such that $A \approxeq \mathbb{Z}/n_1\mathbb{Z} \times \cdots \times \mathbb{Z}/n_r\mathbb{Z}$, ie, A is isomorphic to the Cartesian product of groups of the form $\mathbb{Z}/m\mathbb{Z}$.*

Note $\approxeq$ means is isomorphic to; we say two groups are **isomorphic** if there is a group homomorphism between them which is one-to-one and onto. A **group homomorphism** ϕ is a map which preserves the group structure. Consider two groups G_1 and G_2 and a map $\phi : G_1 \to G_2$. If ϕ is a group homomorphism, then for $x, y \in G_1$, $\phi(x+y) = \phi(x) \bigoplus \phi(y)$, where $+$ is addition in the £rst group and $\bigoplus$ is addition in the second group.

We now de£ne the **Cartesian Product** of two groups. $X \times Y$ is the set of pairs (x, y) where $x \in X$ and $y \in Y$. If we are writing the group action of X and Y additively (say by $\bigoplus_X$ for addition in X and $\bigoplus_Y$ for addition in Y), then $(x_1, y_1) \bigoplus_{X+Y} (x_2, y_2) = (x_1 \bigoplus_X x_2, y_1 \bigoplus_Y y_2)$.

Let us consider a group written in additive notation. Recall the **order of an element** is the number of times you must add the element to itself to get the identity. We de£ne the **exponent** of a group as the least common multiple of the orders of the elements of the group.

Consider $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, the product of two groups. In $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, we have the pairs $(0, 0)$, $(0, 1)$, $(1, 0)$ and $(1, 1)$. $(0, 0)$ is the identity under addition; all other elements (check) have order 2. Thus, the exponent of $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ is 2.

Now consider $\mathbb{Z}/4Z = \{0, 1, 2, 3\}$. 0 is the identity, and under addition 1 and 3 have order 4, and 2 has order 2. Thus, the exponent of this group is 4. **Exercise:** using the exponent, observe that $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ and $\mathbb{Z}/4Z$ cannot be isomorphic.

Fact: if p is prime, then $\mathbb{Z}/p\mathbb{Z} = \mathbb{F}_p$ is a £eld. The non-zero elements have multiplicative inverses. $(\mathbb{Z}/p\mathbb{Z})^* = \mathbb{F}_p^* = \mathbb{F}_p - \{0\}$ (the non-zero elements under multiplication) is a cyclic abelian group with $p - 1$ elements! IE, there is an element g whose powers generate the group. **Exercise:** Prove that if n is composite, then $\mathbb{Z}/n\mathbb{Z}$ is not a £eld. Hint: show some non-zero element has no multiplicative inverse.

Sketch of Proof of Fact: Suppose $\mathbb{F}_p^*$ is not cyclic. Let d be the exponent (the least common multiple of the orders of the elements) of $\mathbb{F}_p^*$. Then $d < p - 1$. As the order of every element divides $p - 1$ (Lagrange's Theorem), we have $d | p - 1$.

For each $x \in Fpf$, $x^d = 1$ (in the £eld $\mathbb{F}_p^*$). This is because d is the least common multiple of the orders of the elements of the group. Thus, if ord(x) is the order of x, $x^{\text{ord}(x)} = 1$. As ord$(x) | d$, we have ord$(x) = kd$ for some $k \in \mathbb{Z}$. Thus, $x^d = x^{k\text{ord}(x)} = (x^{\text{ord}(x)})^k = 1^k = 1$.

Now use the fact that $\mathbb{F}_p^*$ is a £eld. Consider $x^d - 1 = 0$ over $\mathbb{F}_p$ (clearly 0 is not a root). **Over** $\mathbb{F}_p$ means look for solutions to this equation with $x \in \mathbb{F}_p$). A polynomial of degree d has at most d roots (another theorem of Lagrange). But every $x \in \mathbb{F}_p^* = \mathbb{F}_p - \{0\}$ is a root, because d (the exponent of the group) is the least common multiple of the orders of the elements. This is a contradiction: we have $p - 1$ roots (every $x \in \mathbb{F}_p^*$ solves $x^d - 1 \equiv 0 \bmod p$), but by Lagrange there are at most d roots. As $d < p - 1$, contradiction.

Steve Miller will discuss this needed theorem of Lagrange. See also the handout from Davenport's *The Higher Arithmetic*.

3.2.2 Lehmer's Proposition

Proposition 3.2.3 (Lehmer). *If* $\left(\frac{a}{n}\right) \equiv a^{\frac{n-1}{2}}$ *mod* n *for all* a *up to* $C \log^2 n$*, then* n *is prime.*

We assume $n = p_1 \cdots p_t$, $t > 1$, and all primes are distinct and odd. (When there are repeated primes, the proof is easier, and is left as an exercise to the reader). Then $a^{n-1} \equiv 1$ mod n for $(a, n) = 1$ (ie, a and n relatively prime), implies that $a^{n-1} \equiv 1$ mod p_j (p_j one of the prime factors of n).

Thus, $n - 1 \equiv 0$ mod $p_j - 1$. If there were a remainder (ie, if $n - 1$ wasn't equivalent to 0 mod $p_j - 1$ but was equivalent to $r_j \not\equiv 0$ mod $p_j - 1$), if we raised a to this power (r_j), we wouldn't get 1 for all a. Just take a a generator of $\mathbb{F}_{p_j}^*$, ie, an element of maximal order $p_j - 1$. Then $a^{n-1} \equiv a^{r_j} \neq 1$ mod p_j. Note that $a^{p_j-1} \equiv 1$ mod p_j, so $a^{n-1} \equiv a^{r_j}$ mod p_j.

We say p_j (a factor of n) is **of type** 1 if $\frac{n-1}{2} \equiv 0$ mod $p_j - 1$; we say p_j is **of type** 2 if $\frac{n-1}{2} \equiv \frac{p_j-1}{2}$ mod $p_j - 1$.

If at least one of the p_j's (without loss of generality, say p_1) is of type 1, take a a quadratic non-residue mod p_1 and a quadratic residue for $p_2, p_3, \ldots, p_t$. One can £nd such an a by the **Chinese Remainder Theorem** as the primes are distinct. For a statement of the Chinese Remainder Theorem, see the Appendix at the end of the notes.

We have $\left(\frac{a}{p_1}\right) = -1$ and $\left(\frac{a}{p_j}\right) = 1$ for $j > 1$. As we are assuming n is composite, there are at least two primes.

As $\left(\frac{a}{n}\right) = \left(\frac{a}{p_1}\right)\left(\frac{a}{p_2}\right) \cdots \left(\frac{a}{p_t}\right)$, we obtain $\left(\frac{a}{n}\right) = -1 \cdot 1 \cdots 1$.

As we are assuming p_1 is of type 1, we have $\frac{n-1}{2} \equiv 0$ mod $p_1 - 1$. Mod $p_1 - 1$, each non-zero element in $(\mathbb{F}/p_1\mathbb{F})^*$ has order dividing $p_1 - 1$. Thus, $a^{\frac{n-1}{2}} \equiv 1$ mod p_1.

We are assuming that the Miller Test is satis£ed for all a. Thus, $\left(\frac{a}{n}\right) = a^{\frac{n-1}{2}}$ mod n. Mod p_1, we have shown the left hand side is -1 and the right hand side is 1, contradiction!

We are left with the case where all the primes are of type 2. We leave this as an exercise for the reader.

3.3 GRH Implies Just Need To Check Up To $O(\log^2 n)$ in Miller

Why does GRH imply that $\left(\frac{a}{n}\right) \equiv a^{\frac{n-1}{2}} \bmod n$ for all a up to $70\log^2 n$ (and relatively prime to n) implies that $\left(\frac{a}{n}\right) \equiv a^{\frac{n-1}{2}} \bmod n$ for all a relatively prime to n?

3.3.1 Fourier Analysis

Representation Theory (especially characters of Abelian Groups) is the most important things in Mathematics. Let A be a £nite Abelian Group. To each such group, we associate a dual group, denoted $\hat{A}$, where $\hat{A}$ is the set of all homomorphisms from A into $\mathbb{C}^*$, $\mathbb{C}^*$ are the complex numbers invertible under multiplication.

Recall ψ is a *group homomorphism* if $\psi(a+b) = \psi(a)\psi(b)$. Note here $\psi : G_1 \to G_2$, and G_1 has been written additively and G_2 has been written multiplicatively. We call such a ψ a *character*.

3.3.2 Examples

Let $A = \mathbb{Z}/n\mathbb{Z}$, and $\nu \in \mathbb{Z}/n\mathbb{Z}$. Let $e(z) = e^{2\pi i z}$ for $z \in \mathbb{C}$. De£ne $\psi_\nu(x) = e(\frac{\nu x}{n})$. **Exercise:** Show by direct calculation that $\psi_\nu(a+b) = \psi_\nu(a)\psi_\nu(b)$. In fact, for each $\nu \in \mathbb{Z}/n\mathbb{Z}$ we get a character, and the characters are distinct (exercise). We only need to know how ψ_ν acts on 1, as $\psi_\nu(k) = \psi_\nu(1+\cdots+1) = \Big(\psi_\nu(1)\Big)^k$. $\hat{A}$ is canonically isomorphic to A. What this means is, to each $\nu \in A$ there corresponds a character ψ_ν in $\hat{A}$, and to each character $\psi \in \hat{A}$ there corresponds a number $\nu_\psi \in A = \mathbb{Z}/n\mathbb{Z}$.

We can multiply two characters: $(\psi_1\psi_2)(x) = \psi_1(x)\psi_2(x)$. It is easy to see that this is a character. The trivial character (which sends everything to 1) is the identity of the group $\hat{A}$.

Consider two groups A and B. We can use the characters of A and B to get the characters of the cartesian product $A \times B$. If we want the characters of $A \times B$, take a character ψ_a of A and ϕ_b of B and form the character $\psi_a\phi_b$, de£ned by $(\psi_a\phi_b)(x,y) = \psi_a(x)\phi_b(y)$.

3.3.3 Characters of $\mathbb{F}_p^*$: Dirichlet Characters

We denote characters of $\mathbb{F}_p^*$ by χ. Dirichlet proved the best theorem of mathematics, introducing a lot of new math to solve the following:

Theorem 3.3.1 (Primes in Arithmetic Progressions, 1836**,** 1839**).** *Let a and n be relatively prime. There are in£nitely many primes which give a when divided by n; moreover, to £rst order all residue classes a mod n (a, n relatively prime) have the same number of primes!*

In other words, there are in£nitely many x such that $xn + a$ is prime if a and n are relatively prime.

Hard Question: without using Dirichlet's Theorem, can you prove that there must be *one* prime congruent to a mod n if a and n are relatively prime?

Clearly, if a and n are not relatively prime, there cannot be in£nitely many primes congruent to a mod n. Dirichlet shows this is also a suf£cient condition.

Dirichlet introduced characters *without* introducing the concept of a group! They didn't have group notation until later.

Look at $(\mathbb{Z}/n\mathbb{Z})^* = \{x : (x,n) = 1, 0 < x \leq n-1\}$. This is a £nite Abelian group. $\#(\mathbb{Z}/n\mathbb{Z})^* = \phi(n)$, where $\phi(n)$ is the Euler totient function, and $\phi(n)$ is the number of integers (between 0 and n) which are relatively prime to n.

Dirichlet used q instead of n, so we change notation and look at $(\mathbb{Z}/q\mathbb{Z})^*$.

A **Dirichlet Character** is a character χ of $(\mathbb{Z}/q\mathbb{Z})^*$; ie, $\chi : (\mathbb{Z}/q\mathbb{Z})^* \to \mathbb{C}^*$ and $\chi(ab) = \chi(a)\chi(b)$. Thus, χ lies in the dual group of $(\mathbb{Z}/q\mathbb{Z})^*$. Recall that $\mathbb{C}^*$ is the set of complex numbers with multiplicative inverses, ie, $\mathbb{C}^* = \mathbb{C} - \{0\}$.

The **principal or trivial character** takes every $x \in (\mathbb{Z}/q\mathbb{Z})^*$ to 1; we denote the trivial character by χ_0.

If χ is a Dirichlet Character of $(\mathbb{Z}/q\mathbb{Z})^*$, we say χ has **modulus** (also called **conductor**) q. We extend χ to be de£ned on all integers (all of $\mathbb{Z}$) by $\chi(m) = 0$ if m and q are not relatively prime, and $\chi(m) = \chi(m \bmod q)$ otherwise. Clearly χ is periodic, as $\chi(x+\lambda q) = \chi(x)$ for any $\lambda \in \mathbb{Z}$. Thus, we have the map $\chi : \mathbb{Z} \to \mathbb{C}^*$.

If q is prime, we have previously seen the character $\chi(a) = \left(\frac{a}{q}\right)$. This is an extremely important character.

Another example: Let $q = 4$. We de£ne $\chi(n)$ to be 0 if n is even, 1 if $n \equiv 1$ mod 4, and -1 if $n \equiv -1$ mod 4. **Exercise:** Show this is a character.

3.3.4 General Riemann Hypothesis (GRH)

Below, p will always denote a prime. χ will be a Dirichlet character modulo q (q need not be prime).

Conjecture: GRH: *For any $q \geq 1$, $x \geq 3$, if $\chi \neq \chi_0$ (ie, if χ is not the principal character) then*

$$\Big| \sum_{p \leq x} \left(1 - \frac{p}{x}\right) \cdot \log p \cdot \chi(p) \Big| \leq C \log q \cdot \sqrt{x}. \tag{3.1}$$

If $\chi = \chi_0$, then

$$\sum_{p \leq x} \left(1 - \frac{p}{x}\right) \log p \leq \frac{x}{2} + O(\sqrt{x}). \tag{3.2}$$

C is a universal constant, independent of q.

We are summing primes up to x. The $1 - \frac{p}{x}$ factors are just weight factors. If $q = 4$ and χ is the character above, $\chi(p)$ gives a positive sign for primes congruent to $1 \bmod 4$ and a minus sign for primes congruent to $-1 \bmod 4$.

Analysis means cancellation; you want to see cancellation in a sum. The numbers in these sums are ¤ipping (plus one, minus one); we have about $\frac{x}{\log x}$ primes less than x (this is the Prime Number Theorem, £rst proved in 1896).

Random Drunk: each moment he ¤ips a coin. If it is heads he staggers one unit left, tails he staggers one unit right. After n steps, where do you think he'll be? Turns out he'll be about $\sqrt{n}$ units from the origin (with high probability). Steve Miller will prove this in a later lecture or handout.

Motto: random numbers (of size around 1) cancel like the square-root of the number of terms.

What GRH is telling us is that the remainder term behaves like random noise.

3.3.5 Proof of the Miller Test

Theorem 3.3.2. *If GRH is true, then if $\left(\frac{a}{n}\right) \equiv a^{\frac{n-1}{2}} \bmod n$ for all $a \leq C \log^2 n$ for some £xed constant C, then it is true for all a (relatively prime to n, of course).*

Remarkable observation: $S = \{a \in (\mathbb{Z}/n\mathbb{Z})^* : \left(\frac{a}{n}\right) \equiv a^{\frac{n-1}{2}} \bmod n\}$ is a subgroup of $(\mathbb{Z}/n\mathbb{Z})^*$. Why? Exercise (Closure: show if a, b in this set, so is ab. Identity: show 1 is in this set. Inverses: show if $a \in S$, $a^{-1} \bmod n$ is in S. Note associativity is inherited from $(\mathbb{Z}/n\mathbb{Z})^*$).

By assumption, S contains (at least) the £rst $C \log^2 n$ elements. We claim this implies S contains all a relatively prime to n.

Suppose S is not all of $(\mathbb{Z}/n\mathbb{Z})^*$. We talked about cosets (of Abelian Groups) last time. (For more information about cosets and quotient groups, see the following lecture by Steven J. Miller, notes by Alex Barnett). As everything is abelian, we can form a quotient group by dividing our group $(\mathbb{Z}/n\mathbb{Z})^*$ by the abelian subgroup S.

Thus, $A = \frac{(\mathbb{Z}/n\mathbb{Z})^*}{S}$ is a group (a quotient group) and A is a non-trivial group (ie, there is more than one element in this group). Why must A have more than one element? A is the group of representative cosets of $(\mathbb{Z}/n\mathbb{Z})^*$ by the abelian subgroup S. As S is not all of $(\mathbb{Z}/n\mathbb{Z})^*$, there must be at least two cosets. Hence A is not just the identity coset (which is $1 \cdot S$ or just S).

But this quotient (A) is a £nite abelian group. We know for a £nite abelian group that its dual is isomorphic to itself. This means to each element in A we have a character in $\hat{A}$, and vice-versa. Further, the identity of A is mapped to the trivial character of $\hat{A}$.

Thus, there is a non-trivial $\chi : A \to \mathbb{C}^*$. Now $(\mathbb{Z}/n\mathbb{Z})^* \to \frac{(\mathbb{Z}/n\mathbb{Z})^*}{S} \to \mathbb{C}^*$, and thus we have a non-trivial Dirichlet character of $(\mathbb{Z}/n\mathbb{Z})^*$ which is trivial on S (ie, $\chi(s) = 1$ for all $s \in S$).

By construction, χ is a Dirichlet character of $(\mathbb{Z}/n\mathbb{Z})^*$, $\chi \neq \chi_0$, and $\chi|_S = 1$ (the last means that χ, **restricted to** $s \in S$ is the identity map). We know S is all elements up to at least $C \log^2 n$. So, $\chi(a) = 1$ for all $a \leq C \log^2 n$ (from the givens).

So, look at

$$\text{Sum}(p, q, x) = \sum_{p \leq x} \left(1 - \frac{p}{x}\right) \log p \cdot \chi(p). \tag{3.3}$$

By GRH, $\text{Sum}(p, q, x)$ is less than a constant multiple of $\log q \cdot \sqrt{x}$. Calling the constant C_2, we have $\text{Sum}(p, q, x) \leq C_2 \log q \cdot \sqrt{x}$.

If $x \leq C \log^2 n$, then there is a contradiction. Why? In Equation 3.3, $\text{Sum}(p, q, x)$ is going to be $\sum_{p \leq x} \left(1 - \frac{p}{x}\right) \log p$, because all the $\chi(p) = 1$ in this range. The **Prime Number Theorem** states that the number of primes less than x is $\frac{x}{\log x}$ plus an error term which is smaller than $\frac{x}{\log x}$ (ie, in the limit, the size of the error term divided by the number of primes less than x tends to 0 as x goes to in£nity).

Thus,$\text{Sum}(p, q, x)$ is going to look like a multiple of x. Why a multiple of x and not a multiple of $\frac{x}{\log x}$? Remember we have the factor $\log p$ in the sum;

this follows from **Partial Summation**. Partial Summation is a discrete version of Integration by Parts; see the Appendices at the end for a further statement.

By GRH, Sum(p,q,x) is bounded by $C_2 \log q \cdot \sqrt{x}$. Therefore, $x \leq C_2 \log q \cdot \sqrt{x}$. This implies $\sqrt{x} \leq C_2 \log q$, or $x \leq C_2^2 \log^2 q$.

Remember, we've changed notation from n to q. We are assuming that $\left(\frac{a}{q}\right) \equiv a^{\frac{q-1}{2}}$ mod q for all $a \leq C \log^2 q$ for some £xed constant C. Take C greater than C_2^2. Then we have a contradiction! If GRH is true, Sum$(p,q,x) \leq C_2 \log q \cdot \sqrt{x}$ only for $x \leq C_2^2 \log^2 q$. But we are assuming that q passes the Miller Test for all a up to $C \log^2 q$. Thus, we can take x larger than $C_2^2 \log^2 q$.

3.3.6 Review of Proof

Intuition: For random sequences, expect square-root cancellation in sums.

If we have a non-trivial character, if we look at these weighted sums of $\chi(p)$, there is no extra structure; we expect cancellation like $\sqrt{x}$. There has to be *some* q-dependence, but GRH says it is like a universal constant times $\log q$.

In the Miller Test, we test something $C \log^2 n$ times. If the condition is always true for these $C \log^2 n$ elements, we have a subgroup S of $(\mathbb{Z}/n\mathbb{Z})^*$. If S isn't all of $(\mathbb{Z}/n\mathbb{Z})^*$, we can £nd a character, and we have sums with this character (any sum with characters is called **Harmonic Analysis**). Further, this is a non-trivial character which is the identity on the original group. The GRH cannot accomodate a character of this type.

Why does the GRH lead to a contradiction? Basically, GRH says a certain weighted sum of $\chi(p)$ over the primes less than x (where χ is a Dirichlet character with modulus q) cannot be too large. Speci£cally, it is at most $C_2 \log q \cdot \sqrt{x}$.

This implies that there is a lot of noise in the $\chi(p)$; basically, we need to have a good mixing of primes which give $\chi(p) = +1$ with primes giving $\chi(p) = -1$.

However, if n satis£es the Miller Test for a up to $C \log^2 q$ (with $C > C_2^2$), then we can £nd a modulus q and a Dirichlet character χ where we have a very long string of primes giving $\chi(p) = +1$. This forces the weighted sum of $\chi(p)$ over primes less than x (taking $x = C \log^2 q$) to be larger than $C_2 \log q \sqrt{x}$, a contradiction.

3.4 Appendices

3.4.1 Aside: For p Odd, Half the Non-Zero Numbers are Quadratic Residues

Note, for an odd prime p_j, half of the non-zero numbers are quadratic residues, and half are quadratic non-residues. The **Legendre symbol** takes each element $a \in \mathbb{F}_p^*$ to an element in the group $\{-1, 1\}$. This is a homomorphism; not every element has a square. The image is $\{-1, 1\}$; the kernel is all the elements of $\mathbb{F}_p^*$ which are sent to 1. Thus, half the numbers are residues, half are non-residues.

$a \to \left(\frac{a}{p}\right) = a^{\frac{p-1}{2}}$ mod p. Thus, we have a homomorphism (given by the Legendre symbol) from $\mathbb{F}_p^* \to \{-1, 1\}$. We claim the map is onto. The kernel is all elements in $\mathbb{F}_p^*$ which are mapped by the Legendre symbol to 1, ie, the quadratic residues. (One needs to show the Legendre symbol is a group homomorphism: $\left(\frac{xy}{p}\right) = \left(\frac{x}{p}\right) \cdot \left(\frac{y}{p}\right)$).

Standard Group Theory Arguments: $\frac{\#\mathbb{F}_p^*}{\text{kernel}} \approxeq \{-1, 1\}$. Thus, half the numbers in $\mathbb{F}_p^*$ are quadratic residues.

3.4.2 Chinese Remainder Theorem

Theorem 3.4.1 (Chinese Remainder Theorem). *Let $m = m_1 m_2$, m_1 and m_2 relatively prime. Then $\mathbb{Z}/m\mathbb{Z} \approxeq \mathbb{Z}/m_1\mathbb{Z} \times \mathbb{Z}/m_2\mathbb{Z}$.*

This allows us to, given a_1 mod m_1 and a_2 mod m_2, £nd an a mod m such that $a \equiv a_1$ mod m_1 and $a \equiv a_2$ mod m_2. For example, try and solve $x \equiv 3$ mod 5 and $x \equiv 4$ mod 7.

See any book on Algebra.

3.4.3 Partial Summation

Lemma 3.4.2 (Partial Summation: Discrete Version).

$$\sum_{M}^{N} a_n b_n = A_N b_N - A_{M-1} b_M + \sum_{M}^{N-1} A_n (b_n - b_{n+1}) \tag{3.4}$$

Lemma 3.4.3 (Abel's Summation Formula - Integral Version). *Let $h(x)$ be a continuously differentiable function. Let $A(x) = \sum_{n \leq x} a_n$. Then*

$$\sum_{n \leq x} a_n h(n) = A(x)h(x) - \int_1^x A(u)h'(u)du \tag{3.5}$$

See, for example, Walter Rudin, *Principles of Mathematical Analysis* (also known as *The Blue Book*), page 70.

Chapter 4

Cosets, Quotient Groups, and an Introduction to Probability

Quotient groups. Basic probability theory for random walk. Lecture by Steven J. Miller; notes by Alex Barnett and Steven J. Miller.

4.1 Quotient groups

Say we have a £nite Abelian group G (this means for all $x, y \in G$, $xy = yx$) of order m which has a subgroup H of order r. We will use multiplication as our group operation. Recall the *coset* of an element $g \in G$ is de£ned as the set of elements $gH = g\{h_1, h_2, \cdots, h_r\}$. Since G is Abelian (commutative) then $gH = Hg$ and we will make no distinction between left and right cosets here.

The *quotient group* (or *factor group*), symbolized by G/H, is the group formed from the cosets of all elements $g \in G$. We treat each coset g_iH as an element, and de£ne the multiplication operation as usual as g_iHg_jH. Why do we need G to be Abelian? The reason is we can then analyze g_iHg_jH, seeing that it equals g_ig_jHH. We will analyze this further when we prove that the set of cosets is a group.

There are several important facts to note. First, if G is not Abelian, then the set of cosets might not be a group. Second, recall we proved the coset decomposition rule: given a £nite group G (with n elements) and a subgroup H (with r elements) then there exist elements g_1 through g_k such that

$$G = \bigcup_{i=1}^{k} g_i H. \tag{4.1}$$

The choices for the g_i's is clearly not unique. If g_1 through g_k work, so do $g_1 h_1$ through $g_k h_k$, where h_i is any element of H. Recall this was proved by showing any two cosets are either distinct or identical.

We will show below that, for G Abelian, the set of cosets is a group. Note, however, that while it might at £rst appear that there are many different ways to write the coset group, they really are the same. For example, the cosets gH and $gh_1 h_2^4 h_3 H$ are equal. This is similar to looking at integers mod n; mod 12, the integers 5, -7 and 19 are all equal, even though they look different.

We now prove that the set of cosets is a group (for G Abelian).

Closure. By commutivity $g_i H g_j H = g_i g_j HH$. What is "HH"? Just the set of all r^2 possible combinations of elements of H. By closure, and the existence of the identity, this just gives H again (recall no element in a group can appear more than once—duplicates are removed). Therefore $g_i H g_j H = g_i g_j H$. Now, as G is a group and is closed, $g_i g_j \in G$. Thus, there is a α such that $g_i g_j \in g_\alpha H$ (as $G = \bigcup_{\beta=1}^{k} g_\beta H$. Therefore, there is an $h \in H$ such that $g_i g_j = g_\alpha h$, which implies $g_i g_j H = g_\alpha h H = g_\alpha H$. Thus, the set of cosets is closed under coset multiplication. Note, however, that while the coset $g_i g_j H$ is in our set of cosets, it may be written differently.

Identity. If e is identity of G, then $eHg_i H = g_i H$ and $g_i H e H = g_i H$, so eH is the identity of this quotient group.

Associativity. Since as you may have noticed, the quotient group elements behave just like those of G, associativity follows from that of G.

Inverse. It is easy to guess $g^{-1}H$ is the inverse of gH. Check it: $g^{-1}HgH = g^{-1}gH = eH =$ identity, also true the other way round of course by commutativity. Unfortunately, $g^{-1}H$ might not be listed as one of our cosets! Thus, we must be a little more careful. Fortunately, as $g^{-1} \in G = \bigcup_{\beta=1}^{k} g_\beta H$, there is an α such that $g^{-1} \in g_\alpha H$. Then, there is an $h \in H$ with $g^{-1} = g_\alpha h$. Thus, $g^{-1} = g_\alpha h H = g_\alpha H$, and direct calculation will show that the coset $g_\alpha H$ is the inverse (under coset multiplication) of gH.

4.2 Random walk and discrete probability

Each step in a random walk is a random event. We £rst study a single random event, then the combination of two random events, then multiply repeated events.

For other introductory probability theory see
`http://engineering.uow.edu.au/Courses/Stats/File24.html`

4.2.1 Probability distribution for a single event, mean, variance

Our single event consists of one of a set of choices happening. The choices are labelled by $i = 1 \cdots N$, which are exclusive (no more than one can happen), and complete (no less than one can happen). For instance, for a single coin toss,

$$
\begin{aligned}
i &= 1 \qquad\qquad H, \text{ heads,} \\
i &= 2 \qquad\qquad T, \text{ tails.}
\end{aligned}
$$

We take the choice i as a *random variable*, meaning all we know is a probability $p_i \geq 0$ that each choice can happen. $p = 0$ means it never happens, $p = 1$ means it always happens, and most things are somewhere in between (the unbiased coin has $p_i = \frac{1}{2}$, $\forall i$, ignoring the small probabilities of the coin landing on its edge or quantum tunneling through the table.)

The set of $\{p_i\}$ we call the *probability distribution*. Completeness implies

$$
\sum_i p_i = 1. \tag{4.2}
$$

Note the abbreviation $\sum_i$ for $\sum_{i=1}^N$.

We have some quantity f which has a value f_i for each choice i. For instance, f could be the number of dots on each face i of a die, in which case $f_i = i$. *Mean* and *variance* are ways to characterize (summarize) the distribution over f.

Mean. The mean or *expectation value* (expected value) of f over the distribution is

$$
\bar{f} \equiv E[f] = \sum_i p_i f_i. \tag{4.3}
$$

This is just a weighted average of f. Check it for the (unweighted) dice. $p_i = \frac{1}{6}$, $\forall i$. You should get $\bar{f} = \frac{7}{2}$.

Variance. The variance is the square of the *standard deviation* σ_f:

$$\sigma_f^2 \equiv \mathrm{Var}[f] \;=\; \sum_i p_i(f_i - \bar{f})^2. \tag{4.4}$$

Crudely σ_f gives the width of the distribution in f. From the de£nition you can see σ_f is the *root mean square* (rms) deviation from the mean. Expanding out the square, and using earlier results,

$$\begin{aligned} \mathrm{Var}[f] &= \sum_i p_i f_i^2 - 2\bar{f}\cdot\sum_i p_i f_i + \bar{f}^2\cdot\sum_i p_i \\ &= \sum_i p_i f_i^2 - \bar{f}^2 \\ &= E[f^2] - E[f]^2. \end{aligned} \tag{4.5}$$

This is a very useful formula. The £rst term is often known as the *second moment* of f. The *£rst moment* is just the mean. The m^{th} *moment* is de£ned as $\sum_i p_i f_i^m$.

A quick comment about units. If f_i and f are measured in feet (for example), the variance (which gives information on how spread out f_i is) has units feet-squared. If someone asks how tall the people in our class are, one would answer about $5\frac{1}{2}$ feet. If one is further pressed to give a range for the heights of our class, one might say $5\frac{1}{2}$ feet, plus or minus $\frac{1}{4}$ of a foot. One would not give the error range in feet-squared! Thus, in measuring error it is the square-root of the variance that comes into play. Note that if f_i is in feet, the variance is in feet-squared, and the square-root of the variance is in feet.

4.2.2 Multiple events

Consider two random events. Let the £rst have choices $i = 1\cdots N$ (with probabilities p_1 through p_N); let the second event have choices $j = 1\cdots M$ (with probabilities q_1 through q_M). Then there are NM choices (possibilities) for the combined event, which we could label by $k \equiv ij$. Since k is also a random variable, it has probability $r_k = r_{ij}$ (you could think of this as a matrix in i, j). If the two events are *independent* (also called *uncorrelated*), then

$$r_{ij} \;=\; p_i q_j, \;\; \forall ij \qquad\qquad \text{independence.} \tag{4.6}$$

(In other words, the matrix is separable in the i and j directions). Many events we study will be independent. For example, if you ¤ip a fair coin twice, the result of the £rst ¤ip has no effect on the result of the second ¤ip. Or if you roll a fair die, et cetera.

As before we have a quantity f_i associated with each choice i for the £rst event. For the second event we have a (in general different) quantity g_j for its choice j. We want to learn about the sum of these two quantities,

$$s \equiv f + g. \tag{4.7}$$

Note that for each combined choice ij, this quantity s has value $s_{ij} = f_i + g_j$. How is s is distributed? We will show that *independence* of the two events implies a simple law giving the mean and variance of s in term of those of f and g.

We compute the mean of s, following Eq. 4.3, except now we are summing over all combined possibilities ij,

$$\begin{aligned} E[s] &= \sum_{ij} r_{ij} s_{ij} = \sum_{ij} p_i q_j (f_i + g_j) = \sum_i p_i f_i \cdot \sum_j q_j + \sum_i p_i \cdot \sum_j q_j g_j \\ &= \bar{f} \cdot 1 + 1 \cdot \bar{g} = \bar{f} + \bar{g}. \end{aligned} \tag{4.8}$$

So, the means *add*.

We isolate this important fact:

Lemma 4.2.1. *For independent events, the mean of a sum is the sum of the means. Equivalently, the sum of the expected values is the expected value of the sums. Thus, for any independent events A and B, $E[A + B] = E[A] + E[B]$.*

What if we multiply A by a constant c? For example, consider outcomes A_i with probabilities p_i. The mean $\bar{A} = E[A]$. What is the mean of the new event with outcomes cA_i occurring with probabilities p_i?

Well,

$$\begin{aligned} E[cA] &= \sum_i p_i c A_i \\ &= c \sum_i p_i A_i \\ &= cE[A]. \end{aligned} \tag{4.9}$$

We have therefore shown

Lemma 4.2.2. *The mean of a multiple is the multiple of the mean. Equivalently, the expected value of a multiple is the multiple of the expected value. Thus, $E[cA] = cE[A]$.*

We now calculate the variance of a sum, using the above results. For the variance of $s = f + g$ we can use Eq. 4.5 to get

$$\begin{aligned}
\text{Var}[s] &= E[s^2] - E[s]^2 \\
&= E[(f+g)^2] - (E[f+g])^2 \\
&= E[(f+g)^2] - (E[f] + E[g])^2 \text{ by Lemma 4.2.1} \\
&= E[(f+g)^2] - (\bar{f} + \bar{g})^2 \\
&= E[f^2] + 2E[fg] + E[g^2] - (\bar{f} + \bar{g})^2. \qquad (4.10)
\end{aligned}$$

We justify the last step as follows:

$$\begin{aligned}
E[(f+g)^2] &= E[f^2 + 2fg + g^2] \\
&= E[(f^2 + 2fg) + g^2] \\
&= E[(f^2 + 2fg)] + E[g^2] \text{ by Lemma 4.2.1} \\
&= E[f^2] + E[2fg] + E[g^2] \text{ by Lemma 4.2.1} \\
&= E[f^2] + 2E[fg] + E[g^2] \text{ by Lemma 4.2.2.} \qquad (4.11)
\end{aligned}$$

Using independence again we can factorize $E[fg] = \sum_{ij} r_{ij} f_i g_j = \sum_i p_i f_i \cdot \sum_j q_j g_j = E[f]E[g] = \bar{f}\bar{g}$. Elegantly, this term is responsible for cancelling the cross-term in $(\bar{f} + \bar{g})^2$, and collecting the remaining terms leaves

$$\text{Var}[s] = \text{Var}[f] + \text{Var}[g] \qquad (4.12)$$

So, the variances *also* add.

We now take the special case when the second event is identical to the £rst. That is, $M = N$, $q_i = p_i$, $g_i = f_i$, $\forall i$. In this case the above shows that $\bar{s} = 2\bar{f}$ and $\text{Var}[s] = 2\text{Var}[f]$.

We can repeat the above combination law for successively repeated events. Such events are called *identical independently distributed* (iid) events. Suppose we have K repetitions of iid events, with total $s \equiv f + g + \cdots + z$, we can just repeatedly apply the above rules to get

$$\begin{aligned}
\bar{s} &= K\,\bar{f} \qquad (4.13) \\
\text{Var}[s] &= K\,\text{Var}[f]. \qquad (4.14)
\end{aligned}$$

The mean and variance are not the only characteristics of the distribution that add like in this way. Amazingly, there is an in£nite sequence of special combinations of the higher moments, called *cumulants*, which add just like this. The mean and variance are just the £rst two cumulants.

4.2.3 Simplest random walk

We now have the tools to characterize a random walk. We choose $N = 2$ and $p_1 = p_2 = \frac{1}{2}$, just as with the coin toss, and de£ne the "step" displacements $f_1 = +1$ and $f_2 = -1$. This corresponds to a drunkard taking (uncorrelated) steps of unit length along the integer line.

We use Eqs. 4.3 and 4.5 to evaluate the mean and variance of a single step event.

$$\begin{aligned} \bar{f} &= +\frac{1}{2} - \frac{1}{2} = 0 && (4.15) \\ \text{Var}[f] &= \frac{1}{2}.(1)^2 + \frac{1}{2}.(-1)^2 - (0)^2 = 1. && (4.16) \end{aligned}$$

For K steps, starting from the origin, we have the £nal displacement of s, the sum of all the steps, using the formulae above,

$$\begin{aligned} \bar{s} &= K \cdot 0 = 0 && (4.17) \\ \text{Var}[s] &= K \cdot 1 = K. && (4.18) \end{aligned}$$

So the standard deviation, *i.e.* the width of the distribution of s, has value

$$\sigma_s \equiv \sqrt{\text{Var}[s]} = \sqrt{K}. \qquad (4.19)$$

This is our £rst vital fact about all but the most pathological [1] random walks: the distribution has width which scales like $K^{1/2}$. This means that a *typical* distance from the origin is $\sqrt{K}$. This is called a *diffusion process* and is very common in the real world.

Again, remember that if the person walks in feet, the variance (which is a measure of how much the distribution spreads out) will be in feet-squared. By taking the square-root we again have units of feet.

4.2.4 Central Limit Theorem

The Central Limit Theorem (CLT) states that the distribution on s tends to a *Gaussian distribution*,

$$p(s) \approx \mathcal{N}(\mu, \sigma^2) \equiv \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(s-\mu)^2}{2\sigma^2}} \qquad (4.20)$$

[1] Random walks which do *not* exhibit this power law have in£nite variance and exhibit *anomalous diffusion*. For more on this, see M. Bazant's excellent course at `http://www-math.mit.edu/~bazant/teach/18.325`

with mean $\mu = \bar{s}$ and variance $\sigma^2 = \mathrm{Var}[s]$ as given above, as $K \to \infty$. This is a very common "bell curve" with width σ centered about μ. We have not de£ned $p(s)$ very rigorously—it is simply the probability of being displaced s from the origin at the end of the K-step random walk. An exact formula for $p(s)$ involves counting all the ways that ± 1 can be added K times to get exactly s. We will postpone this, and the proof of CLT, for next time.

Remarkably the CLT applies to *any* N with any discrete step distribution $\{p_i\}$ and any step displacements $\{f_i\}$. It also applies to the case of continous-valued steps with distribution $p(f)$ along the real line. However a criterion for validity is always the *£niteness of the second moment* of $p(f)$.

Chapter 5

Quadratic Reciprocity, Central Limit Theorem and Graph Problems

We give Eisenstein's proof of Quadratic Reciprocity, and then introduce the Graph Theory problems. Lecture by Steven J. Miller and Peter Sarnak; notes by Steven J. Miller.

5.1 Eisenstein's Proof of Quadratic Reciprocity

5.1.1 Preliminaries

Theorem 5.1.1 (Quadratic Reciprocity). *Let p and q be distinct odd primes. Then*

$$\left(\frac{q}{p}\right)\left(\frac{p}{q}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}. \tag{5.1}$$

As p and q are distinct, odd primes, both $\left(\frac{q}{p}\right)$ and $\left(\frac{p}{q}\right)$ are ± 1. The dif£culty is £guring out which signs are correct, and how the two signs are related.

We use Euler's Criterion, proved in a previous lecture:

Lemma 5.1.2 (Euler's Criterion).

$$\left(\frac{q}{p}\right) \equiv q^{\frac{p-1}{2}} \; mod \, p. \tag{5.2}$$

The idea behind Eisenstein's proof is as follows: $\left(\frac{q}{p}\right)\left(\frac{p}{q}\right)$ is -1 to a power. Further, we only need to determine the power mod 2. Eisenstein shows many expressions are equivalent, mod 2, to this power. Eventually, we arrive at an expression which is trivial to calculate (mod 2).

5.1.2 First Stage

Consider all even multiples of q by an $a \leq p-1$: $\{2q, 4q, 6q, \ldots, (p-1)q\}$. Denote a generic multiple by aq. Recall $[x]$ is the greatest integer less than or equal to x. By integer division,

$$aq = \left[\frac{aq}{p}\right] p + r_a, \ 0 \leq r_a < p-1. \tag{5.3}$$

Thus, r_a is the least non-negative number equivalent to qa mod p.

The numbers $(-1)^{r_a} r_a$ are equivalent to even numbers in $\{0, \ldots, p-1\}$. If r_a is even this is clear; if r_a is odd, then $(-1)^{r_a} r_a \equiv p - r_a$ mod p, and as p and r_a are odd, this is even.

Lemma 5.1.3. *If* $(-1)^{r_a} r_a \equiv (-1)^{r_b} r_b$, *then* $a = b$.

Proof: We quickly get $\pm r_a \equiv r_b$ mod p. If the plus sign holds, then $r_a \equiv r_b$ mod p implies $qa \equiv qb$ mod p. As q is invertible mod p, we get $a \equiv b$ mod p, which yields $a = b$ (as a and b are even integers between 0 and $p-1$).

If the minus sign holds, then $r_a + r_b \equiv 0$ mod p, or $qa + qb \equiv 0$ mod p. Multiplying by q^{-1} mod p now gives $a + b \equiv 0$ mod p. As a and b are even integers between 0 and $p-1$, $0 < a + b \leq 2(p-1)$. The only integer strictly between 0 and $2p$ which is equivalent to 0 mod p is p; however, p is odd and $a+b$ is even. Thus, the minus sign cannot hold, and the elements are all distinct. $\Box$.

Lemma 5.1.4.

$$\left(\frac{q}{p}\right) = (-1)^{\sum_{a \text{ even}} r_a}. \tag{5.4}$$

Proof: For each even a, $qa \equiv r_a$ mod p. Thus, mod p:

$$\begin{aligned} \prod_{a \text{ even}} qa &\equiv \prod_{a \text{ even}} r_a \\ q^{\frac{p-1}{2}} \prod_{a \text{ even}} a &\equiv \prod_{a \text{ even}} r_a \\ \left(\frac{q}{p}\right) \prod_{a \text{ even}} a &\equiv \prod_{a \text{ even}} r_a, \end{aligned} \tag{5.5}$$

where the above follows from the fact that we have $\frac{p-1}{2}$ choices for an even a (to get $q^{\frac{p-1}{2}}$ and Euler's Criterion (to replace $q^{\frac{p-1}{2}}$ with $\left(\frac{q}{p}\right)$).

As a ranges over all even number from 0 to $p-1$, so too do the numbers $(-1)^{r_a} r_a$ mod p. Thus, mod p,

$$\begin{aligned} \prod_{a \text{ even}} a &\equiv \prod_{a \text{ even}} (-1)^{r_a} r_a \\ \prod_{a \text{ even}} a &= (-1)^{\sum_{a \text{ even}} r_a} \prod_{a \text{ even}} r_a. \end{aligned} \tag{5.6}$$

Combining gives

$$\left(\frac{q}{p}\right) (-1)^{\sum_{a \text{ even}} r_a} \prod_{a \text{ even}} r_a \equiv \prod_{a \text{ even}} r_a. \tag{5.7}$$

As each r_a is invertible mod p, so is the product. Thus,

$$\left(\frac{q}{p}\right) (-1)^{\sum_{a \text{ even}} r_a} \equiv 1 \bmod p. \tag{5.8}$$

As $\left(\frac{q}{p}\right)$ is its own inverse, the Lemma now follows by multiplying both sides by $\left(\frac{q}{p}\right)$. $\Box$.

Therefore, it is suf£cient to determine $\sum_{a \text{ even}} r_a$ mod 2.

We make one last simpli£cation. By integer division, we have

$$\begin{aligned} \sum_{a \text{ even}} qa &= \sum_{a \text{ even}} \left(\left[\frac{qa}{p}\right] p + r_a \right) \\ &= \sum_{a \text{ even}} \left[\frac{qa}{p}\right] p + \sum_{a \text{ even}} r_a. \end{aligned} \tag{5.9}$$

As we are summing over even a, the Left Hand Side above is even. Thus, the Right Hand Side is even, so

$$\begin{aligned} \sum_{a \text{ even}} \left[\frac{qa}{p}\right] p &\equiv \sum_{a \text{ even}} r_a \bmod 2 \\ p \sum_{a \text{ even}} \left[\frac{qa}{p}\right] &\equiv \sum_{a \text{ even}} r_a \bmod 2 \\ \sum_{a \text{ even}} \left[\frac{qa}{p}\right] &\equiv \sum_{a \text{ even}} r_a \bmod 2, \end{aligned} \tag{5.10}$$

where the last line follows from the fact that p is odd, so mod 2, dropping the factor of p from the Left Hand Side doesn't change the parity.

We have shown

Lemma 5.1.5. *It is suf£cient to calculate* $\sum_{a \text{ even}} \left[\frac{qa}{p}\right]$

5.1.3 Second Stage

Consider the rectangle with vertices at $A = (0, 0)$, $B = (p, 0)$, $C = (p, q)$ and $D = (0, q)$. The upward slopping vertical is given by the equation $y = \frac{q}{p}x$. As p and q are distinct odd primes, there are no pairs of integers (x, y) on the line AC.

We now interpret $\sum_{a \text{ even}} \left[\frac{qa}{p}\right]$. Consider the vertical line with x coordinate a. Then $\left[\frac{qa}{p}\right]$ gives the number of pairs (x, y) with x-coordinate equal to a and y-coordinate an integer at most $\left[\frac{qa}{p}\right]$. Thus, $\sum_{a \text{ even}} \left[\frac{qa}{p}\right]$ is the number of integer pairs (in the rectangle $ABCD$) with even x-coordinate that are below the line AC.

We add some additional points: $E = (\frac{p}{2}, 0)$, $F = (\frac{p}{2}, \frac{q}{2})$, $G = (0, \frac{q}{2})$ and $H = (\frac{p}{2}, q)$. We prove

Lemma 5.1.6. *The number of integer pairs under the line AC (inside the rectangle) with even x-coordinate is congruent mod 2 to the number of integer pairs under the line AF.*

Let $a > \frac{p}{2}$ be an even integer. The integer pairs on the line $x = a$ are $(a, 0)$, $(a, 1), \ldots, (a, q)$. There are $q + 1$ pairs. As q is odd, there are an even number of integer pairs on the line $x = a$. As there are no integer pairs on the line AC, for a

£xed $a > \frac{p}{2}$, mod 2 there are the same number of integer pairs *above* AC as there are *below* AC.

Further, the number of integer pairs *above* AC is equivalent mod 2 to the number of integer pairs below AF on the line $x = p - a$. To see this, consider the map which takes (x, y) to $(p - x, q - y)$. As $a > \frac{p}{2}$ and is even, $p - a < \frac{p}{2}$ and is odd. Further, every odd $a < \frac{p}{2}$ is hit (given $a_{odd} < \frac{p}{2}$, start with the even number $p - a_{odd} > \frac{p}{2}$).

Let $\#FCH_{even}$ be the number of integer pairs (x, y) in triangle FCH with x even.

Let $\#EBCH$ be the number of integer pairs in the rectangle $EBCH$; $\#EBCH \equiv 0$ mod 2 (we've shown each vertical line has an even number of pairs).

Let $\#AFE_{even}$ be the number of integer pairs (x, y) in the triangle AFE with x even, and let $\#AFE$ be the number of integer pairs in the triangle AFE.

We need to calculate $\sum_{a \text{ even}} \left[\frac{qa}{p}\right]$ mod 2:

$$\begin{aligned}
\sum_{a \text{ even}} \left[\frac{qa}{p}\right] &= \#AFE_{even} + \#EBCH - \#FCH \\
&\equiv \#AFE_{even} + \#EBCH + \#FCH \\
&= \#AFE_{even} + \#FCH + \#EBCH \\
&= \#AFE + \#EBCH \\
&= \#AFE. \qquad (5.11)
\end{aligned}$$

Therefore, $\mu = \sum_{a \text{ even}} \left[\frac{qa}{p}\right] \equiv \#AFE$ mod 2, and we have

$$\left(\frac{q}{p}\right) = (-1)^{\mu}. \qquad (5.12)$$

Reversing the rolls of p and q, we see that

$$\left(\frac{p}{q}\right) = (-1)^{\nu}, \qquad (5.13)$$

where $\nu \equiv \#AFG$ mod 2, with $\#AFG$ equal to the number of integer pairs in the triangle AFG.

Now, $\mu + \nu = \#AFE + \#AFG$, which is the number of integer pairs in the rectangle $AEFG$. There are $\frac{p-1}{2}$ choices for x and $\frac{q-1}{2}$ choices for y, giving $\frac{p-1}{2}\frac{q-1}{2}$ pairs of integers in the rectangle $AEFG$.

Thus,

$$\begin{aligned}\left(\frac{q}{p}\right)\left(\frac{p}{q}\right) &= (-1)^{\mu+\nu} \\ &= (-1)^{\#AFE+\#AFG} \\ &= (-1)^{\frac{p-1}{2}\frac{q-1}{2}},\end{aligned} \tag{5.14}$$

which completes the proof of Quadratic Reciprocity. $\Box$.

5.2 Central Limit Theorem

$X_1, X_2, X_3, \dots$ an in£nite sequence of random variables such that the X_j are independent identically distributed random variables (abbreviated i.i.d.r.v.) with $E[X_j] = \bar{X}_j = 0$ (can always renormalize by shifting) and variance $E[X_j^2] = 1$. Let $S_N = \sum_{j=1}^{N} X_j$.

Theorem 5.2.1. *Fix $-\infty < a \le b < \infty$. Then as $N \to \infty$,*

$$\textit{Prob}\Big(\frac{S_N}{\sqrt{N}} \in [a,b]\Big) \to \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{t^2}{2}} dt. \tag{5.15}$$

The probability function is called the Gaussian or the Normal distribution. This is the universal curve of probability. Note how robust the Central Limit Theorem is: it doesn't depend on £ne properties of the X_j.

5.3 Possible Problems

5.3.1 Combinatorics and Probability

Combinatorics is the number of ways of doing something. A **graph** is a set of **vertices** (V) and **edges** (E) between them. Thus, edges are unordered pairs of vertices.

Four Color Theorem (proved by an exhaustive search by the computer). Say you have a graph in the plane (thus, if you draw the vertices and edges in the plane, no two edges cross). Call such a graph a **planar graph**. Can you color the vertices such that if two vertices are joined by an edge, they have different colors? Sure, by using $|V|$ colors! What is the least number of colors needed?

Theorem 5.3.1 (Four Coloring Theorem). *You can color the vertices of a planar graph using at most four colors such that no two joined vertices have the same color.*

A k**-regular graph** is a graph such that there are k-edges out of each vertex. A 2-regular graph has no freedom: you get a closed cycle.

Consider 3-regular graphs. To each graph associate the **adjacency matrix** A. (First to study may be Kirchhoff). Say $G = (V, E)$ has $|V| = n$ vertices. For now, assume there are no **multiple edges** (ie, between any two vertices is at most one edge, and there are no edges connecting a vertex to itself). A is an $n \times n$ matrix, rows and columns indexed by the vertices, and $A_{ij} = 1$ if there is an edge from v_i to v_j and 0 otherwise.

Thus, the adjacency matrix is a matrix with 0s and 1s, and is symmetric.

Problem: What is the second largest eigenvalue of A? How does it vary? What do we expect?

In Linear Algebra, we learn we can diagonalize a real symmetric matrix. The eigenvalues are real, and satisfy $p_A(\lambda) = \det(\lambda I - A)$, the characteristic polynomial. This is a polynomial in λ of degree n with integer coef£cients. Thus, the eigenvalues are algebraic numbers. The leading coef£cient is λ^n, the constant term is $\det(A)$.

Thus, $p_A(\lambda) = \lambda^n + \cdots + \det(A)$, and by the Fundamental Theorem of Algebra, there are n complex roots. If the leading coef£cient of the de£ning polynomial is 1, we say the roots are **algebraic integers**. These roots are the eigenvalues of A.

Why must the eigenvalues be real? Want $Av = \lambda v$, v a non-zero vector.

Fact: if $v = (1, 1, \ldots, 1)^T$, v is an eigenvector of A with eigenvalue k. Why? As each vertex is connected to k distinct vertices, each row has exactly k entries that are 1 and $n - k$ entries that are 0. Thus, k is an eigenvalue, denote by λ_0.

Exercise 5.3.2. *Show, for such adjacency matrices A, that all eigenvalues satisfy $-k \le \lambda \le k$.*

Consider connected graphs G. How big is $\lambda_1(G)$ for the random 3-regular graph?

Theorem 5.3.3 (Kirchhoff's Theorem). *Let $\det^*(kI - A)$ be the product of the non-zero eigenvalues of A. Then $\det^*(kI - A)$ equals the number of vertices times the number of spanning trees.*

A tree is a connected graph with no cycles. A **spanning tree** is a connected sub-graph containing all the vertices. Sometimes called complexity of the graph.

Chapter 6

Ef£cient Algorithms, Probability, Alg+Transcendental, Pidgeon Hole, Chebychev

We review many basic number theory results. We give ef£cient algorithms for polynomial evaluation, calculating x^n, and £nding the greatest common divisor. We brie¤y review probability theory. After an introduction to Algebraic and Transcendental Numbers, we review Dirichlet's Box Principle (aka the Pidgeonhole Principle), and give an application. We prove a weak version of Chebyshev's Theorem on the approximate number of primes. Lecture by Steven J. Miller. Notes by Steven J. Miller (and Florin Spinu, who helped write up the notes on Dirichlet's Box Principle and Chebyshev).

6.1 Notation

1. $\mathbb{W}$: the set of whole numbers: $\{1, 2, 3, 4, \dots\}$.
2. $\mathbb{N}$: the set of natural numbers: $\{0, 1, 2, 3, \dots\}$.
3. $\mathbb{Z}$: the set of integers: $\{\dots, -2, -1, 0, 1, 2, \dots\}$.
4. $\mathbb{Q}$: the set of rational numbers: $\{x : x = \frac{p}{q}, p, q \in \mathbb{Z}, q \neq 0\}$.
5. $\mathbb{R}$: the set of real numbers.
6. $\mathbb{C}$: the set of complex numbers: $\{z : z = x + iy,\ x, y \in \mathbb{R}\}$.

7. $\mathbb{Z}/n\mathbb{Z}$: the additive group of integers mod n.
8. $(\mathbb{Z}/n\mathbb{Z})^*$: the multiplicative group of invertible elements mod n.
9. $a|b$: a divides b, *i.e.* the remainder after integer division $\frac{b}{a}$ is 0.
10. (a, b) : greatest common divisor (gcd) of a and b, often written $\gcd(a, b)$.
11. $x \equiv y(\mod n)$: there exists an integer a such that $x = y + an$.
12. wlog : without loss of generality.
13. s.t. : such that.
14. $\forall$: for all.
15. $\exists$: there exists.
16. big O notation : $A(x) = O(B(x))$, read "$A(x)$ is of order $B(x)$", means $\exists C > 0$ such that $\forall x$, $|A(x)| \leq C\,B(x)$.
17. $|S|$ or $\#S$: number of elements in the set S.
18. p : usually a prime number.
19. n : usually an integer.

6.2 Ef£cient Algorithms

For computational purposes, often having an algorithm to compute a quantity is not enough; we need an algorithm which will compute *quickly*. Below we study three standard problems, and show how to either rearrange the operations more ef£ciently, or give a more ef£cient algorithm than the obvious candidate.

6.2.1 Polynomial Evaluation

Let $f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$. The obvious way to evaluate is to calculate x^n and multiply by a_n (n multiplications), calculate x^{n-1} and multiply by a_{n-1} ($n-1$ multiplications) and add, et cetera. There are n additions and $\sum_{k=0}^{n} k$ multiplications, for a total of $n + \frac{n(n+1)}{2}$ operations. Thus, the standard method leads to $O(n^2)$ computations.

Instead, consider the following:

$$\left(\left((a_n x + a_{n-1})x + a_{n-2}\right)x + \cdots + a_1\right)x + a_0. \tag{6.1}$$

For example,

$$7x^4 + 4x^3 - 3x^2 - 11x + 2 = \left(\left((7x+4)x - 3\right)x - 11\right)x + 2. \tag{6.2}$$

Evaluating the long way takes 14 steps; cleverly rearranging takes 8 steps.

Exercise 6.2.1. *Prove that the second method takes at most $2n$ steps to evaluate $a_n x^n + \cdots a_0$.*

6.2.2 Exponentiation

Consider x^n. The obvious way to evaluate involves $n-1$ multiplications. By writing n in base two, we can evaluate x^n in at most $2\log_2 n$ steps.

Let k be the largest integer such that $2^k \leq n$. Then $\exists a_i \in \{0,1\}$ such that

$$n \;=\; a_k 2^k + a_{k-1}2^{k-1} + \cdots + a_1 2 + a_0. \tag{6.3}$$

It costs k multiplications to evaluate x^{2^i}, $i \leq k$. How? Consider $y_0 = x^{2^0}$, $y_1 = y_0 \cdot y_0 = x^{2^0} \cdot x^{2^0} = x^{2^1}$, $y_2 = y_1 \cdot y_1 = x^{2^2}, \dots, y_k = y^{k-1} \cdot y^{k-1} = x^{2^k}$.

Then

$$\begin{aligned} x^n &= x^{a_k 2^k + a_{k-1}2^{k-1} + \cdots + a_1 2 + a_0} \\ &= x^{a_k 2^k} \cdot x^{a_{k-1}2^{k-1}} \cdots x^{a_1 2} \cdot x^{a_0} \\ &= \left(x^{2^k}\right)^{a_k} \cdot \left(x^{2^{k-1}}\right)^{a_{k-1}} \cdots \left(x^2\right)^{a_1} \cdot \left(x^1\right)^{a_0} \\ &= y_k^{a_k} \cdot y_{k-1}^{a_{k-1}} \cdots y_1^{a_1} \cdot y_0^{a_0}. \end{aligned} \tag{6.4}$$

As each $a_i \in \{0,1\}$, we have at most $k+1$ multiplications above (if $a_i = 1$ we have the term y_i in the product, if $a_i = 0$ we don't).

Thus, it costs k multiplications to evaluate the x^{2^i} ($i \leq k$), and at most another k multiplications to £nish calculating x^n. As $k \leq \log_2 n$, we see that x^n can be determined in at most $2\log_2 n$ steps.

Note, however, that we do need more storage space for this method, as we need to store the values $y_i = x^{2^i}$, $i \leq \log_2 n$.

Exercise 6.2.2. *Instead of expanding n in base two, expand n in base three. How many calculations are needed to evaluate x^n this way? Why is it preferable to expand in base two rather than any other base?*

6.2.3 Euclidean Algorithm

The Euclidean Algorithm is an ef£cient way to determine the greatest common divisor of x and y, denoted $\gcd(x, y)$ or (x, y). Without loss of generality, assume $1 < x < y$.

The obvious way to determine $\gcd(x, y)$ is to divide x and y by all positive integers up to x. This takes at most $2x$ steps.

Let $[z]$ denote the greatest integer less than or equal to z. We write

$$y = \frac{y}{x} \cdot x + r_1, \; 0 \leq r_1 < x. \tag{6.5}$$

Exercise 6.2.3. *Prove that $r_1 \in \{0, 1, \dots, x - 1\}$.*

Exercise 6.2.4. *Prove $\gcd(x, y) = \gcd(r_1, x)$. Hint: $r_1 = y - \frac{y}{x} \cdot x$.*

We proceed in this manner until r_k equals zero or one. As each execution results in $r_i < r_{i-1}$, we proceed at most x times (although later we prove we need to apply these steps at most $2 \log_2 x$ times).

$$\begin{aligned}
x &= \frac{x}{r_1} \cdot r_1 + r_2, \; 0 \leq r_2 < r_1 \\
r_1 &= \frac{r_1}{r_2} \cdot r_2 + r_3, \; 0 \leq r_3 < r_2 \\
r_2 &= \frac{r_2}{r_3} \cdot r_3 + r_4, \; 0 \leq r_4 < r_3 \\
&\vdots \\
r_{k-2} &= \frac{r_{k-2}}{r_{k-1}} \cdot r_{k-1} + r_k, \; 0 \leq r_k < r_{k-1}.
\end{aligned} \tag{6.6}$$

Exercise 6.2.5. *Prove that if $r_k = 0$, then $\gcd(x, y) = r_{k-1}$, while if $r_k = 1$, then $\gcd(x, y) = 1$.*

We now analyze how large k can be. The key observation is the following:

Lemma 6.2.6. *Consider three adjacent remainders in the expansion: r_{i-1}, r_i and r_{i+1} (where $y = r_{-1}$ and $x = r_0$). Then $\gcd(r_i, r_{i-1}) = \gcd(r_{i+1}, r_i)$, and $r_{i+1} < \frac{r_{i-1}}{2}$.*

Proof: We have the following relation:

$$r_{i-1} = \frac{r_{i-1}}{r_i} \cdot r_i + r_{i+1},\ 0 \le r_{i+1} < r_i. \tag{6.7}$$

If $r_i \le \frac{r_{i-1}}{2}$, then as $r_{i+1} < r_i$, we immediately conclude that $r_{i+1} < r_i$. If $r_i > \frac{r_{i-1}}{2}$, then we note that

$$r_{i+1} = r_{i-1} - \frac{r_{i-1}}{r_i} \cdot r_i. \tag{6.8}$$

But $\frac{r_{i-1}}{r_i} = 1$ (easy exercise). Thus $r_{i-1} < \frac{r_{i-1}}{2}$. $\Box$

We count how often we apply Euclid's Algorithm. Going from $(x, y) = (r_0, r_{-1})$ to (r_1, r_0) costs one application. Every two applications leads to the £rst entry in the last pair being at most half of the second entry of the £rst pair.

Thus, if k is the largest integer such that $2^k \le x$, we see we apply Euclid's Algorithm at most $1 + 2k \le 1 + 2\log_2 x$ times. Each application requires one integer division, where the remainder is the input for the next step.

We have proven

Lemma 6.2.7. *Euclid's Algorithm requires at most $1 + 2\log_2 x$ divisions to £nd the greatest common denominator of x and y.*

Let us assume that $r_i = \gcd(x, y)$. Thus, the last equation before Euclid's Algorithm terminated was

$$r_{i-2} = \frac{r_{i-2}}{r_{i-1}} \cdot r_{i-1} + r_i,\ 0 \le r_i < r_{i-1}. \tag{6.9}$$

Therefore, we can £nd integers a_{i-1} and b_{i-2} such that

$$r_i = a_{i-1} r_{i-1} + b_{i-2} r_{i-2}. \tag{6.10}$$

Looking at the second to last application of Euclid's algorithm, we £nd that there are integers a'_{i-2} and b'_{i-3} such that

$$r_{i-1} = a'_{i-2} r_{i-2} + b'_{i-3} r_{i-3}. \tag{6.11}$$

Substituting for $r_{i-1} = r_{i-1}(r_{i-2}, r_{i-3})$ in the expansion of r_i yields that there are integers a_{i-2} and b_{i-3} such that

$$r_i = a_{i-2} r_{i-2} + b_{i-3} r_{i-3}. \tag{6.12}$$

Continuing by induction, and recalling $r_i = \gcd(x, y)$ yields

Lemma 6.2.8. *There exist integers a and b such that $\gcd(x, y) = ax + by$. Moreover, Euclid's Algorithm gives a constructive procedure to find a and b.*

Exercise 6.2.9. *Find a and b such that $a \cdot 244 + b \cdot 313 = \gcd(244, 313)$.*

Exercise 6.2.10. *Add details to complete an alternate proof of the existence of a and b with $ax + by = \gcd(x, y)$:*

1. *Let d be the smallest positive value attained by $ax + by$ as we vary $a, b \in \mathbb{Z}$. Such a d exists: consider $(a, b) = (1, 0)$ or $(0, 1)$. Thus, $d = ax + by$. We now show $d = \gcd(x, y)$.*

2. $\gcd(x, y)|d$.

3. *Let $e = Ax + By > 0$. Then $d|e$. Therefore, for any choice of $A, B \in \mathbb{Z}$, $d|(Ax + By)$.*

4. *$d|x$ and $d|y$ (consider clever choices of A and B; one choice gives $d|x$, one gives $d|y$). Therefore $d|\gcd(x, y)$. As we've shown $\gcd(x, y)|d$, this completes the proof.*

Note this is a non-constructive proof. By minimizing $ax + by$, we obtain $\gcd(x, y)$, but we have no idea how many steps is required. Prove that a solution will be found either among pairs (a, b) with $a \in \{1, \dots, y - 1\}$ and $-b \in \{1, \dots, x - 1\}$, or $-a \in \{1, \dots, y - 1\}$ and $b \in \{1, \dots, x - 1\}$.

6.3 Probabilities of Discrete Events

6.3.1 Introduction

Let $\Omega = \{\omega_1, \omega_2, \omega_3, \dots\}$ be an at most countable set of events. We call Ω the **sample (or outcome) space**. We call the elements $\omega \in \Omega$ the **events**. Let $x : \Omega \to \mathbb{R}$. That is, for each event $\omega \in \Omega$, we attach a real number $x(\omega)$. We call x a **random variable**.

Example 6.3.1. *Flip a fair coin 3 times. The possible outcomes are $\Omega = \{HHH, HHT, HTH, THH,$ One possible random variable is $x(\omega)$ equals the number of heads in ω. Thus, $x(HHT) = 2$ and $x(TTT) = 0$.*

Example 6.3.2. *Let Ω be the space of all ¤ips of a fair coin where all but the last ¤ip are tails, and the last is a head. Thus, $\Omega = \{H, TH, TTH, TTTH, \dots\}$. One possible random variable is $x(\omega)$ is the number of tails; another is $x(\omega)$ equals the number of the ¤ip which is a head.*

We say $p(\omega)$ is a **probability function** on Ω if

1. $0 \leq p(\omega_i) \leq 1$ for all $\omega_i \in \Omega$.
2. $p(\omega) = 0$ if $\omega \notin \Omega$.
3. $\sum_i p(\omega_i) = 1$.

We call $p(\omega)$ the probability of event ω.

Often, we have a random variables where $x(\omega) = \omega$. In a convenient abuse of notation, we write X for Ω and x for $x(\omega)$ and ω. For example, consider two rolls of a fair die. Let X be the result of the £rst roll, and Y of the second. Then the sample space is $X = Y = \{1, 2, 3, 4, 5, 6\}$.

In general, consider X and Y with x_i occurring with probability $p(x_i)$ and y_j occurring with probability $q(y_j)$. We analyze the **joint probability** $r(x, y)$ of observing x and y.

X and Y are **independent** if $\forall x, y$, $r(x, y) = p(x)q(y)$. In the example of rolling a fair die twice, $r(x, y) = p(x)q(y) = \frac{1}{6} \cdot \frac{1}{6}$ if $x, y \in X = Y$, and 0 otherwise.

Exercise 6.3.3. *Consider again two rolls of a fair die. Now, let X represent the £rst roll, and Y the sum of the £rst two rolls. Prove X and Y are not independent.*

Events X_1 through X_N are **independent** if $p(x_1, \dots, x_N) = p_1(x_1) \cdots p(x_N)$.

Exercise 6.3.4. *Construct three events such that any two are independent, but all three are not independent. Hint: roll a fair die twice.*

6.3.2 Means

If $x(\omega) = \omega$, the **mean (or expected value)** of an event x is de£ned by

$$\bar{x} = \sum_i x_i p(x_i). \tag{6.13}$$

More generally, for a sample space Ω with events ω and a random variable $x(\omega)$, we have

$$\bar{x}(\omega) \;=\; \sum_i x(\omega_i)p(\omega_i). \tag{6.14}$$

For example, the mean of one roll of a fair die is 3.5.

Exercise 6.3.5. *Let X be the number of tosses of a fair coin needed before getting the £rst head. Thus, $X = \{1, 2, \dots\}$. Calculate $p(x_i)$ and $\bar{x}$. We could let Ω be the space of all tosses of a fair coin where all but the last toss are tails, and the last toss is a head. Then $x(\omega)$ is the number of tosses of ω.*

Instead of writing $\bar{x}$, we often write $E[x]$ or $E[X]$, read as **the expected value of** x **or** X. More generally, we would have $\bar{x}(\omega)$ and $E[x(\omega)]$.

The k^{th} **moment of** X is the expected value of x^k:

$$E[x^k] \;=\; \sum_i x_i^k p(x_i) \tag{6.15}$$

or

$$E[x^k(\omega)] \;=\; \sum_i x^k(\omega_i)p(\omega_i). \tag{6.16}$$

Lemma 6.3.6 (Additivity of the Means). *Let X and Y be two independent events with joint probability $r(x, y) = p(x)q(y)$. Let $z = x+y$. Then $E[z] = E[x+y] = E[x] + E[y]$.*

Proof:

$$\begin{aligned}
E[x+y] &= \sum_{(i,j)} (x_i + y_j) r(x_i, y_j) \\
&= \sum_i \sum_j (x_i + y_j) p(x_i) q(y_j) \\
&= \sum_i \sum_j x_i p(x_i) q(y_j) + \sum_i \sum_j y_j p(x_i) q(y_j) \\
&= \sum_i x_i p(x_i) \sum_j q(y_j) + \sum_i p(x_i) \sum_j y_j q(y_j) \\
&= E[x] \cdot 1 + 1 \cdot E[y] = E[x] + E[y].
\end{aligned} \tag{6.17}$$

The astute reader may notice that some care is needed to interchange the order of summations. If $\sum_i \sum_j |x_i y_j| r(x_i, y_j) < \infty$, then Fubini's Theorem is applicable, and we may interchange the summations at will.

We used the two events were independent to go from $\sum_{(i,j)} x_i r(x_i, y_j)$ to $\sum_i x_i p(x_i) \sum_j q(y_j) = E[x]$. Lemma 6.3.6 is true even if the two events are not independent.

If the events are not independent, we encounter sums like $\sum_i \sum_j x_i r(x_i, y_j)$; however, $\sum_j r(x_i, y_j) = p(x_i)$. Why? By summing over all possible y, we are asking what is the probability that $x = x_i$; we do not care what y is. Thus, $\sum_i \sum_j x_i r(x_i, y_j) = \sum_i x_i p(x_i) = E[x]$, and similarly for the other piece.

Exercise 6.3.7. *Write out the proof of the generalization of Lemma 6.3.6, where X and Y are not assumed independent.*

Given an outcome space $X = \{x_1, x_2, \dots\}$ with probabilities $p(x_i)$, let aX be shorthand for the event a times X with outcome space $\{ax_1, ax_2, \dots\}$ and probabilities $p_a(ax_i) = p(x_i)$.

Lemma 6.3.8. *Let X_1 through X_N be a £nite collection of independent events. Let a_1 through a_N be real constants. Then*

$$E[a_1 x_1 + \cdots + a_N x_N] = a_1 E[x_1] + \cdots + a_N E[x_N]. \tag{6.18}$$

Lemma 6.3.9. *Let X and Y be independent events. Then $E[xy] = E[x]E[y]$.*

Exercise 6.3.10. *Prove Lemmas 6.3.8 and 6.3.9.*

6.3.3 Variances

The **variance** σ_x^2 (and its square-root, the **standard deviation** σ_x) measure how spread out a probability distribution is. Assume $x(\omega) = \omega$. Given an event X with mean $\bar{x}$, we de£ne the standard deviation σ_x^2 by

$$\sigma_x^2 \;=\; \sum_i (x_i - \bar{x}) p(x_i). \tag{6.19}$$

More generally, given a sample space Ω, events ω, and a random variable $x : \Omega \to \mathbb{R}$,

$$\sigma_{x(\omega)}^2 \;=\; \sum_i \Big(x(\omega_i) - \bar{x}(\omega) \Big) p(\omega_i). \tag{6.20}$$

Exercise 6.3.11. *Let* $X = \{0, 25, 50, 75, 100\}$ *with probabilities* $\{.2, .2, .2, .2, .2\}$. *Let* Y *be the same outcome space, but with probabilities* $\{.1, .25, .3, .25, .1\}$. *Calculate the means and the variances of* X *and* Y.

For computing variances, instead of equation 6.19 one often uses

Lemma 6.3.12. $\sigma_x^2 = E[x^2] - E[x]^2$.

Proof: Recall $\bar{x} = E[x]$. Then

$$\begin{aligned} \sigma_x^2 &= \sum_i \Big(x_i - E[x]\Big)^2 p(x_i) \\ &= \sum_i (x_i^2 - 2x_i E[x] + E[x]^2) p(x_i) \\ &= \sum_i x_i^2 p(x_i) - 2E[x] \sum_i x_i p(x_i) + E[x]^2 \sum_i p(x_i) \\ &= E[x^2] - 2E[x]^2 + E[x]^2 = E[x^2] - E[x]^2. \end{aligned} \tag{6.21}$$

The main result on variances is

Lemma 6.3.13 (Variance of a Sum). *Let* X *and* Y *be two independent events. Then* $\sigma_{x+y}^2 = \sigma_x^2 + \sigma_y^2$.

Proof: We constantly use the expected value of a sum of independent events is the sum of expected values (Lemma 6.3.6 and Lemma 6.3.8).

$$\begin{aligned} \sigma_{x+y}^2 &= E[(x+y)^2] - E[(x+y)]^2 \\ &= E[x^2 + 2xy + y^2] - \Big(E[x] + E[y]\Big)^2 \\ &= \Big(E[x^2] + 2E[xy] + E[y^2]\Big) - \Big(E[x]^2 + 2E[x]E[y] + E[y]^2\Big) \\ &= \Big(E[x^2] - E[x]^2\Big) + \Big(E[y^2] - E[y]^2\Big) + 2\Big(E[xy] - E[x]E[y]\Big) \\ &= \sigma_x^2 + \sigma_y^2 + 2\Big(E[xy] - E[x]E[y]\Big). \end{aligned} \tag{6.22}$$

By Lemma 6.3.9, $E[xy] = E[x]E[y]$, completing the proof.

Lemma 6.3.14. *Consider* n *independent copies of the same event (for example,* n *¤ips of a coin or* n *rolls of a die). Then* $\sigma_{nx} = \sqrt{n}\sigma_x$.

Exercise 6.3.15. *Prove Lemma 6.3.14.*

Note that, if the event X has units of meters, then the variance σ_x^2 has units meters-squared, and the standard deviation σ_x and the mean $\bar{x}$ have units meters. Thus, it is the standard deviation that gives a good measure of the deviations of an event around the mean.

There are, of course, alternate measures one can use. For example, one could consider

$$\sum_i (x_i - \bar{x}) p(x_i). \tag{6.23}$$

Unfortunately, this is a signed quantity, and large positive deviations can cancel with large negatives. This leads us to consider

$$\sum_i |x_i - \bar{x}| p(x_i). \tag{6.24}$$

While this has the advantage of avoiding cancellation of errors (as well as having the same units as the events), the absolute value function is not a good function analytically. For example, it is not differentiable. This is primarily why we consider the standard deviation (the square-root of the variance).

Exercise 6.3.16. *Consider the following set of data: for $i \in \{1, \dots, n\}$, given x_i one observes y_i. Believing that X and Y are linearly related, £nd the best £t straight line. Namely, determine constants a and b that minimize the error (calculated via the variance)*

$$\sum_{i=1}^{n} \Big(y_i - (ax_i + b)\Big)^2 = \sum_{i=1}^{n} \Big(\textit{Observed}_i - \textit{Predicted}_i\Big)^2. \tag{6.25}$$

Hint: use Multi-variable Calculus to £nd linear equations for a and b, and then solve with Linear Algebra.

If instead of measuring total error by the squares of the individual error (for example, using the absolute value), closed form expressions for a and b become signi£cantly harder.

If one requires that $a = 0$, show that the b leading to least error is $b = \bar{y} = \frac{1}{n}\sum_i y_i$.

6.3.4 Random Walks

Consider the classical problem of a drunk staggering home from a lamp post late at night. We ¤ip a fair coin N times. With probability $\frac{1}{2}$ we get heads (tails). For each head (tail) the drunk staggers one unit to the right (left). How far do we expect the drunk to be?

It is very unlikely the drunk will be very far to the left or right.

Exercise 6.3.17. *Let x be $+1$ if we ¤ip a head, -1 for a tail. For a fair coin, prove $E[x] = 0$, $\sigma_x^2 = 1$, $\sigma_x = 1$.*

Exercise 6.3.18. *Let $p_N(y)$ be the probability that after N ¤ips of a fair coin, the drunk is y units to the right of the origin (lamp post).*

1. *Prove $p_N(y) = p_N(-y)$.*
2. *Consider $N = 2M$. Prove $p_{2M}(2k) = \binom{2M}{M+k}\frac{1}{2^{2M}}$, where $\binom{n}{r} = \frac{n!}{r!(n-r)!}$*
3. *Use Stirling's formula ($n! \approx n^n e^{-n}\sqrt{2\pi n} = \sqrt{2\pi} n^{n+\frac{1}{2}} e^{-n}$) to approximate $p_N(y)$.*

Label the coin tosses X_1 through X_N. Let X denote a generic toss of the coin, and Y_N be the distance of the drunkard after N tosses. By Lemma 6.3.8, $E[y_N] = E[x_1 + \cdots + x_N] = E[x_1] + \cdots + E[x_N]$. As each $E[x_i] = E[x] = 0$, $E[y_N] = 0$.

Thus, we expect the drunkard to be at the lamp post. How spread out is his expected position? By Lemma 6.3.14,

$$\sigma_{y_N} = \sigma_{Nx} = \sqrt{N}\sigma_x = \sqrt{N}. \tag{6.26}$$

This means that a *typical* distance from the origin is $\sqrt{N}$. This is called a *diffusion process* and is very common in the real world.

6.3.5 Bernoulli Process

Recall $\binom{N}{r} = \frac{N!}{r!(N-r)!}$ is the number of ways to choose r objects from N objects when order does not matter. Consider n independent repetitions of an event with only two possible outcomes. We typically call one outcome **success** and the other **failure**, the event a **Bernoulli Trial**, and a collection of independent Bernoulli Trials a **Bernoulli Process**.

In each Bernoulli Trial, let there be probability p of success and $q = 1 - p$ of failure. Often, we represent a success with 1 and a failure with 0.

Exercise 6.3.19. *For a Bernoulli Trial, show $\bar{x} = p$, $\sigma_x^2 = pq$, and $\sigma_x = \sqrt{pq}$.*

Let Y_N be the number of successes in N trials. Clearly, the possible values are $Y_N = \{0, 1, \cdots, N\}$. We analyze $p_N(k)$. Rigorously, the sample space Ω is all possible sequences of N trials, and the random variable $y_N : \Omega \to \mathbb{R}$ is given by $y_N(\omega)$ equals the number of successes in ω.

If $k \in Y_N$, we need k successes and $N - k$ failures. We don't care what order we have them (ie, if $k = 4$ and $N = 6$ then $SSFSSF$ and $FSSSSF$ both contribute). Each such string of k successes and $N - k$ failures has probability of $p^k \cdot (1-p)^{N-k}$. There are $\binom{N}{k}$ such strings.

Thus, $p_N(k) = \binom{N}{k} p^k \cdot (1-p)^{N-k}$ if $k \in \{0, 1, \cdots, N\}$ and 0 otherwise.

By clever algebraic manipulations, one can directly evaluate the mean $\overline{y_N}$ and the variance $\sigma_{y_N}^2$; however, Lemmas 6.3.8 and 6.3.14 allow one to calculate both quantities immediately, once one knows the mean and variance for one occurrence.

Lemma 6.3.20. *For a Bernouilli Process with N trials, each having probability p of success, the expected number of successes is $\overline{y_N} = Np$, and the variance is $\sigma_{y_N}^2 = Npq$.*

Exercise 6.3.21. *Prove Lemma 6.3.20.*

Consider the following problem: Let $Z = \{0, 1, 2, \dots\}$ be the number of trials before the £rst success. What is $\bar{z}$ and σ_z^2?

First, we determine $p(k)$, the probability that the £rst success occurs after k trials. Clearly this probability is non-zero only for k a positive integer, in which case the string of results must be $k-1$ failures followed by 1 success. Therefore,

$$p(k) \ = \ p \cdot (1-p)^{k-1} \text{ if } k \in \{1, 2, \dots\}, \text{ and 0 otherwise.} \tag{6.27}$$

To determine the mean $\bar{z}$ we must evaluate

$$\begin{aligned} \bar{z} &= \sum_{k=1}^{\infty} k \cdot p \cdot (1-p)^{k-1} \\ &= p \sum_{k=1}^{\infty} k q^{k-1}, \ 0 < q = 1 - p < 1. \end{aligned} \tag{6.28}$$

Consider the geometric series

$$f(q) \;=\; \sum_{k=0}^{\infty} q^k \;=\; \frac{1}{1-q}. \tag{6.29}$$

A careful analysis shows we can differentiate term by term if $0 \le q < 1$. Then

$$f'(q) \;=\; \sum_{k=0}^{\infty} kq^{k-1} \;=\; \frac{1}{(1-q)^2}. \tag{6.30}$$

Recalling $q = 1 - p$ and substituting yields

$$\begin{aligned}\bar{z} &= p\sum_{k=1}^{\infty} kq^{k-1} \\ &= \frac{p}{\Big(1-(1-p)\Big)^2} \;=\; \frac{1}{p}.\end{aligned} \tag{6.31}$$

Differentiating under the summation sign is a powerful tool in Probability Theory.

Exercise 6.3.22. *Calculate σ_z^2. Hint: differentiate $f(q)$ twice.*

6.3.6 Poisson Distribution

Divide the unit interval into N equal pieces. Consider N independent Bernoulli Trials, one for each sub-interval. If the probability of a success is $\frac{\lambda}{N}$, then by Lemma 6.3.20 the expected number of successes is $N \cdot \frac{\lambda}{N} = \lambda$.

We consider the limit as $N \to \infty$. Obviously, we still expect λ successes in each interval, but what is the probability of 3λ successes? How long do we expect to wait between successes?

We call this a **Poisson process with parameter** λ. For example, look at the midpoints of the N intervals. At each midpoint we have a Bernoulli Trial with probability of success $\frac{\lambda}{N}$ and failure $1 - \frac{\lambda}{N}$.

We determine the $N \to \infty$ limits. For £xed N, the probability of k successes in a unit interval is

$$
\begin{aligned}
p_N(k) &= \binom{N}{k}\left(\frac{\lambda}{N}\right)^k\left(1-\frac{\lambda}{N}\right)^{N-k} \\
&= \frac{N!}{k!(N-k)!}\frac{\lambda^k}{N^k}\left(1-\frac{\lambda}{N}\right)^{N-k} \\
&= \frac{N\cdot(N-1)\cdots(N-k+1)}{N\cdot N\cdots N}\frac{\lambda^k}{k!}\left(1-\frac{\lambda}{N}\right)^N\left(1-\frac{\lambda}{N}\right)^{-k} \\
&= 1\cdot\left(1-\frac{1}{N}\right)\cdots\left(1-\frac{k-1}{N}\right)\frac{\lambda^k}{k!}\left(1-\frac{\lambda}{N}\right)^N\left(1-\frac{\lambda}{N}\right)^{-k} \qquad (6.32)
\end{aligned}
$$

For £xed, £nite k, as $N \to \infty$, the £rst k factors in $p_N(k)$ tend to 1, $\left(1-\frac{\lambda}{N}\right)^N \to e^{-\lambda}$, and $\left(1-\frac{\lambda}{N}\right)^{-k} \to 1$.

Thus, we are led to the **Poisson Distribution**: Given a parameter λ (interpreted as the expected number of occurrences per unit interval), the probability of k occurrences in a unit interval is $p(k) = \frac{\lambda^k}{k!}e^{-\lambda}$ for $k \in \{0, 1, 2, \ldots\}$.

Exercise 6.3.23. *Check that $p(k)$ given above is a probability distribution. Namely, show $\sum_{k\geq 0} p(k) = 1$.*

Exercise 6.3.24. *Show, for the Poisson Distribution, that the mean $\bar{x} = \lambda$ and the variance $\sigma_x^2 = \lambda$. Hint: let*

$$
f(\lambda) = \sum_{k=0}^{\infty}\frac{\lambda^k}{k!} = e^\lambda. \qquad (6.33)
$$

Differentiate once to determine the mean, twice to determine the variance.

6.3.7 Continuous Poisson Distribution

We calculate a very important quantity related to the Poisson Distribution (with parameter λ), namely, how long does one expect to wait between successes?

We've discussed that we expect λ successes per unit interval, and we've calculated the probability of k successes per unit interval.

Start counting at 0, and assume the £rst success is at x. What is $p_S(x)$? As before, we divide each unit interval into N equal pieces, and consider a Bernoulli Trial at the midpoint of each sub-interval, with probability $\frac{\lambda}{N}$ of success.

We have approximately $\frac{x-0}{1/N} = Nx$ midpoints from 0 to x (with N midpoints per unit interval). Let $\lceil y \rceil$ be the smallest integer greater than or equal to y. Then we have $\lceil Nx \rceil$ midpoints, where the results of the Bernoulli Trials of the £rst $\lceil Nx \rceil - 1$ midpoints are all failures and the last is a success.

Thus, the probability of the £rst success occuring in an interval of length $\frac{1}{N}$ containing x (with N divisions per unit interval) is

$$p_{N,S}(x) \quad = \quad \left(1 - \frac{\lambda}{N}\right)^{\lceil Nx \rceil - 1} \cdot \left(\frac{\lambda}{N}\right)^{1}. \tag{6.34}$$

For N large, the above converges to $e^{-\lambda x}\frac{\lambda}{N}$.

We say $p(x)$ is a **continuous probability distribution on** $\mathbb{R}$ if

1. $p(x) \geq 0$ for all $x \in \mathbb{R}$.
2. $\int_{\mathbb{R}} p(x)dx = 1$.
3. Probability$(a \leq x \leq b) = \int_a^b p(x)dx$.

We call $p(x)$ the **probability density function**.

Thus, as $N \to \infty$, we see the probability density function $p_S(x) = \lambda e^{-\lambda x}$. In the special case of $\lambda = 1$, we get the standard exponential decay, e^{-x}.

For instance, let $\pi(M)$ be the number of primes that are at most M. The Prime Number Theorem states $\pi(M) = \frac{M}{\log M}$ plus lower order terms.

Thus, the average spacing between primes around M is about $\log M$. We can model the distribution of primes as a Poisson Process, with parameter $\lambda = \lambda_M = \frac{1}{\log M}$. While possible locations of primes (obviously) is discrete (it must be an integer, and in fact the location of primes aren't independent), a Poisson model often gives very good heuristics.

We can often renormalize so that $\lambda = 1$. This is denoted **unit mean spacing**. For example, one can show the M^{th} prime p_M is about $M \log M$, and spacings between primes around p_M is about $\log M$. Then the normalized primes, $q_M \approx \frac{p_M}{\log M}$ will have unit mean spacing and $\lambda = 1$.

6.3.8 Central Limit Theorem

$X_1, X_2, X_3, \ldots$ are an in£nite sequence of random variables such that the X_j are independent identically distributed random variables (abbreviated i.i.d.r.v.) with $E[X_j] = \bar{X}_j = 0$ (can always renormalize by shifting) and variance $E[X_j^2] = 1$. Let $S_N = \sum_{j=1}^{N} X_j$.

Theorem 6.3.25. *Fix $-\infty < a \leq b < \infty$. Then as $N \to \infty$,*

$$\textit{Prob}\Big(\frac{S_N}{\sqrt{N}} \in [a, b]\Big) \to \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{t^2}{2}} dt. \tag{6.35}$$

The probability function is called the Gaussian or the Normal distribution. This is the universal curve of probability. Note how robust the Central Limit Theorem is: it doesn't depend on £ne properties of the X_j.

6.4 Algebraic and Transcendental Numbers

6.4.1 De£nitions

A function $f : A \to B$ is **one-to-one** if $f(x) = f(y)$ implies $x = y$; f is **onto** if given any $b \in B$, $\exists a \in A$ with $f(a) = b$. f is a **bijection** if f is a one-to-one and onto function.

We say two sets A and B **have the same cardinality** (ie, are the same size) if there is a bijection $f : A \to B$. We denote this by $|A| = |B|$. If A has £nitely many elements (say n elements), A is **£nite** and $|A| = n < \infty$.

Exercise 6.4.1. *Show two £nite sets have the same cardinality if and only if they have the same number of elements.*

Exercise 6.4.2. *If f is a bijection from A to B, prove there is a bijection $g = f^{-1}$ from B to A.*

A is **countable** if there is a bijection between A and the integers $\mathbb{Z}$. A is **at most countable** if A is either £nite or countable.

Recall a binary relation R is an **equivalence relation** if

1. Re¤exive: $R(x, x)$ is true (x is equivalent to x).
2. Symmetric: $R(x, y)$ true implies $R(y, x)$ is true (if x is equivalent to y then y is equivalent to x).

3. Transitive: $R(x, y)$ and $R(y, z)$ true imply $R(x, z)$ is true (if x is equivalent to y and y is equivalent to z, then x is equivalent to z).

We often denote equivalence by $\equiv$ or $=$.

Exercise 6.4.3. *Let $x, y, z \in \mathbb{Z}$, and let $n \in \mathbb{Z}$ be given. De£ne $R(x, y)$ to be true if $n|(x - y)$ and false otherwise. Prove R is an equivalence relation. We denote it by $x \equiv y$.*

Exercise 6.4.4. *Let x, y, z be subsets of X (for example, $X = \mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{R}^n$, et cetera). De£ne $R(x, y)$ to be true if $|x| = |y|$ (the two sets have the same cardinality), and false otherwise. Prove R is an equivalence relation.*

6.4.2 Countable Sets

We show several sets are countable. Consider the set of non-negative integers $\mathbb{N}$. De£ne $f : \mathbb{N} \to \mathbb{Z}$ by $f(2n) = n$, $f(2n + 1) = -n - 1$. By inspection, we see f gives the desired bijection.

Consider $\mathbb{W} = \{1, 2, 3, \dots\}$ (the positive integers). Then $f : \mathbb{W} \to \mathbb{Z}$ de£ned by $f(2n) = n$, $f(2n + 1) = -n$ gives the desired bijection.

Thus, we have proved

Lemma 6.4.5. *To show a set S is countable, it is suf£cient to £nd a bijection from S to either $\mathbb{Z}$, $\mathbb{N}$ or $\mathbb{W}$.*

We need the intuitively plausible

Lemma 6.4.6. *If $A \subset B$, then $|A| \leq |B|$.*

We can then prove

Lemma 6.4.7. *If $f : A \to C$ is a one-to-one function (not necessarily onto), then $|A| \leq |C|$. Further, if $C \subset A$, then $|A| = |C|$.*

Exercise 6.4.8. *Prove Lemmas 6.4.6 and 6.4.7.*

If A and B are sets, the **cartesian product** $A \times B$ is $\{(a, b) : a \in A, b \in B\}$.

Theorem 6.4.9. *If A and B are countable, so is $A \cup B$ and $A \times B$.*

Proof: we have bijections $f : \mathbb{N} \to A$ and $g : \mathbb{N} \to B$. Thus, we can label the elements of A and B by

$$\begin{aligned} A &= \{a_0, a_1, a_2, a_3, \dots\} \\ B &= \{b_0, b_1, b_2, b_3, \dots\}. \end{aligned} \tag{6.36}$$

Assume $A \cap B$ is empty. De£ne $h : \mathbb{N} \to A \cup B$ by $h(2n) = a_n$ and $h(2n+1) = b_{n-1}$. We leave to the reader the case when $A \cap B$ is not empty.

To prove the second claim, consider the following function $h : W \to A \times B$:

$$\begin{aligned} & h(1) = (a_0, b_0) \\ & h(2) = (a_1, b_0), h(3) = (a_1, b_1), h(4) = (a_0, b_1) \\ & h(5) = (a_2, b_0), h(6) = (a_2, b_1), h(7) = (a_2, b_2), h(8) = (a_1, b_2), h(9) = (a_0, b_2) \\ & \quad \vdots \\ & h(n^2+1) = (a_n, b_0), h(n^2+2) = (a_n, b_{n-1}), \dots, \\ & \qquad h(n^2+n+1) = (a_n, b_n), h(n^2+n+2) = (a_{n-1}, b_n), \dots, \\ & \qquad h((n+1)^2) = (a_0, b_n) \\ & \quad \vdots \end{aligned} \tag{6.37}$$

Basically, look at all pairs of integers in the £rst quadrant (including those on the axes). Thus, we have pairs (a_x, b_y). The above function h starts at $(0, 0)$, and then moves through the £rst quadrant, hitting each pair once and only once, by going up and over. Draw the picture! □

Corollary 6.4.10. *Let A_i be countable $\forall i \in \mathbb{N}$. Then for any n, $A_1 \cup \cdots \cup A_n$ and $A_1 \times \cdots \times A_n$ are countable, where the last set is all n-tuples $(a_1, \dots, a_n)$, $a_i \in A_i$. Further, $\cup_{i=0}^{\infty} A_i$ is countable. If each A_i is at most countable, then $\cup_{i=0}^{\infty} A_i$ is at most countable.*

Exercise 6.4.11. *Prove Corollary 6.4.10. Hint: for $\cup_{i=0}^{\infty} A_i$, mimic the proof used to show $A \times B$ is countable.*

As the natural numbers, integers and rationals are countable, by taking each $A_i = \mathbb{N}$, $\mathbb{Z}$ or $\mathbb{Q}$ we immediately obtain

Corollary 6.4.12. *$\mathbb{N}^n$, $\mathbb{Z}^n$ and $\mathbb{Q}^n$ are countable. Hint: proceed by induction. For example write $\mathbb{Q}^{n+1}$ as $\mathbb{Q}^n \times \mathbb{Q}$.*

Exercise 6.4.13. *Prove there are countably many rationals in the interval $[0, 1]$.*

6.4.3 Algebraic Numbers

Consider a polynomial $f(x) = 0$ with rational coef£cients. By multiplying by the least common multiple of the denominators, we can clear the fractions. Thus, without loss of generality it is suf£cient to consider polynomials with integer coef£cients.

The **algebraic numbers**, $\mathcal{A}$, are the set of all $x \in \mathbb{C}$ such that there is a polynomial of £nite degree and integer coef£cients (depending on x, of course!) such that $f(x) = 0$. The remaining complex numbers are the **transcendentals**.

The **algebraic numbers of degree** n, $\mathcal{A}_n$, are the set of all $x \in \mathcal{A}$ such that

1. there exists a polynomial with integer coef£cients of degree n such that $f(x) = 0$

2. there is no polynomial g with integer coef£cients and degree less than n with $g(x) = 0$.

Thus, $\mathcal{A}_n$ is the subset of algebraic numbers x where for each $x \in \mathcal{A}_n$, the degree of the smallest polynomial f with integer coef£cients and $f(x) = 0$ is n.

Exercise 6.4.14. *Show the following are algebraic: any rational, the square-root of any rational, the cube-root of any rational, $r^{\frac{p}{q}}$ where $r, p, q \in \mathbb{Q}$, $i = \sqrt{-1}$, $\sqrt{3\sqrt{2} - 5}$.*

Theorem 6.4.15. *The Algebraic Numbers are countable.*

Proof: If we show each $\mathcal{A}_n$ is at most countable, then as $\mathcal{A} = \cup_{n=1}^{\infty} \mathcal{A}_n$, by Corollary 6.4.10 $\mathcal{A}$ is at most countable.

Recall the **Fundamental Theorem of Algebra (FTA):** Let $f(x)$ be a polynomial of degree n with complex coef£cients. Then $f(x)$ has n (not necessarily distinct) roots. Of course, we will only need a weaker version, namely that the Fundamental Theorem of Algebra holds for polynomials with integer coef£cients.

Fix an $n \in \mathbb{N}$. We now show $\mathcal{A}_n$ is at most countable. We can represent every integral polynomial $f(x) = a_n x^n + \cdots + a_0$ by an $(n+1)$-tuple $(a_0, \ldots, a_n)$. By Corollary 6.4.12, the set of all $(n+1)$-tuples with integer coef£cients ($\mathbb{Z}^{n+1}$) is countable. Thus, there is a bijection from $\mathbb{N}$ to $\mathbb{Z}^{n+1}$, and we can index each $(n+1)$-tuple $a \in \mathbb{Z}^{n+1}$:

$$\{a : a \in \mathbb{Z}^{n+1}\} = \bigcup_{i=1}^{\infty} \{\alpha_i\}, \tag{6.38}$$

where each $\alpha_i \in \mathbb{Z}^{n+1}$.

For each tuple α_i (or $a \in \mathbb{Z}^{n+1}$), there are n roots. Let R_{α_i} be the roots of the integer polynomial associated to α_i. The roots in R_{α_i} need not be distinct, and the roots may solve an integer polynomial of smaller degree. For example, $f(x) = (x^2-1)^4$ is a degree 8 polynomial. It has two roots, $x = 1$ with multiplicity 4 and $x = -1$ with multiplicity 4, and each root is a root of a degree 1 polynomial.

Let $R_n = \{x \in \mathbb{C} : x \text{ is a root of a degree } n \text{ polynomial}\}$. One can show that

$$R_n \;=\; \bigcup_{i=1}^{\infty} R_{\alpha_i} \;\supset\; \mathcal{A}_n. \tag{6.39}$$

By Lemma 6.4.10, R_n is countable. Thus, by Lemma 6.4.6, as R_n is at most countable, A_n is at most countable.

Therefore, each $\mathcal{A}_n$ is at most countable, so by Corollary 6.4.10 $\mathcal{A}$ is at most countable. As $\mathcal{A}_1 \supset \mathbb{Q}$ (given $\frac{p}{q} \in \mathbb{Q}$, consider $qx - p = 0$), $\mathcal{A}_1$ is at least countable. As we've shown $\mathcal{A}_1$ is at most countable, this implies $\mathcal{A}_1$ is countable. Thus, $\mathcal{A}$ is countable. □

6.4.4 Transcendental Numbers

A set is **uncountable** if there is no bijection between it and the rationals (or the integers, or any countable set).

Theorem 6.4.16. *The set of irrationals in* $[0,1]$ *is uncountable.*

Proof: Let $I = [0,1] - \mathbb{Q} = \{x : 0 \le x \le 1 \text{ and } x \notin \mathbb{Q}\}$. Assume that I is countable (the case where I is £nite is even easier).

We can write every number in I in a base two expansion, say $y = .y_1y_2y_3y_4\cdots$, $y_i \in \{0,1\}$, $y = \sum_i y_i 2^{-1}$. Certain numbers can be written two different ways. For example, $0.010011111111111\cdots = .0101$. As we are assuming I is countable, including both representations of these numbers is equivalent to taking the union of two countable sets, which by Theorem 6.4.9 is countable.

Further, we can add back all the rationals in $[0,1]$, as there are countably many rationals in $[0,1]$. Call this set S (the union of the irrationals, the alternate representation of some of the irrationals, and the rationals). As X is contained in the union of three at most countable sets (and two are countable), X is countable by Theorem 6.4.9.

There is therefore a bijection between $\mathbb{N}$ and X. We can enumerate the elements by $\{x_1, x_2, x_3, \dots\}$.

For each x_i, let $.x_{i1}x_{i2}x_{i3}\cdots x_{ii}\cdots$ be its binary expansion. We list the countable members of X:

$$\begin{aligned}
x_1 &= x_{11}x_{12}x_{13}x_{14}\cdots \\
x_2 &= x_{21}x_{22}x_{23}x_{24}\cdots \\
x_3 &= x_{31}x_{32}x_{33}x_{34}\cdots \\
&\vdots \\
x_n &= x_{n1}x_{n2}x_{n3}x_{n4}\cdots x_{nn}\cdots \\
&\vdots
\end{aligned} \tag{6.40}$$

We construct a real number $x \in [0,1]$ not in X. As this was supposed to be (more than a) complete list of all reals in $[0,1]$, this will contradict the assumption that I is countable.

Consider the number $z = .z_1z_2z_3\cdots z_n\cdots$ de£ned by $z_n = 1 - x_{nn}$. Can z be one of the numbers in our list? For example, could $z = x_m$?

No, as they differ in the m^{th} digit. Thus, z is not on our list, violating the assumption that we had a complete enumeration. Note we had to be careful and make sure we included all equivalent ways of writing the same number. Thus, while z disagrees with the base two expansion of x_m, it cannot be an equivalent way of representing x_m, as all equivalent ways of representing x_m are in our list. This is merely an annoying technical detail.

Thus, the set of irrationals in $[0,1]$ is not countable. $\Box$.

The above proof is due to Cantor $(1873 - 1874)$, and is known as **Cantor's Diagonalization Argument**. Note Cantor's proof shows that *most* numbers are transcendental, though it doesn't tell us *which* numbers are transcendental. We can easily show many numbers (such as $\sqrt{3 + 2^{\frac{3}{5}}\sqrt{7}}$) are algebraic. What of other numbers, such as π and e?

Lambert (1761), Legendre (1794), Hermite (1873) and others proved π irrational; Legendre (1794) also proved π irrational. In 1882 Lindemann proved π transcental.

What about e? Euler (1737) proved that e and e^2 are irrational, Liouville (1844) proved e is not an algebraic number of degree 2, and Hermite (1873) proved e is transcendental.

Liouville (1851) showed transcendental numbers exist; we will discuss his construction later.

6.5 Introduction to Number Theory

6.5.1 Dirichlet's Box Principle

De£nition 6.5.1 (Dirichlet's Box Principle / Pidgeon Hole Principle). *Consider n boxes, and place $n+1$ objects in the n boxes. Then some box contains at least two objects.*

We will use Dirichlet's Box Principle to £nd good rational approximations to irrational numbers.

Approximation by Rationals

Let $\alpha \in \mathbb{R} - \mathbb{Q}$ be an irrational number. We are looking for a rational number $\frac{p}{q}$ such that $\left|\alpha - \frac{p}{q}\right|$ is small, so that $\frac{p}{q}$ is a good rational approximation to α.

Lemma 6.5.2. *Let $\alpha \in \mathbb{R} - \mathbb{Q}$. Then there exist $p, q \in \mathbb{Z}, q \neq 0$ such that:*

$$\left|\alpha - \frac{p}{q}\right| \leq \frac{1}{q} \tag{6.1}$$

Proof. It is enough to prove this for $\alpha \in (0,1)$. Let $q \geq 1$ and divide the interval $[0,1)$ into q intervals $[\frac{p}{q}, \frac{p+1}{q})$ of length $\frac{1}{q}$. Then α belongs to one of these intervals. For some $0 < p < q$ we then have:

$$\alpha \in \left[\frac{p}{q}, \frac{p+1}{q}\right) \Rightarrow \left|\alpha - \frac{p}{q}\right| \leq \frac{1}{q}. \tag{6.2}$$

To obtain a better approximation, we start with an irrational number $\alpha \in (0,1)$ and an integer parameter $Q > 1$. As before, divide the interval $(0,1)$ into Q equal pieces $(\frac{a}{Q}, \frac{a+1}{Q})$. Consider the $Q+1$ numbers inside the interval $(0,1)$:

$$\{\alpha\}, \{2\alpha\}, ..., \{(Q+1)\alpha\}, \tag{6.3}$$

where $\{x\}$ denotes the fractional part of x. Letting $[x]$ denote the greatest integer less than or equal to x, we have $x = [x] + \{x\}$.

By Dirichlet's Box Principle, at least two of these numbers, say $\{q_1\alpha\}$ and $\{q_2\alpha\}$, belong to a common interval of length $\frac{1}{Q}$. Without loss of generality, we may take $1 \leq q_1 < q_2 \leq Q+1$.

Hence

$$\Big|\{q_2\alpha\} - \{q_1\alpha\}\Big| \leq \frac{1}{Q} \tag{6.4}$$

and

$$|(q_2\alpha - n_2) - (q_1\alpha - n_1)| \leq \frac{1}{Q}, \quad n_i = [q_i\alpha]. \tag{6.5}$$

Now let $q = q_1 - q_2$, $1 \leq q \leq Q$ and $p = n_1 - n_2 \in \mathbb{Z}$. Then

$$\Big|q\alpha - p\Big| \leq \frac{1}{Q} \tag{6.6}$$

and hence

$$\Big|\alpha - \frac{p}{q}\Big| \leq \frac{1}{qQ} \leq \frac{1}{q^2}. \tag{6.7}$$

We have proven

Theorem 6.5.3. *Given $\alpha \in \mathbb{R}$, there exist $p, q \in \mathbb{Z}, q \neq 0$, such that*

$$\Big|\alpha - \frac{p}{q}\Big| < \frac{1}{q^2}. \tag{6.8}$$

6.5.2 Counting the Number of Primes

Euclid

Lemma 6.5.4 (Euclid). *There are in£nitely many primes.*

Proof by contradiction: Assume there are only £nitely many primes, say $p_1, p_2, \dots, p_n$. Consider

$$x = p_1p_2..p_n + 1. \tag{6.9}$$

x cannot be prime, as we are assuming p_1 through p_n is a complete list of primes. Thus, x is composite, and divisible by a prime. However, p_i cannot divide x, as it gives a remainder of 1. Thus, x would have to be divisible by some prime not in our list, again contradicting the assumption that p_1 through p_n is a complete enumeration of the primes. □

Exercise 6.5.5. *Try, using Euclid's argument, to £nd an explicit lower bound (as weak as you like) to the function:*

$$\pi(X) = \#\{p : p \text{ is prime and } p \leq X\}. \tag{6.10}$$

Dirichlet's Theorem

Theorem 6.5.6 (Primes in Arithmetic Progressions). *Let a and b be relatively prime integers. Then there are in£nitely many primes in the progression $an + b$. Further, for a £xed a, to £rst order all relatively prime b give progressions having the same number of primes.*

Notice that the condition $(a, b) = 1$ is necessary. If $\gcd(a, b) > 1$, $an + b$ can never be prime. Dirichlet's remarkable result is that this condition is also suf£cient.

Exercise 6.5.7. *Dirichlet's theorem is not easy to prove, but try to prove it in the particular case $a = 4, b = -1$, i.e. for the arithmetic progression $4n - 1$, using an argument similar to Euclid's. Proving there are in£nitely many primes of the form $4n + 1$ is a lot harder.*

Prime Number Theorem

Theorem 6.5.8. *(Prime Number Theorem or PNT) As $X \to \infty$,*

$$\pi(X) \sim \frac{X}{\log X} \tag{6.11}$$

The Prime Number Theorem was proved in 1896 by Jacques Hadamard and Charles Jean Gustave Nicolas Baron de la Vallee Poussin. Of course, we need to quantify what $\pi(X) \sim \frac{X}{\log X}$ means. Basically, there is an error function $E(X)$ such that $|\pi(X) - \frac{X}{\log X}| \leq E(X)$, and $E(X)$ grows slower than $\frac{X}{\log X}$.

A weaker version was proved by Pafnuty Chebyshev (around 1850).

Theorem 6.5.9 (Chebyshev). *There exist explicit positive constants A and B such that, for $n > 30$:*

$$\frac{AX}{\log X} \leq \frac{\pi(X)}{X} \leq \frac{BX}{\log X}. \tag{6.12}$$

Chebyshev showed one can take $A = \log\left(\frac{2^{\frac{1}{2}}3^{\frac{1}{3}}4^{\frac{1}{4}}}{30^{\frac{1}{30}}}\right) \approx .921$ and $B = \frac{6A}{5} \approx 1.105$, which are indeed very close to 1. To highlight the method, we will use cruder arguments and prove the theorem for a smaller A and a larger B.

Chebyshev's argument uses an identity using von Mangoldt's Lambda function $\Lambda(n)$, where $\Lambda(n) = \log p$ if $m = p^k$ for some prime p, and 0 otherwise.

De£ne the function

$$T(X) = \sum_{1 \leq n \leq X} \Lambda(n) \left[\frac{X}{n}\right] = \sum_{n \geq 1} \Lambda(n) \left[\frac{X}{n}\right]. \tag{6.13}$$

Exercise 6.5.10. *Show that* $T(X) = \sum_{n \leq X} \log n$.

Now, it is easy to see (compare upper and lower sums) that

$$\sum_{n \leq X} \log n = \int_1^X \log \text{t dt} + O(\log X) = X \log X - X + O(\log X), \tag{6.14}$$

giving a good approximation to the function $T(X)$. The trick is to look at

$$T(X) - 2T\Big(\frac{X}{2}\Big) = \sum_n \Lambda(n) \left(\left[\frac{X}{n}\right] - 2\left[\frac{X}{2n}\right] \right) \tag{6.15}$$

By the previous remarks, the LHS $= X \log 2 + O(\log X)$. Also,

$$\text{RHS} \leq \sum_{p \leq X} (\log p) \frac{\log X}{\log p} = \pi(X) \log X. \tag{6.16}$$

Hence we immediately obtain the lower bound:

$$\pi(X) \geq \frac{X \log 2}{\log X} + O(\log X) \tag{6.17}$$

Exercise 6.5.11. *Prove the bound in Equation 6.16.*

To obtain an upper bound for $\pi(X)$, we notice that, since $[2\alpha] \geq 2[\alpha]$, the sum in equation (6.15) has only positive terms. By dropping terms we get a lower bound.

$$\begin{aligned} T(X) - 2T\Big(\frac{X}{2}\Big) &\geq \sum_{X/2<n\leq X} \Lambda(n)\left(\left[\frac{X}{n}\right] - 2\left[\frac{X}{2n}\right]\right) \\ &\geq \sum_{X/2<p\leq X} \log p \\ &\geq \log\Big(\frac{X}{2}\Big) \sum_{X/2<p\leq X} 1 \\ &= \log\Big(\frac{X}{2}\Big)\left(\pi(X) - \pi\Big(\frac{X}{2}\Big)\right) \qquad (6.18) \end{aligned}$$

Hence we obtain an upper bound for the number of primes between $\frac{X}{2}$ and X:

$$\pi(X) - \pi(X/2) \leq \frac{X\log 2}{\log(\frac{X}{2})} + O(1) \qquad (6.19)$$

Now, if we write inequality (6.19) for $X, \frac{X}{2}, \frac{X}{2^2}, \dots$ we get

$$\begin{aligned} \pi(X) - \pi(X/2) &\leq 2\frac{X/2}{\log(X/2)} \\ \pi(X/2) - \pi(X/2^2) &\leq 2\frac{X/2^2}{\log(X/2^2)} \\ &\vdots \\ \pi(X/2^{k-1}) - \pi(X/2^k) &\leq 2\frac{X/2^k}{\log(X/2^k)} \qquad (6.20) \end{aligned}$$

as long as $\frac{X}{2^k} \geq 1$, i.e. $k \leq [\log_2 X] = k_0$. Summing the above inequalities we get on the left hand side a telescoping sum. All the terms cancel, except for the leading term $\pi(X)$ and $\pi(X/2^{k_0}) = 0$.

Thus

$$\pi(X) \leq 2\sum_{k=1}^{k_0} \frac{X/2^k}{\log(X/2^k)} \qquad (6.21)$$

To evaluate the sum in the above inequality we split it into two parts, k "small" and k "large". More precisely, let $n_0 = \log_2(X^{1/10})$ so that $2^{n_0} = X^{1/10}$ and note that:

$$2 \sum_{k>n_0} \frac{X/2^k}{\log(X/2^k)} \leq 2 \sum_{k>n_0} \frac{X}{2^k} \leq \frac{2X}{2^{n_0}} = \frac{2X}{X^{1/10}} = 2X^{9/10}. \tag{6.22}$$

Hence the contribution from k "large" is very small compared to what we expect (i.e. order of magnitude $\frac{X}{\log X}$), or we can say that the main term comes from the sum over k small.

We now evaluate the contribution from small k.

$$2 \sum_{k=1}^{n_0} \frac{X}{2^k} \frac{1}{\log(X/2^k)} \leq \frac{2X}{\log(X/2^{n_0})} \sum_{k=1}^{n_0} \frac{1}{2^{n_0}} \leq \frac{2X}{\log(X^{9/10})} = \frac{20}{9} \frac{X}{\log X} \tag{6.23}$$

Hence the right hand side of the equation (6.21) is made up of two parts, a main term of size $\frac{BX}{\log X}$ coming from equation (6.23), and a lower order term coming from equation (6.22).

For X suf£ciently large,

$$\pi(X) \leq \frac{BX}{\log X} \tag{6.24}$$

where B can be any constant strictly bigger than $\frac{20}{9}$.

To obtain Chebyshev's better constant we would have to work a little harder along these lines, but it is the same method.

Gathering equations (6.17) and (6.24) we see we have proven

$$\frac{AX}{\log X} \leq \pi(X) \leq \frac{BX}{\log X}. \tag{6.25}$$

While this is not an asymptotic for $\pi(X)$, it does give the right order of magnitude for $\pi(X)$, namely $\frac{X}{\log X}$.

Exercise 6.5.12. *Using Chebyshev's Theorem, Prove Bertrand's Postulate: for any integer $n \geq 1$, there is always a prime number between n and $2n$.*

Chapter 7

More of an Introduction to Graph Theory

We review some basic de£nitions of Graph Theory, and prove a simple result about the size of the eigenvalues of adjacency graphs. Lecture by Peter Sarnak; notes by Steven J. Miller.

7.1 De£nitions

De£nition 7.1.1. *For a graph G, let $V(G)$ be the set of vertices and $E(G)$ the set of edges (an edge is a pair (v, w) with $v, w \in V$). We often just write V and E.*

De£nition 7.1.2 (Connected Graph). *A graph G is connected if for any two vertices $v, w \in G$, there is a path of edges in E starting at v and ending at w.*

De£nition 7.1.3 (Boundary). $\partial A = \{v \in V - A$ *: there is a* $w \in A$ *with* $(v, w) \in E\}$.

De£nition 7.1.4 (k-regular). *A graph G is k-regular if there are k edges coming out from each vertex.*

De£nition 7.1.5 (Expander Graph). *Let G be a connected graph with n vertices. Let $A \subset V(G)$ be any subset of vertices with $|A| \leq \frac{n}{2}$ (ie, at most half of the vertices). We say G is a (n, c, k) expander if for any such A, $|\partial A| \geq c|A|$.*

Example 7.1.6. *Let G be a 2-regular graph with n vertices. For de£nateness, label the vertices $\{1, 2, \dots, n\}$. Let the edges be $(1, 2)$, $(2, 3)$, $(3, 4), \dots, (n, 1)$.*

Let A be the £rst half of the vertices: $A = \{1, 2, \ldots, \left[\frac{n}{2}\right]\}$. Then $|\partial A| = 2$, and G is not an $(n, c, 2)$ expander.

7.2 Size of Eigenvalues

Consider a 3-regular graph with n vertices. We want the graph to have certain desirable connectivity properties.

Let $A = (a_{vw})$ is the adjacency matrix attached to the graph G, and $a_{vw} = 1$ if there is an edge from v to w and 0 otherwise. This is a very sparse matrix (only three non-zero entries in each row or column). Compute its eigenvalues (real numbers).

Lemma 7.2.1. *The eigenvalues $\lambda_i \in [-3, 3]$.*

Proof: Let $f : V \to \mathbb{R}$, de£ne the action of the adjacency matrix A on f by

$$Af(v) = \sum_{(v,w)\in E} f(w). \tag{7.1}$$

As there are £nitely many vertices (n, in fact), we can regard the function $f(v)$ as living in $\mathbb{R}^n$, with coordinates $\Big(f(v_1), \ldots, f(v_n)\Big)$.

Suppose $\forall w \in V$, $Af(w) = \lambda f(w)$, $f \neq 0$. Then λ is an eigenvalue (which must be real as A is real symmetric).

We use the Maximum Modulus Principle. As $f \neq 0$, let w_0 be such that $f(w_0)$ is the maximum value of $f(w)$ (not zero, and exists as we have £nitely many vertices). Then

$$f(w_0) = \frac{1}{\lambda} \sum_{(w,w_0)} f(w). \tag{7.2}$$

If $\lambda > 3$, this cannot happen (we're assuming the graph is k-regular and connected). If $\lambda = 3$, then f is constant. Working with absolute values, we similarly obtain $\lambda > -3$.

Exercise 7.2.2. *Prove a k-regular graph G is connected if and only if $\lambda = k$ is a simple eigenvalue.*

Let λ_1 be the second largest eigenvalue; $\lambda = k$ is always an eigenvalue for a k-regular connected graph. How big can the gap be between λ_1 and k?

Chapter 8

Linear Algebra Review, especially Spectral Theorem for Real Symmetric Matrice

We review some basic facts about Linear Algebra and Matrix Groups, and give an introduction to Random Matrix Theory. Lecture by Steven J. Miller; notes by Steven J. Miller and Alex Barnett.

8.1 Linear Algebra Review

Matrices can either be thought of as rectangular (often square) arrays of numbers, or as linear transformations from one space to another (or possible to the same space). The former picture is the simplest starting point, but as Professor Sarnak emphasized, it is the latter, geometric view that gives a deeper understanding.

To connect with the simpler vector case, a vector can be thought of as a list of real numbers which change in a certain way when the coordinate system changes, or as a geometric object with length and direction. The latter object is *coordinate-independent*, and has different representations in different choices of coordinate axes. Try to keep the geometric picture in mind for matrices.

8.1.1 De£nitions

Given an $n \times m$ matrix A (where n is the number of rows and m is the number of columns), the **transpose of** A, denoted A^T, is the $m \times n$ matrix where the rows of

A^T are the columns of A (or, equivalently, the columns of A^T are the rows of A).

Lemma 8.1.1. $(AB)^T = B^T A^T$ *and* $(A^T)^T = A$.

We leave the proof to the reader.

If an $n \times n$ matrix (also called a **square** matrix) A satis£es $A^T = A$, then we say A is **symmetric**

Example 8.1.2. *Let A be the matrix*

$$\begin{pmatrix} 2 & 2 & 4 & 2 \\ 1 & 1 & -2 & 2 \\ -2 & 0 & 0 & 1 \\ 1 & 1 & 2 & 1 \end{pmatrix} \tag{8.1}$$

Then A^T is

$$\begin{pmatrix} 2 & 1 & -2 & 1 \\ 2 & 1 & 0 & 1 \\ 4 & -2 & 0 & 2 \\ 2 & 2 & 1 & 1 \end{pmatrix} \tag{8.2}$$

Note the above matrix is not symmetric.

The number of *degrees of freedom* in a symmetric matrix (*i.e.* independent real numbers needed to completely specify the matrix) is $n(n+1)/2$. Why? There are n^2 entries, n on the diagonal. If you specify all entries above the diagonal and all entries on the diagonal, then you know the symmetric matrix.

There are $n^2 - n$ non-diagonal entries (half above the diagonal, half below). Thus, one needs to specify $\frac{n^2-n}{2} + n = \frac{n^2+n}{2}$ entries.

Exercise 8.1.3. *If A and B are symmetric, show AB is symmetric.*

Matrix multiplication. We call the element in the i^{th} row and j^{th} column a_{ij}. Think of $i = 1 \cdots n$ going down the left side, and $j = 1 \cdots M$ going across the top. A vector v we represent as a column of elements with the i^{th} being v_i. A nice way to see matrix-vector multiplication is that the v_i give the *coef£cients* by which the columns of A are linearly mixed together. For the product $w = Av$ to make sense, the length (dimension) of v must equal m, and the dimension of w will be n. A is therefore a linear transform from m-dim space to n-dim space.

Multiple transformations appear written backwards: if we apply A then B then C to a vector, we write

$$w = CBAv. \tag{8.3}$$

Note that taking the product of two $n \times n$ matrices requires $O(n^3)$ effort.

Exercise 8.1.4. *Show that there are two ways to evaluate triple matrix products of the type CBA. The slow way involves $O(n^4)$ effort. How about the fast way? How do these results scale for the case of a product of k matrices?*

De£nition 8.1.5 (Invertible Martices). *A is invertible if a matrix B can be found such that $BA = AB = I$. The inverse is then written $B = A^{-1}$. Invertibility requires A to be square.*

Transformations of a matrix. Just as with vectors, we can £nd out how the components of a square matrix A change under transformation. Say we have a scalar quantity $x = w^T Av$. We transform our coordinate system linearly such that the vector v has components $v' = Mv$, where M is some invertible matrix representing the transformation. Therefore also $w' = Mw$. The only way that x can remain unchanged by the transformation (as any scalar must), for all choices of v and w, is if the transformed matrix is written $A' = M^{-T}AM^{-1}$. Check this via

$$x' = w'^T A'v' = (Mw)^T(M^{-T}AM^{-1})(Mv) = w^T IAIv = w^T Av = x. \quad (8.4)$$

This is called a *similarity transformation*, or a *conjugation*. Really we have one object, the transformation A, but it may have different representations by a matrix of numbers, depending on the choice of basis.

De£nition 8.1.6 (Orthogonal Matrices). *Q is an orthogonal $n \times n$ matrix if it has real entries and $Q^TQ = QQ^T = I$.*

Q is invertible, with inverse Q^T. The geometric meaning of Q is a *rotation*: the vector $w = Qv$ is just v rotated (about the origin).

The number of degrees of freedom in an orthogonal matrix is $n(n-1)/2$.

Exercise 8.1.7. *In 3 dimensions a general rotation involves 3 angles (for example, azimuth, elevation, and 'roll'). How many angles are needed in 4 dimensions? In 3d you rotate about a line-like axis (the set of points which do not move under rotation); what object do you rotate about in 4d?*

Exercise 8.1.8. *Show that the identity matrix I, always has representation $I_{ij} = \delta_{ij}$ regardless of the choice of basis. Hint: perform orthogonal tranformation on the matrix δ_{ij}.*

The set of orthogonal matrices of order n forms a *continuous* (or *topological*) group, which we call $O(n)$. (Not to be confused with "of order N"). Group properties:

- Associativity follows from that of matrix multiplication.
- The identity matrix acts as an identity element, since it is in the group.
- Inverse is the transpose (see above): $Q^{-1} = Q^T$.
- Closure is satis£ed because any product QR of orthogonal matrices is itself orthogonal.

Exercise 8.1.9. *Prove the last assertion.*

However, not all the elements of $O(n)$ can 'talk' to each other, *i.e.* you cannot reach all the elements by continuous transformation from the identity I.

Example for $n = 2$: a general order-2 orthogonal matrix can be written

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} \cos\theta & -\sin\theta \\ -\sin\theta & -\cos\theta \end{pmatrix}, \tag{8.5}$$

where $0 \leq \theta < 2\pi$ is a real angle. The £rst has determinant $+1$ and de£nes the 'special' (*i.e.* unit determinant) group $SO(2)$ which is a subgroup of $O(2)$ with identity I. The second has determinant -1 and corresponds to rotations with a re¤ection; this subgroup is disjoint from $SO(2)$, and has the weird (re¤ecting) identity can be written in some basis as

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{8.6}$$

Note that $SO(2)$, alternatively written as the family of planar rotations $R(\theta)$, is *isomorphic* to the unit length complex numbers under the multiplication operation:

$$R(\theta) \;\longleftrightarrow\; e^{i\theta}. \tag{8.7}$$

Therefore we have $R(\theta_1)R(\theta_2) = R(\theta_1 + \theta_2)$. This commutativity relation does *not* hold in higher $n > 2$.

Orthogonal transformations. If an orthogonal matrix Q is used for conjugation of a general square matrix A, then the rule Eq. 8.4 for transformation becomes,

$$A' \;=\; QAQ^T. \tag{8.8}$$

This tells you how to 'rotate' a (square) matrix.

De£nition 8.1.10 (Complex Conjugate Transpose). *Let A be an $n \times m$ matrix. Then the complex conjugate transpose of A, denoted A^*, is obtained by the following: (1) take the complex conjugate of A; ie, replace every entry $a_{jk} = x_{jk} + iy_{jk}$ with $\overline{a_{jk}} = x_{jk} - iy_{jk}$, and call this matrix A_1; (2) take the transpose of A_1.*

Exercise 8.1.11. *Prove that $(AB)^* = B^* A^*$.*

De£nition 8.1.12 (Dot or Inner Product). *The dot (or inner) product of two real vectors v and w is de£ned as $v^T w$; if the vectors are complex, we instead use $v^* w$.*

Exercise 8.1.13. *Show that the dot product is invariant under orthogonal transformation. That is, show that given two vectors, transforming them using the same orthogonal matrix leaves their dot product unchanged.*

De£nition 8.1.14 (Length of a vector). *The length of a real vector v is $|v|^2 = v^T v$; for a complex vector, we have $|v|^2 = v^* v$.*

De£nition 8.1.15 (Orthogonality). *Two real vectors are orthogonal (also called perpendicular) if $v^T w = 0$; for two complex vectors, the equivalent condition is $v^* w = 0$.*

De£nition 8.1.16 (Eigenvalue, Eigenvector). *Recall λ is an **eigenvalue** and v is an **eigenvector** if $Av = \lambda v$ and v is not the zero vector.*

Exercise 8.1.17. *If v is an eigenvector of A with eigenvalue λ, show $w = av$, $a \in \mathbb{C}$, is also an eigenvector of A with eigenvalue λ.*

Exercise 8.1.18. *Show that given an eigenvalue λ and an eigenvector v, you can always £nd an eigenvector w with the same eigenvalue, but $|w| = 1$.*

To £nd the eigenvalues, we solve the equation $\det(\lambda I - A) = 0$. This gives a polynomial $p(\lambda) = 0$. We call $p(\lambda)$ the **characteristic polynomial**.

The **trace** of a matrix A, denote $\text{Tr}(A)$ is the sum of the diagonal entries of A:

$$\text{Tr}(A) = \sum_{i=1}^{n} a_{ii}. \tag{8.9}$$

Lemma 8.1.19. *$Tr(A) = \sum_{i=1}^{n} \lambda_i$.*

The proof relies on writing out the characteristic equation and comparing powers of λ with the factorized version. By the fact that the polynomial has roots λ_i we can write

$$\det(\lambda I - A) \;=\; p(\lambda) \;=\; \prod_{i=1}^{n}(\lambda - \lambda_i). \tag{8.10}$$

Note the coef£cient of λ^n is 1, thus we have $\prod_i(\lambda - \lambda_i)$ and not $c\prod_i(\lambda - \lambda_i)$ for some constant c.

By expanding out the RHS, the coef£cient of λ^{n-1} is $-\sum_{i=1}^{n}\lambda_i$, which we will show is $-\mathrm{Tr}(A)$. Expanding the LHS, we want to £nd the corresponding coef£cient in

$$\begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & & \\ \vdots & & \ddots & \\ -a_{n1} & & & \lambda - a_{nn} \end{vmatrix}.$$

We have to remember the expansion of the determinant. Taking the top-left-most 2×2 block, we see its determinant is $(\lambda - a_{11})(\lambda - a_{22}) - a_{12}a_{21} = \lambda^2 - (a_{11} + a_{22})\lambda + (a_{11}a_{22} - a_{12}a_{21})$. The determinant of the top-left-most 3×3 block is then formed by $(\lambda - a_{33})$ times the above 2×2 determinant, plus two other multiples of determinants which can give only a highest power of λ of λ^1. Thus we see that the coef£cient in λ^2 is $-(a_{11} + a_{22} + a_{33})$. Repeating this argument for 4×4 block up to $n \times n$ gives us the coef£cient of λ^{n-1} in the full determinant is $-\sum_{i=1}^{n} a_{ii}$. Since the LHS and RHS must be equal $\forall\lambda$, the LHS and RHS coef£cients in λ^{n-1} are equal. $\square$

Corollary 8.1.20. *Tr(A) is invariant under rotation of basis.*

The proof follows immediately from the invariance of the eigenvalues under rotation of basis. We need the following:

Lemma 8.1.21. $\det(AB) = \det(A)\det(B)$. *Further, by induction one can show* $\det(AB\cdots Z) = \det(A)\det(B)\cdots\det(Z)$. *Further,* $\det(I) = 1$.

Proof of Corollary: Let $A = Q^T BQ$. We show A and B have the same trace by showing A and B have the same eigenvalues. To £nd the eigenvalues of A we must solve:

$$
\begin{aligned}
\det(\lambda I - A) &= \det(\lambda I - Q^T B Q) \\
&= \det(\lambda Q^T Q - Q^T B Q) \\
&= \det(Q^T \lambda I Q - Q^T B Q) \\
&= \det\Big(Q^T(\lambda I - B)Q\Big) \\
&= \det(Q^T)\det(\lambda I - B)\det(Q) \\
&= \det(Q^T)\det(Q)\det(\lambda I - B) \\
&= \det(Q^T Q)\det(\lambda I - B) = \det(I)\det(\lambda I - B) = \det(\lambda I - B).
\end{aligned}
\tag{8.11}
$$

As the eigenvalues of A and B satisfy the same equation, they are equal. □

8.1.2 Spectral Theorem for Real Symmetric Matrices

The main theorem we will prove is

Theorem 8.1.22 (Spectral Theorem). *Let A be a real symmetric $n \times n$ matrix. Then there exists an orthogonal $n \times n$ matrix Q and a diagonal matrix Λ such that $Q^T A Q = \Lambda$. Moreover, the n eigenvalues of A are the diagonal entries of Λ.*

This result is remarkable: it tells you that any real, symmetric matrix is diagonal when rotated into an appropriate basis (recall the rotation effect of conjugation using Q). In other words, the operation of matrix A on a vector v can be broken down into three steps:

$$
Av = Q\Lambda Q^T v = (\text{undo the rotation})(\text{stretch along coord axes})(\text{rotation})v.
\tag{8.12}
$$

Recall the ordering of transformations is read like Hebrew, right to left. The rotation is just the rotation into the basis of eigenvectors.

Furthermore, the eigenvalues λ_i (= diag els of Λ) are a set of numbers invariant under rotations of A. In other words, if $A' = PAP^T$ is an orthogonally-conjugated (*i.e.* P is orthogonal) version of A, then A' has the same $\{\lambda_i\}$ as A. Of course the ordering of the λ_i has to be chosen the same.

For the Spectral Theorem we prove a sequence of needed lemmas:

Lemma 8.1.23. *The eigenvalues of a real symmetric matrix are real.*

Let A be a real symmetric matrix with eigenvalue λ and eigenvector v. Note that we do not yet know that v has only real coordinates!

Therefore, $Av = \lambda v$. Take the dot (or inner) product of both sides with the vector v^*, the complex conjugate transpose of v:

$$v^*Av = \lambda v^*v. \tag{8.13}$$

But the left hand side is real. The two sides are clearly complex numbers (ie, 1-dimensional matrices). Taking the complex conjugate transpose of the LHS gives

$$\Big(v^*(Av)\Big)^* = (Av)^*(v^*)^* = v^*Av. \tag{8.14}$$

Therefore, the LHS is real, implying the RHS is real. But clearly v^*v is real (similar calculation). Thus, λ is real. $\Box$

We will only prove the Spectral Theorem when all the eigenvalues are distinct. Henceforth, we shall always assume A is a real symmetric matrix.

Lemma 8.1.24. *The eigenvectors of a real symmetric matrix are real.*

The eigenvalues solve the equation $(\lambda I - A)v = 0$. Let λ be an eigenvalue. Then $\det(\lambda I - A) = 0$. Therefore the matrix $B = \lambda I - A$ is not invertible. Therefore it send a vector to 0 (standard linear algebra calculation).

Lemma 8.1.25. *If λ_1 and λ_2 are two distinct eigenvalues of a real symmetric matrix A, then their corresponding eigenvectors are perpendicular.*

We study $v_1^T A v_2$. Now

$$v_1^T A v_2 = v_1^T(Av_2) = v_1^T(\lambda_2 v_2) = \lambda_2 v_1^T v_2. \tag{8.15}$$

Also,

$$v_1^T A v_2 = v_1^T A^T v_2 = (v_1^T A^T) v_2 = (Av_1)^T v_2 = (\lambda_1 v_1)^T v_2 = \lambda_1 v_1^T v_2. \tag{8.16}$$

Therefore

$$\lambda_2 v_1^T v_2 = \lambda_1 v_1^T v_2 \text{ or } (\lambda_1 - \lambda_2) v_1^T v_2 = 0. \tag{8.17}$$

As $\lambda_1 \neq \lambda_2$, $v_1^T v_2 = 0$. Thus, the eigenvectors v_1 and v_2 are perpendicular. $\Box$

We can now prove the Spectral Theorem for real symmetric matrices **if there are n distinct eigenvectors**.

Let λ_1 to λ_n be the n distinct eigenvectors, and let v_1 to v_n be the corresponding eigenvectors chosen so that each v_i has length 1.

Consider the matrix Q, where the £rst column of Q is v_1, the second column of Q is v_2, all the way to the last column of Q which is v_n:

$$Q = \begin{pmatrix} \uparrow & \uparrow & & \uparrow \\ v_1 & v_2 & \cdots & v_n \\ \downarrow & \downarrow & & \downarrow \end{pmatrix} \tag{8.18}$$

The transpose of Q is

$$Q^T = \begin{pmatrix} \leftarrow & v_1 & \rightarrow \\ & \vdots & \\ \leftarrow & v_n & \rightarrow \end{pmatrix} \tag{8.19}$$

Exercise 8.1.26. *Show that Q is an orthogonal matrix. Use the fact that the v_i all have length one, and are orthogonal (perpendicular) to each other.*

Consider Q^TAQ. This is a matrix, call it B. To £nd its entry in the i^{th} row and j^{th} column, we look at

$$e_i^T B e_j \tag{8.20}$$

where the e_k are column vectors which are 1 in the k^{th} position and 0 elsewhere:

$$e_k = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{8.21}$$

Thus, we need only show that $e_i^T B e_j = 0$ if $i \neq j$ and equals λ_j if $i = j$.

Exercise 8.1.27. *Show $Qe_i = v_i$ and $Q^T v_i = e_i$.*

We calculate

$$\begin{aligned}
e_i^T B e_j &= e_i^T Q^T A Q e_j \\
&= (e_i^T Q^T) A (Q e_j) \\
&= (Q e_i)^T A (Q e_j) \\
&= v_i^T A v_j \\
&= v_i^T (A v_j) \\
&= v_i^T \lambda_j v_j = \lambda_j v_i^T v_j. \qquad (8.22)
\end{aligned}$$

As $v_i^T v_j$ equals 0 if $i \neq j$ and 1 if $i = j$, this proves the claim.

Thus, the off-diagonal entries of $Q^T A Q$ are zero, and the diagonal entries are the eigenvalues λ_j. This shows that $Q^T A Q$ is a diagonal matrix whose entries are the n eigenvalues of A. $\Box$

Note that, in the case of n distinct eigenvalues, not only can we write down the diagonal matrix, we can easily write down what Q should be. Further, by reordering the columns of Q, we see we reorder the positioning of the eigenvalues on the diagonal.

Chapter 9

Central Limit Theorem, Spectral Theorem for Real Symmetric, Spectral Gaps

Proof of the Central Limit Theorem (via the Fourier Transform). Proof of the Spectral Theorem for Real Symmetric Matrices (via maximization). Spectral Gap and Families of Expanders. Lecture by Peter Sarnak; notes by Steven J. Miller.

9.1 Central Limit Theorem

Let $X : \Omega \to \mathbb{R}$ be a random variable ($x(\omega) \in \mathbb{R}$). De£ne the density function μ on $\mathbb{R}$ by

$$\mu[a, b] = \text{Prob}\Big(\omega : x(\omega) \in [a, b]\Big). \tag{9.1}$$

Thus, X gives rise to the probability measure μ on $\mathbb{R}$.
Assume

1. $X_1, X_2, \dots$ are independent identically distributed random variables (iidrv) with density μ.
2. $E(X) = \int_\Omega x(\omega)dp(\omega) = \int_\mathbb{R} xd\mu(x)$.
3. Var$(X) = \int_\Omega x^2(\omega)dp(\omega) = \int_R x^2 d\mu(x) = 1$.

Define $S_N = \sum X_j$, and let $N(0,1)$ denote the standard Gaussian or normal distribution (mean zero, variance one).

Theorem 9.1.1.

$$\frac{S_N}{\sqrt{N}} \quad \underset{in\, probability}{\longrightarrow} \quad N(0,1), \tag{9.2}$$

ie,

$$Prob\Big(\frac{S_N}{\sqrt{N}} \in [a,b]\Big) \to \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx. \tag{9.3}$$

Proof: We have a random variable X, induces a probability measure μ on the real line $\mathbb{R}$. Thus, $d\mu = f(x)dx$, $f(x)$ is a nice function. As μ is a probability measure on $\mathbb{R}$, we must have $f(x) \geq 0$ and $\int_R f(x)dx = 1$.

The given assumptions about the X_is imply

1. $\int_R xf(x)dx = 0$.
2. $\int_R x^2 f(x)dx = 1$.

Definition 9.1.2 (Fourier Transform).

$$\hat{f}(\xi) \;=\; \int_R f(x)e^{-2\pi i x\xi} dx. \tag{9.4}$$

Clearly, $|\hat{f}(\xi) \leq \int_R f(x)dx \leq 1$. Further, $\hat{f}(0) = \int_R f(x)dx = 1$.
Now

$$\hat{f}'(\xi) = \int_R (2\pi i x) f(x) e^{-2\pi i x\xi} dx. \tag{9.5}$$

Thus, $\hat{f}'(0) = 0$ (from $E(x) = 0$).
We will assume $\hat{f}$ is continuous (although this is implied by our assumptions).
Further,

$$\hat{f}''(\xi) = -4\pi^2 \int_R x^2 f(x) e^{-2\pi i x\xi} dx. \tag{9.6}$$

Therefore, $\hat{f}''(0) = -4\pi^2$ (by our assumption on the variance).
Using Taylor to expand $\hat{f}$ we obtain

$$\begin{aligned}\hat{f}(\xi) &= 1 + \frac{f''(0)}{2}\xi^2 + \cdots \\ &= 1 - 2\pi^2\xi^2 + \cdots \end{aligned} \tag{9.7}$$

Near the origin, the above shows $\hat{f}$ looks like a concave down parabola.

Theorem 9.1.3 (Fourier Inversion). *If f is a nice function (so all the quantities below make sense, ie, f decays fast enough):*

$$f(x) = \int_R \hat{f}(\xi)e^{2\pi i x\xi}d\xi. \tag{9.8}$$

De£nition 9.1.4. $e(y) = e^{2\pi i y}$.

Exercise 9.1.5. *If $\phi_\xi(x+y) = e\Big((x+y)\xi\Big)$, prove $\phi_\xi(x+y) = \phi_\xi(x)\phi_\xi(y)$. ϕ_ξ is a **character** of $(\mathbb{R}, +)$.*

Exercise 9.1.6. *Is there a $\psi : \mathbb{R} \to \mathbb{R}$ such that $\psi(x+y) = \psi(x) + \psi(y)$? IE, can you £nd a homomorphism that takes addition to addition? If yes (of course: see above!) what can you say about ψ? If we assume ψ is continuous, it must be of the form ϕ_ξ.*

Think of $\mathbb{R}$ as a vector space over $\mathbb{Q}$ (ie, the scalars are $\mathbb{Q}$). What is the dimension of $\mathbb{R}$ over $\mathbb{Q}$, and what is a basis? As the reals are uncountable and the rationals are countable, there are uncountably many basis vectors.

*Any linear transformation will satisfy the desired condition! For example, choose any basis (called a **Hamel Basis**) of $\mathbb{R}$ over $\mathbb{Q}$ (very hard! need the Axiom of Choice).*

To show every character of the reals is of the form ϕ_ξ, you need more. If you assume the character is continous, then it must be of the form ϕ_ξ.

Suppose we have random variables X and Y with measures μ and ν, with induced functions $f(x)$ and $g(x)$. If we choose X and Y independently, what is the distribution of $X+Y$?

Lemma 9.1.7. *The distribution of $X+Y$ clearly cannot be the sum (as that won't be a probability measure). It is $f * g$, and $f * g$ is a probability measure.*

De£nition 9.1.8 (Convolution).

$$(f * g)(x) = \int_R f(x-y)g(y)dy \tag{9.9}$$

We show the convolution is a probability measure. We assume our functions f and g are nice (ie, that we may use Fubini to interchange the order of integration).

$$\begin{aligned}
\int_R (f*g)(x)dx &= \int_R\int_R f(x-y)g(y)dydx \\
&= \int_R\int_R f(x-y)g(y)dxdy \\
&= \int_R g(y)\Bigg(\int_R f(x-y)dx\Bigg)dy \\
&= \int_R g(y)\Bigg(\int_R f(t)dt\Bigg)dy \\
&= \int_R g(y)\cdot 1dy = 1. \qquad (9.10)
\end{aligned}$$

Exercise 9.1.9. $\widehat{f*g}(\xi) = \hat{f}(\xi)\cdot\hat{g}(\xi)$. *Thus, Fourier Transform converts convolution to multiplication.*

Let $Z = X + Y$. What is

$$\text{Prob}\Big(Z \in [z, z+dz]\Big) = h(z)dz. \qquad (9.11)$$

x can be anywhere; given x, y must lie between $z - x$ and $z - x + dz$.

$$\begin{aligned}
\text{Prob}\Big(X+Y \in [z,z+dz]\Big) &= \int_{x\in\mathbb{R}} \text{Prob}\Big(X \in [x, x+dx] \text{ and} \\
&\qquad\qquad Y \in [z-x, z-x+dz]\Big) f(x)dx \\
&= \int_{x\in\mathbb{R}} f(x)g(z-x)dx = h(z), \qquad (9.12)
\end{aligned}$$

where the last step follows from the de£nition of the density functions.

By Induction, we see $X_1 + \cdots X_N$ has distribution $f * \cdots * f$ (as the random variables are iidrv).

However, we want to study $S_N = \frac{X_1+\cdots X_N}{\sqrt{N}}$.

De£nition 9.1.10 (FT). *Let $FT(f) = \hat{f}$; FT denotes the Fourier Transform.*

Lemma 9.1.11.

$$FT\left(\text{the distribution of } \frac{X_1 + \cdots + X_N}{\sqrt{N}}\right) = \left[\hat{f}\Big(\frac{\xi}{\sqrt{N}}\Big)\right]^N. \tag{9.13}$$

We take the limit as $N \to \infty$ for **£xed** ξ. Recall we showed that $\hat{f}(\xi) = 1 - 2\pi^2\xi^2 + \cdots$. Thus, we have to study

$$\left[1 - \frac{2\pi^2\xi^2}{N} + O\left(\frac{|\xi|^3}{N^{\frac{3}{2}}}\right)\right]^N. \tag{9.14}$$

But as $N \to \infty$, the above goes to

$$e^{-2\pi\xi^2}. \tag{9.15}$$

The universality arises because *only* terms up to the second order in the Taylor Series contribute.

Exercise 9.1.12. *Show the Fourier Transform of the Gaussian is the Gaussian.*

Key point:

- Used Fourier Analysis to study the sum of independent identically distributed random variables, as it converts convolution to multiplication.

Exercise 9.1.13. *Fix g a nice, smooth, rapidly decreasing function. Consider the linear transformation A:*

$$(Af)(x) = \int_R g(x-y)f(y)dy. \tag{9.16}$$

Any convolution operator is diagonalized by these characters (exponentials). At a formal level,

$$\begin{aligned}
(A\phi_\xi)(x) &= \int_R g(x-y)\phi_\xi(y)dy \\
&= \int_R g(x-y)e^{-2\pi i\xi y}dy \\
&= \int_R g(t)e^{-2\pi i\xi(x-t)}dt \\
&= e^{-2\pi ix\xi}\int_R g(t)e^{2\pi i\xi t}dt = \hat{g}(\xi)\phi_\xi(x).
\end{aligned} \tag{9.17}$$

Thus, ϕ_ξ is an eigenvector of A with eigenvalue $\hat{g}(\xi)$.

9.2 Spectral Theorem for Real Symmetric Matrices

Let A be a real symmetric matrix acting on $\mathbb{R}^n$. Then A has an orthonormal basis $v_1, \ldots, v_n$ such that $Av_j = \lambda_j v_j$.

A simpler proof, assuming all eigenvalues are distinct, is available in the September 25^{th} *lecture notes.*

Write the inner or dot product $\langle v, w\rangle = v^t w$. As A is symmetric, $\langle Av, w\rangle = \langle v, Aw\rangle$.

De£nition 9.2.1. $V^\perp = \{w : \forall v \in V, \langle w, v\rangle = 0\}$.

Lemma 9.2.2. *Suppose* $V \subset \mathbb{R}^n$ *is an* ***invariant vector subspace*** *under* A *(if* $v \in V$*, then* $Av \in V$*). Then* $V^\perp$ *is also* A*-invariant:* $A(V^\perp) \subset V^\perp$.

This proves the spectral theorem. Suppose we £nd a $v_0 \neq 0$ such that $Av_0 = \lambda_0 v_0$. Take $V = \{\mu v_0 : \mu \in \mathbb{R}\}$ for the invariant subspace.

By Lemma 9.2.2, $V^\perp$ is left invariant under A, and is one dimension less. Thus, by whatever method we used to £nd an eigenvector, we apply the same method on $V^\perp$.

Thus, all we must show is given an A-invariant subspace, there is an eigenvector.

Consider

$$\max_{v \ with \ \langle v,v\rangle = 1} \Big\{ \langle Av, v\rangle \Big\}. \tag{9.18}$$

Standard fact: every continuous function on a compact set attains its maximum (not necessarily uniquely). See, for example, W. Rudin, *Principles of Mathematical Analysis*.

Let v_0 be a vector giving the maximum value, and denote this maximum value by λ_0. As $\langle v_0, v_0\rangle = 1$, v_0 is not the zero vector.

Lemma 9.2.3. $Av_0 = \lambda_0 v_0$.

Clearly, if Av_0 is a multiple of v_0 it has to be λ_0 (from the de£nition of v_0 and λ_0).

Thus, it is suf£cient to show

Lemma 9.2.4. $\{\mu v_0 : \mu \in \mathbb{R}\}$ *is an* A*-invariant subspace.*

Proof: let w be an arbitrary vector perpendicular to v_0, and ϵ be an arbitrary small real number. Consider

$$\langle A(v_0 + \epsilon w), v_0 + \epsilon w \rangle \tag{9.19}$$

We need to renormalize, as $v_0 + \epsilon w$ is not unit length; it has length $1 + \epsilon^2 \langle w, w \rangle$. As v_0 was chosen to maximize $\langle Av, v \rangle$ subject to $\langle v, v \rangle = 1$, after normalizing the above cannot be larger. Thus,

$$\langle A(v_0 + \epsilon w), v_0 + \epsilon w \rangle \;=\; \langle Av_0, v_0 \rangle + 2\epsilon \langle Av_0, w \rangle + \epsilon^2 \langle w, w \rangle. \tag{9.20}$$

Normalizing the vector $v_0 + \epsilon w$ by its length, we see that in Equation 9.20, the order ϵ terms must be zero. Thus,

$$\langle Av_0, w \rangle = 0; \tag{9.21}$$

however, this implies Av_0 is in the space spanned by v_0 (as w was an arbitrary vector perpendicular to v_0), completing our proof. □

Corollary 9.2.5. *Any local maximum will lead to an eigenvalue-eigenvector pair.*

The second largest eigenvector (denoted λ_1) is

$$\lambda_1 \;=\; \max_{\langle v, v_0 \rangle = 0} \frac{\langle Av, v \rangle}{\langle v, v \rangle}. \tag{9.22}$$

We can either divide by $\langle v, v \rangle$, or restrict to unit length vectors.

9.3 Applications to Graph Theory

Let G be a k-regular graph, $f : V \to \mathbb{R}$. Recall $v \sim w$ if there is an edge connecting v and w. Let A be the adjacency matrix of G, and de£ne

$$\begin{aligned} Af(v) &= \sum_{v \sim w} f(w) \\ \langle f, g \rangle &= \sum_{v \in V} f(v)\overline{g}(v). \end{aligned} \tag{9.23}$$

Consider the function $f_0(v) = 1$ for all $v \in V$. Then

$$Af_0(v) = \sum_{v \sim w} f_0(w) = kf_0(v). \tag{9.24}$$

Thus, f_0 is an eigenfunction with eigenvalue k.

Theorem 9.3.1 (Expander Families). *Fix k. Suppose we have a sequence of k-regular graphs G_j with $|V_j| \to \infty$ and suppose $k - \lambda_1(G) \geq \delta > 0$, δ £xed. Then G_j is an expander family.*

Remark 9.3.2. *If you have an algorithm that has a random element, then one can show there is another algorithm which does what this algorithm does without having a random component. (More or less, some slight of hand).*

Suppose we have a **bipartite graph**: there are two sets of vertices I (inputs) and O (outputs). Edges run only between I and O; there are no edges between two vertices in I or between two vertices in O. Let there be n inputs and n outputs, and join each input with k outputs.

Fix $\delta_0 > 0$. We want, for any $B \subset I$, $|\partial B| \geq \delta_0 |B|$.

We give a sketch of the proof. We will show that knowledge of a spectral gap ensures that the boundary of *any* subset of I will be big. We do bipartite for simplicity.

Let $B \subset I$. De£ne

$$f(v) = \begin{cases} 2n - |B| & \text{for } v \in B \\ -|B| & \text{otherwise.} \end{cases} \tag{9.25}$$

Then

$$\sum_{v \in V} f(v) = |B| \cdot (2n - |B|) + (2n - |B|) \cdot (-|B|) = 0. \tag{9.26}$$

Then

$$\lambda_1 = \max_{\langle \tilde{f}, f_0 \rangle = 0} = \frac{\langle A\tilde{f}, \tilde{f} \rangle}{\langle \tilde{f}, \tilde{f} \rangle}. \tag{9.27}$$

In particular,

$$\frac{\langle Af_1, f_1 \rangle}{\langle f_1, f_1 \rangle} \leq \lambda_1. \tag{9.28}$$

De£nition 9.3.3 (Laplacian). $\Delta = kI - A$.

Thus, $\Delta f_0 = kf_0 - Af_0 = 0 \cdot f_0$.
The eigenvalues of Δ are trivially related to the eigenvalues of A.

Remark 9.3.4 (Motivation for Laplacian). *On the line we have $\frac{d^2}{dx^2}$. A discrete version is (exercise)*

$$\frac{f(x+h) - 2f(x) + f(x-h)}{h^2}. \tag{9.29}$$

In the plane, we would have $\frac{d^2}{dx_1^2} + \frac{d^2}{dx_2^2}$. Integrating by parts we have

$$\begin{aligned} \int_\Omega (\Delta f) \cdot g dx_1 dx_2 &= -\int_\Omega \nabla f \cdot \nabla g dx_1 dx_2 \\ &= \int_\Omega f \cdot (\Delta g) dx_1 dx_2. \end{aligned} \tag{9.30}$$

We want to integrate by parts on a graph!

$$(\Delta F)(x) = kF(x) - \sum_{x \sim y} F(y). \tag{9.31}$$

Therefore

$$\begin{aligned} \langle \Delta F, F \rangle &= \sum_{x \in V} \left(kF(x) - \sum_{y \sim x} F(y) \right) F(x) \\ &= k \sum_{x \in V} F(x)^2 - \sum_{x \in V} \sum_{x \sim y} F(x)F(y). \end{aligned} \tag{9.32}$$

For each edge e, orient it by e^+ and e^-. The analogue of the Laplacian becomes

$$\begin{aligned} \sum_e \Big(F(e^+) - F(e^-) \Big)^2 &= \sum_e F^2(e^+) - 2F(e^+)F(e^-) + F^2(e^-) \\ &= 2\langle \Delta F, F \rangle, \end{aligned} \tag{9.33}$$

where the last line follows by thinking about what it means for vertices to be connected, and how often each vertex is hit.

Recall

$$k - \lambda_1 \;=\; \min_{\langle F, f_0\rangle = 0} \frac{\langle \Delta F, F\rangle}{\langle F, F\rangle} \;=\; \min_{\langle F, f_0\rangle = 0} \frac{1}{2}\frac{||dF||^2}{\langle F, F\rangle}. \tag{9.34}$$

Plug in the function f de£ned above, namely,

$$f(v) = \begin{cases} 2n - |B| & \text{for } v \in B \\ -|B| & \text{otherwise.} \end{cases} \tag{9.35}$$

If the edge runs from input to input or output to output, we get zero. The only way we get non-zero contribution is from an input to an output (or vice-versa). For our f,

$$\frac{1}{2}\sum_{e\in E}\Big|f(e^+) - f(e^-)\Big|^2 = \frac{1}{2}(2n)^2 \cdot \#\{\text{edges e running from } B \text{ to } B^c\}, \tag{9.36}$$

where B^c is the complement of B. We want $0 < \delta_0 = k - \lambda_1$. Divide the previous equation by $\langle F, F\rangle$, where $\langle F, F\rangle = (2n - |B|)^2 \cdot |B| + |B|^2 \cdot (2n - |B|)$ $= |B| \cdot (2n - |B|) \cdot 2n$.

Thus,

$$\delta_0 \le \frac{\frac{1}{2}(2n)\#\{edges\}}{|B| \cdot (2n - |B|) \cdot 2n}. \tag{9.37}$$

Therefore,

$$\begin{aligned} \#\{edges\} \;&\ge\; \frac{\delta_0 |B| \cdot (2n - |B|)}{n} \\ &\ge\; \delta_0 |B|, \end{aligned} \tag{9.38}$$

as $|B| \le n$.

Thus, the total number of edges is at least $\delta_0|B|$. But each vertex gets k edges. Thus,

$$|\partial B| \;\ge\; \frac{\delta_0 |B|}{k}. \tag{9.39}$$

9.4 $2\sqrt{k-1}$

Let G be a k-regular connected graph. A is the adjacency matrix. Biggest eigenvalue is $\lambda_0 = k$. Eigenvalues cannot be smaller than $-k$. How big is the gap between k and λ_1?

Theorem 9.4.1 (Alon-Boppana). *Fix k. Take any sequencce of graphs where the number of vertices $|G| \to \infty$. Then*

$$\overline{lim}_{|G|\to\infty}\lambda_1(A_G) \;\geq\; 2\sqrt{k-1}. \tag{9.40}$$

Remark 9.4.2. *In 1-dimension, with probability one the drunk returns home; same in 2-dimensions. He escapes with £nite probability in 3 and higher dimensions!*

Consider a 3 regular graph with many vertices. Go to v, look locally. If there are no short circuits, know what it looks like locally: it will look like a tree. (This is p-adic hyperbolic geometry).

Let T be the in£nite tree where each vertex is connected to three other vertices (and a vertex cannot be connected to itself). Suppose a drunk is walking on a tree. The only way he can get back is to exactly undo what he's done.

Consider the following operator: consider an in£nite dimensional Hilbert space $l_2(V)$, the set of all $f : V \to \mathbb{R}$ such that $\sum_v |f(v)|^2 < \infty$. This space is in£nite dimensional (for each v, take the function $f_v(w) = 1$ if $w = v$ and 0 otherwise).

Using $|ab| \leq \frac{a^2+b^2}{2}$,

$$\langle f, g\rangle = \sum_v f(v)g(v) \tag{9.41}$$

exists (and is our inner product). We de£ne

$$Af(v) = \sum_{w\sim v} f(w). \tag{9.42}$$

It is not obvious that there are any eigenvectors (and, in fact, there are no eigenvectors!). There is still a notion of spectrum. We will show the spectrum of this operator ($k = 3$) is $[-2\sqrt{2}, 2\sqrt{2}]$.

Note the constant function is horrendously not in this space (not even *close* to being square-integrable.

Chapter 10

Properties of Eigenvalues of Adjacency Matrices of Random Graphs

We discuss properties of eigenvalues of adjacency matrices arising from Random Graphs. Lecture by Peter Sarnak; notes by Steven J. Miller.

10.1 De£nitions

Let G be a connected, simple (no multiple bonds or edges) k-regular graph with adjacency matrix $A = (a_{v,w})$. Here

$$a_{v,w} = \begin{cases} 1 & \text{if } v \sim w \\ 0 & \text{otherwise} \end{cases} \tag{10.1}$$

Thus, $a_{v,w}$ is the number of paths of length one from v to w.

Let $A^2 = (a_{v,w}^{(2)})$, and here

$$a_{v,w}^{(2)} = \text{number of paths of length 2 from } v \text{ to } w. \tag{10.2}$$

Thus,

$$a_{v,w}^{(2)} = \sum_{v'} a_{v,v'} a_{v',w}. \tag{10.3}$$

Similarly let $A^n = (a_{v,w}^{(n)})$, and here

$$a_{v,w}^{(n)} = \text{number of paths of length } n \text{ from } v \text{ to } w. \tag{10.4}$$

Recall

$$\text{Trace}(B) \ = \ Tr(B) \ = \ \sum_v b_{v,v}. \tag{10.5}$$

Given an adjacency matrix A, let D be the diagonal matrix of eigenvalues.

$$D = \begin{pmatrix} \lambda_0 & & & \\ & \lambda_1 & & \\ & & \ddots & \\ & & & \lambda_{N-1} \end{pmatrix} \tag{10.6}$$

$A = Q^{-1}DQ$ for some orthogonal matrix Q. Thus, $A^n = Q^{-1}DQ$, and we £nd:

Lemma 10.1.1. $Tr(A^n) = Tr(D^n)$.

Lemma 10.1.2 (Trace Formula). *For any* $n \geq 0$*,*

$$\sum_{j=0}^{N-1} \lambda_j^n \ = \ \sum_{v \in V} a_{v,v}^{(n)}. \tag{10.7}$$

10.2 $\rho_k(2n)$ and $\lambda_{\max}$

To count walks of length n from v to v, it is clearly at least as many walks as there are on a k-regular in£nite tree.

A tree is a homogeneous object: any vertex looks exactly the same as any other. There is no special vertex on a tree, though we will often name a vertex **the root**.

De£nition 10.2.1. $\rho_k(n)$ *is the number of paths of length* n *from* v *to* v*, where* v *is* any *vertex of* the k*-regular tree.*

Remark 10.2.2. $\rho_k(n) = 0$ *for* n *odd.*

Remark 10.2.3. *The number of paths of length n from v to v on our graph G is at least the number of paths of length n from any vertex to itself on the in£nite tree. Thus,*

$$\forall v \in V,\ a_{v,v}^{(n)} \ \geq\ \rho_k(n). \tag{10.8}$$

Remember that we've labeled the N eigenvalues by $\lambda_0 = k, \dots, \lambda_{N-1}$.
In the trace formula, we £nd

$$\sum_{j=0}^{N-1} \lambda_j^{2n} \ \geq\ N\rho_k(2n). \tag{10.9}$$

Therefore

$$\frac{1}{N}\sum_{j=0}^{N-1} \lambda_j^{2n} \ \geq\ \rho_k(2n). \tag{10.10}$$

and as $\lambda_0 = k$

$$\frac{k^{2n}}{N} + \frac{1}{N}\sum_{j=1}^{N-1} \lambda_j^{2n} \ \geq\ \rho_k(2n). \tag{10.11}$$

Fix n and let $N \to \infty$.
Let $\lambda_{\max} = \max\Big(|\lambda_1|, |\lambda_{N-1}|\Big)$.
Thus, substituting into Equation 10.11 we £nd

$$\frac{k^{2n}}{N} + \lambda_{\max}^{2n} \geq \rho_k(2n) \tag{10.12}$$

in the limit as $N \to \infty$ (as we have λ_{max} a total of $N-1$ times, and we divide by N; in the limit, $\frac{N-1}{N} \to 1$.
As $N \to \infty$, we £nd

$$\begin{aligned} \lambda_{\max}^{2n} &\geq \rho_k(2n) \\ \text{or } \lambda_{\max} &\geq \Big(\rho_k(2n)\Big)^{\frac{1}{2n}}. \end{aligned} \tag{10.13}$$

Exercise 10.2.4. *Show*

1. $\rho_k(2n) \geq \frac{1}{m}\binom{2m-2}{m-1}k(k-1)^{m-1}$.

2. $\left(\rho_k(2n)\right)^{\frac{1}{2n}} \to 2\sqrt{k-1}$.

Using the above exercise, we now £nd that

$$\lambda_{\max} \geq 2\sqrt{k-1}. \tag{10.14}$$

10.3 Measure from the Eigenvalues

The trace formula told us that

$$\sum_{j=0}^{N-1} \lambda_j^n = \sum_{v \in V} a_{v,v}^{(n)}. \tag{10.15}$$

De£nition 10.3.1 (girth). *The girth of a graph is the length of the shortest closed cycle that returns to the starting vertex without any backtracking.*

Assume that the girth of G_N tends to ∞ as $N \to \infty$.

Fix n, let N be very large. Then by assumption the girth is greater than say $2n+1$. Thus, $a_{v,v}^{(n)}$ cannot have any contribution from cycles without backtracking. Thus, locally, to calculate $a_{v,v}^{(n)}$, we look like a tree, and we £nd $a_{v,v}^{(n)} = \rho_k(n)$ for every $v \in V$. It is *essential* that we have £xed n.

Therefore, we now have (for £xed n under our assumption) that

$$\begin{aligned} \sum_{j=0}^{N-1} \lambda_j^n &= N\rho_k(n) \\ \frac{1}{N}\sum_{j=0}^{N-1} \lambda_j^n &= \rho_k(n). \end{aligned} \tag{10.16}$$

The left hand side looks like a Riemann sum.

Suppose the density of the eigenvalues of the 3-regular graph is $d\mu = f(x)dx$. We have just shown, for polynomials $p_n(x) = x^n$, that

$$\frac{1}{N}\sum_{j=0}^{N-1} p_n(\lambda_j) \to \int_{-3}^{3} p_n(x)d\mu(x), \tag{10.17}$$

where the above converges to $\rho_k(n)$ (here $k = 3$).

Thus, we are looking for a density function $f(x)$ such that

$$\int_{-3}^{3} x^n f(x)dx = \rho_3(n), \ n \geq 0. \tag{10.18}$$

Is there such a function? Is it unique? What does it look like? This is the **Inverse Moment Problem**.

If there were two such functions, they would have to be equal by the **Weierstrass Approximation Theorem**, as their difference integrates to zero against any polynomial.

Exercise 10.3.2. *Compute the generating function*

$$F(z) = \sum_{n=0}^{\infty} \rho_3(n) z^n, \tag{10.19}$$

which is something like

$$\frac{1}{\sqrt{4(k-1)^2 - z^2}} \tag{10.20}$$

if $|z|$ is small (or maybe complex and outside $[-k, k]$).

$$\begin{aligned}
\sum_{n=0}^{\infty} \rho_3(n) z^n &= \sum_{n=0}^{\infty} \left(\int_{-3}^{3} x^n f(x) dx \right) z^n \\
&= \int_{-3}^{3} f(x) \sum_{n=0}^{\infty} (zx)^n dx \\
&= \int_{-3}^{3} \frac{f(x)dx}{1 - zx} = F(z).
\end{aligned} \tag{10.21}$$

If we let $z = \frac{1}{w}$ we £nd

$$\begin{aligned}
F(w) &= \sum_{n=0}^{\infty} \rho_3(n) \frac{1}{w^n} \\
&= w \int_{-3}^{3} \frac{f(x)dx}{w - x}.
\end{aligned} \tag{10.22}$$

Let $w \in C$ be such that $w \notin [-3,3]$. Letting $w \to a \in [-3,3]$, one gets a different value (a jump) if w approaches from above or below, and the jump is basically $f(a)$.

Look at $a + ib$ and $a - ib$, $b \to 0$.

10.4 Summary

The above is all based on the assumption that the girth was big. For the random graph, there are very few short closed cycles. Thus, when we use the trace formula, we now have

$$\frac{1}{N}\sum_{j=0}^{N-1} \lambda_j^n = \rho_k(n) + O\left(\frac{1}{N}\right). \tag{10.23}$$

Chapter 11

Spacings of Eigenvalues of Real Symmetric Matrices; Semi-Circle Law

Joint Density Function for eigenvalues of real symmetric matrices; spacing of eigenvalues for 2×2 real symmetric matrices; Semi-Circle Rule. Lecture by Steven J. Miller; notes by Steven J. Miller and Alex Barnett.

11.1 Joint density function of eigenvalues of real symmetric matrices ('GOE')

11.1.1 Dirac Notation

The derivation handed out in lecture used physics notation which should be explained. The matrix is called the 'Hamiltonian' (meaning that it happened to arise in a quantum physics problem). Vectors are often called *states* (referring to quantum states), however they can be thought of as your usual vectors. (Quantum mechanics is just linear algebra, amazingly). A general vector in 2D is written

$$|u\rangle \quad \text{equivalent to} \quad \mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \tag{11.1}$$

the latter being its coordinate representation in some basis. The unit vectors are

$$|1\rangle\,,\ |2\rangle \quad \text{equivalent to} \quad \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{11.2}$$

The $|u\rangle$ is a column vector, and $\langle u| \equiv |u\rangle^T$ is a row vector. Inner product can be written as $\langle v\,|u\rangle = \mathbf{v}^T \cdot \mathbf{u}$. General bilinear product can be written $\langle v\,|M|\,u\rangle = \mathbf{v}^T \cdot M \cdot \mathbf{u}$.

11.1.2 2×2 Gaussian Orthogonal Ensemble (GOE)

We consider 2×2 real symmetric matrices,

$$A \equiv \begin{pmatrix} x & y \\ y & z \end{pmatrix}. \tag{11.3}$$

Understanding this case is *vital* to building intuition about Random Matrix Theory for $N \times N$ matrices.

A can always be diagonalized by an orthogonal matrix Q as follows,

$$Q^T \begin{pmatrix} x & y \\ y & z \end{pmatrix} Q = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \equiv D. \tag{11.4}$$

In 2×2 case, the characteristic equation $\det(A - \lambda I) = 0$ is quadratic:

$$\lambda^2 - \text{Tr}(A)\lambda + \det(A) = 0, \tag{11.5}$$

where

$$\text{Tr}(A) = x + z,\ \det(A) = xz - y^2. \tag{11.6}$$

Solutions are

$$\lambda_{1,2} = \frac{x+z}{2} \pm \sqrt{\left(\frac{x-z}{2}\right)^2 + y^2}, \tag{11.7}$$

where 1 is the $+$ case, 2 the $-$.

If the two eigenvalues are equal, we say the matrix is degenerate. Initially we are in a three-dimensional space (as x, y and z are arbitrary). Degeneracy requires that x, y and z satisfy

$$\left(\frac{x-z}{2}\right)^2 + y^2 = 0, \tag{11.8}$$

or, equivalently,

$$x - z = 0,\ y = 0. \tag{11.9}$$

Thus, we lose two degrees of freedom, because there are two equations which must be satis£ed. The set of solutions is $\{(x, y, z) = (x, 0, x)\}$.

Exercise 11.1.1. *Show that $\lambda_1 - \lambda_2$ is twice the distance from the origin in this 2D subspace.*

Corresponding eigenvectors are,

$$\mathbf{v}_1 = \begin{pmatrix} c \\ s \end{pmatrix}, \qquad \mathbf{v}_2 = \begin{pmatrix} -s \\ c \end{pmatrix}. \tag{11.10}$$

We use abbreviations $c \equiv \cos\theta$ and $s \equiv \sin\theta$.

Why can we write the eigenvectors as above? We can always normalize the eigenvector attached to a given eigenvalue to have length 1. We have previously shown that, if the eigenvalues are distinct, then the eigenvectors of a real symmetric matrix are perpendicular. This forces the above form for the two eigenvectors, at least when $\lambda_1 \neq \lambda_2$.

One rotation angle θ de£nes the orthogonal matrix,

$$Q = Q(\theta) = \begin{pmatrix} \mathbf{v}_1 \; \mathbf{v}_2 \end{pmatrix} = \begin{pmatrix} c & -s \\ s & c \end{pmatrix}. \tag{11.11}$$

The structure of the eigenvectors is actually quite rich.

Exercise 11.1.2. *Find θ in terms of x, y, z. Hint: use trigonometric identities to simplify the resulting form. Hint: solve $(A - \lambda_1 \mathbf{v}_1) = \mathbf{0}$.*

Exercise 11.1.3. *Show that a general A can be written*

$$A \;=\; \alpha \begin{pmatrix} \cos\beta & \sin\beta \\ \sin\beta & -\cos\beta \end{pmatrix} \;+\; \gamma \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{11.12}$$

Exercise 11.1.4. *Find $\lambda_{1,2}$ in terms of α, β, γ. Show that the eigenvector angle is given by $\theta = \beta/2$. This result is quite deep; for instance notice that taking a complete 2π cycle in β reverses the signs of the eigenvectors! This isn't that relevant for the rest of this lecture.*

We adopt two assumptions about the joint distribution over A, called $p(A) \equiv p(x, y, z)$:

1. Invariance of p under orthogonal transformations (aka 'basis-invariance'), $p(M^T A M) = p(A)$ for all orthogonal M.

2. Independence of distributions of individual matrix elements, $p(x, y, z) = p_x(x) p_y(y) p_z(z)$.

Section 2C-1 of the handout reminds us that these two assumptions taken together demand a unique form of distribution,

$$p(x, y, z) \;\propto\; e^{-C\mathrm{Tr}(A^2)}, \tag{11.13}$$

depending on only one parameter C; we choose $C = 1$. Note $\propto$ means proportional to; the constant of proportionality is what is needed to make $P9x, y, z)$ a probability distribution (ie, the integral of $\int\int\int p(x, y, z)dxdydz = 1$.

This corresponds to Gaussian distributions of matrix elements,

$$\begin{aligned} p_y(y) &= \sqrt{\frac{2}{\pi}} e^{-2y^2} && \text{off-diag} \\ p_x(x) = p_z(x) &= \frac{1}{\sqrt{\pi}} e^{-x^2} && \text{diag.} \end{aligned} \tag{11.14}$$

Note that the diag elements have variance $\frac{1}{2}$, the off-diag variance $\frac{1}{4}$. We show how to compute the normalization prefactors later on. This form (for general C) is the so-called GOE. The $n \times n$ case is derived in Miller's handout of 9/25/02.

11.1.3 Transformation to diagonal representation

The operation of diagonalizing A can be viewed as the transformation from one 3D space to another 3D space,

$$\mathbf{r} \equiv \underbrace{(x, y, z)}_{A} \;\longleftrightarrow\; \mathbf{r}' \equiv \underbrace{(\lambda_1, \lambda_2, \theta)}_{D,Q}. \tag{11.15}$$

This is 1-to-1 apart from the set of measure zero (ie, a lower dimensional subspace) corresponding to degenerate eigenvalues. Looking at Eq. 11.4 we can see the transformation is linear in the eigenvalues, nonlinear in θ. We are interested in the *marginal* distribution of the eigenvalues,

$$p'(\lambda_1, \lambda_2) \;\equiv\; \int d\theta\, p'(\lambda_1, \lambda_2, \theta), \tag{11.16}$$

in other words we don't care what θ is. We use primes to signify distributions over £nal (Q, D) variables.

We want to know how to transform probability density from $\mathbf{r}$ space to $\mathbf{r}'$ space. In general this must follow the law,

$$p(\mathbf{r})d\mathbf{r} \;=\; p'(\mathbf{r}')d\mathbf{r}', \tag{11.17}$$

giving

$$p'(\mathbf{r}') \;=\; \det(J)p(\mathbf{r}). \tag{11.18}$$

The ratio of the volume elements is $|\det J|$ where J is the 3×3 Jacobean matrix of the transformation. J has elements $J_{ij} = \partial r_j / \partial r'_i$.

Inverting Eq. 11.4 we can write $A(\mathbf{r}')$ as

$$\begin{aligned}
\begin{pmatrix} x & y \\ y & z \end{pmatrix} = QDQ^T \;&=\; \begin{pmatrix} c & -s \\ s & c \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} c & s \\ -s & c \end{pmatrix} \\
&=\; \begin{pmatrix} \lambda_1 c^2 + \lambda_2 s^2 & (\lambda_1 - \lambda_2) sc \\ (\lambda_1 - \lambda_2) sc & \lambda_1 s^2 + \lambda_2 c^2 \end{pmatrix}
\end{aligned} \tag{11.19}$$

We evaluate J for this case,

$$J \;\equiv\; \begin{pmatrix} \frac{\partial x}{\partial \lambda_1} & \frac{\partial y}{\partial \lambda_1} & \frac{\partial z}{\partial \lambda_1} \\ \frac{\partial x}{\partial \lambda_2} & \ddots & \\ \frac{\partial x}{\partial \theta} & & \end{pmatrix}. \tag{11.20}$$

We see λ's only appear in the bottom three entries, and furthermore they only appear as factors $(\lambda_1 - \lambda_2)$ in each entry.

Exercise 11.1.5. *Evaluate the bottom row of J to prove the above.*

Therefore this factor of a row of J can be brought out in evaluating the determinant:

$$\det(J) = \left| \text{messy } \theta\text{-dep } 3 \times 3 \text{ matrix} \right| \cdot (\lambda_1 - \lambda_2) \;=\; g(\theta)(\lambda_1 - \lambda_2). \tag{11.21}$$

Warning! The Jacobian is the absolute value of the determinant. Thus, we need $|\lambda_1 - \lambda_2|$ above, or we need to adopt the convention that we label the eigenvalues so that $\lambda_1 \geq \lambda_2$.

The only dependence on the λ's is given by the second factor. Plugging into Eq. 11.18 and marginalizing over θ gives,

$$\begin{aligned}
p'(\lambda_1, \lambda_2) \;&=\; \int d\theta \, g(\theta) \, (\lambda_1 - \lambda_2) e^{-(\lambda_1^2 + \lambda_2^2)} \\
&\propto\; (\lambda_1 - \lambda_2) e^{-(\lambda_1^2 + \lambda_2^2)}.
\end{aligned} \tag{11.22}$$

Note that we do not need the absolute value sign around $(\lambda_1 - \lambda_2)$ because we chose $\lambda_1 > \lambda_2$. This is the joint density of the eigenvalues in 2×2 GOE.

11.1.4 Generalization to $n \times n$ case

The above generalizes quite easily, with the dimension of the two spaces being $N = \frac{1}{2}n(n+1)$. The $\frac{1}{2}n(n+1)$ degrees of freedom in A equal n degrees of freedom in D (namely the eigenvalues $\{\lambda_i\}$) plus $\frac{1}{2}n(n-1)$ degrees of freedom in Q (namely the generalized angles Ω). We sort the eigenvalues such that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$.

Theorem 11.1.6. *If $\lambda_i = \lambda_j$ for some $1 \leq i < j \leq n$, then the Jacobean of the transformation $A \leftrightarrow (D, Q)$ vanishes, that is $\det(J) = 0$.*

The proof relies on realising that the two eigenvectors $\mathbf{v}_i$ and $\mathbf{v}_j$ span a 2D subspace, invariant under A. (Recall here we refer to a subspace of the n-dim vector space upon which A operates by multiplication). The invariance means that the choices of directions of the eigenvectors is arbitrary in this 2D plane. Therefore there is one angle degree of freedom in Ω which in not constrained by A, that is, it is independent of the elements of A. Now think of the inverse transformation from $(D, Q) \to A$. An in£nitesimal volume element is transformed as

$$d\mathbf{r} \;=\; \det(J) d\mathbf{r}'. \tag{11.23}$$

Changes of eigenvector angle within the 2D subspace have no effect on A, so the volume element $d\mathbf{r}$ is collapsed to zero. (Another way of putting this is that J acquires a null-space of dimension 1). Therefore $\det(J) = 0$. $\square$.

This vanishing of the Jacobean at degeneracies renders the non-uniqueness of the forward map $A \to (D, Q)$ at these points harmless in the following.

The upper n rows of J are messy functions of angles Ω, and the bottom $\frac{1}{2}n(n-1)$ rows contain entries each which is *linear* in the eigenvalues. Therefore $\det(J)$ is a polynomial of degree $\frac{1}{2}n(n-1)$ in the eigenvalues λ_i. Further, $\det(J) = 0$ if any two eigenvalues are equal.

Consider the polynomial $\prod_{1 \leq i < j \leq n}(\lambda_i - \lambda_j)$. First, note that this polynomial vanishes whenever two eigenvalues are the same. We claim it is a polynomial of degree $\frac{1}{2}n(n-1)$ in the eigenvalues. For each j, there are $j-1$ choices for i. Thus, the degree is

$$\sum_{j=2}^{n} j - 1 = \sum_{k=1}^{n-1} k = \frac{(n-1)(n-1+1)}{2} = \frac{n(n-1)}{2}. \tag{11.24}$$

Thus, $\det(J)$ and $\prod_{1 \leq i < j \leq n}(\lambda_i - \lambda_j)$ both vanish whenever two eigenvalues are equal, and they have the same degree. Therefore, they must be scalar multiples of each other.

So,

$$\det(J) \propto \prod_{1\le i<j\le n} (\lambda_i - \lambda_j). \tag{11.25}$$

Combining with the GOE form of $p(A)$ gives, after marginalizing over Ω as before,

$$p(\{\lambda_i\}) \;=\; \prod_{1\le i<j\le n} (\lambda_i - \lambda_j) \cdot e^{-\sum_{i=1}^{n} \lambda_i^2}. \tag{11.26}$$

The vanishing of this probability density as any two eigenvalues come close is called *level repulsion*.

11.2 Eigenvalue spacing distribution in 2×2 real symmetric matrices

11.2.1 Reminder: Integral of the Gaussian

We want

$$I = \int_{-\infty}^{\infty} e^{-x^2} dx. \tag{11.27}$$

Square it and rearrange the summation over area by using polar coordinates:

$$\begin{aligned} I^2 &= \int_{-\infty}^{\infty} e^{-x^2} dx \cdot \int_{-\infty}^{\infty} e^{-y^2} dy \;=\; \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-x^2-y^2} dx\, dy \\ &= \int_0^{2\pi} d\theta \int_0^{\infty} r\, dr\, e^{-r^2} \;=\; 2\pi \cdot \left[-\frac{1}{2} e^{-r^2}\right]_0^{\infty} \\ &= \pi. \end{aligned} \tag{11.28}$$

Introduction of the radius factor r produced re^{-r^2}, a known differential. So,

$$I \;=\; \sqrt{\pi}. \tag{11.29}$$

Changing the variable in the above, and rearranging, gives

$$\frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2\sigma^2}} dx \;=\; 1. \tag{11.30}$$

This is therefore the correct normalization for a 1D Gaussian probability density, of variance σ^2.

11.2.2 Spacing distribution

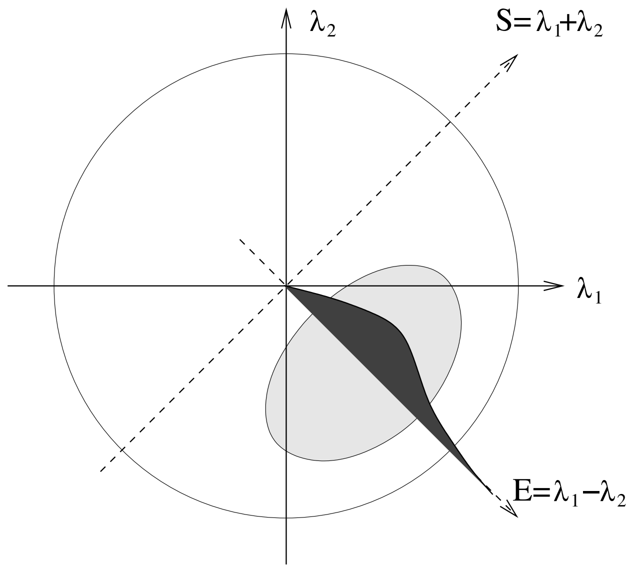

Figure 11.1: Change of coordinates to get spacing distribution in E. Light shading suggests form of density across the 2D plane. Dark shading shows a graph of its projection onto the E axis.

Here for convenience we present a slightly simpler derivation than in lecture. Given the 2D density $p'(\lambda_1, \lambda_2)$ we want the 1D density of the difference $E \equiv \lambda_1 - \lambda_2$. This will require marginalizing again, since there is a reduction in dimensionality. We de£ne $S \equiv \lambda_1 + \lambda_2$. The linear transformation $(\lambda_1, \lambda_2) \to (E, S)$ has £xed Jacobean (it is a rotation by -45^o and a compression by $\sqrt{2}$ in each axis). See Fig. 11.1.

Therefore, substituting in $\lambda_1 = (S+E)/2$ and $\lambda_1 = (S-E)/2$ into Eq. 11.22 gives

$$\begin{aligned} p'(E,S) &\propto p'(\lambda_1(E,S), \lambda_2(E,S)) \;=\; Ee^{-\frac{1}{4}[(S+E)^2+(S-E)^2]} \\ &= Ee^{-E^2/2} \cdot e^{-S^2/2}, \end{aligned} \tag{11.31}$$

which is separable. Therefore integrating over S gives an E-independent number, and

$$p'(E) \;\propto\; Ee^{-E^2/2}. \tag{11.32}$$

This is the so-called 'Wigner Surmise' for the eigenvalue spacing density. Remarkably, in the $n \times n$ case, even for large n, this density is very close to the true spacing distribution of adjacent eigenvalues. The limiting powerlaw $\lim_{E\to 0} p'(E) \propto E^\beta$ with $\beta = 1$ is intimately related to the matrix symmetry class GOE that we close. It is also possible to achieve $\beta = 2$ and $\beta = 4$ by choosing different symmetry classes.

Finally, let's say you couldn't be bothered to construct your second variable S. Instead you could derive the above using the Dirac delta-function (see below) to marginalize:

$$p'(E) \;=\; \int_{-\infty}^{\infty} d\lambda_1 \int_{-\infty}^{\lambda_1} d\lambda_2 \, p'(\lambda_1, \lambda_2)\, \delta(E - (\lambda_1 - \lambda_2)). \tag{11.33}$$

Apart from the unusual limits (due to ordering of eigenvalues), this is the standard procedure to extract a marginal density.

Exercise 11.2.1. *Simplify the above to arrive at* $p'(E)$.

11.3 Delta Function(al)

Let $f(x)$ be a nice function; for example, let $f(x)$ be an in£nitely differentiable function whose Taylor Series converges to $f(x)$:

$$f(x) = f(0) + \frac{f'(0)}{1!}x + \frac{f''(0)}{2!}x^2 + \cdots \tag{11.34}$$

Let

$$\delta_n(x) = \begin{cases} n & \text{if } x \in \left[-\frac{1}{2n}, \frac{1}{2n}\right] \\ 0 & \text{otherwise} \end{cases} \tag{11.35}$$

Exercise 11.3.1. *Show that*

$$\int_{-\infty}^{\infty} f(x)\delta_n(x)dx = f(0) + O\Big(\frac{1}{n}\Big). \tag{11.36}$$

Let δ be the limit as $n \to \infty$ of δ_n. We £nd / de£ne

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} f(x)\delta_n(x)dx \;=\; \int_{-\infty}^{\infty} f(x)\delta(x)dx \;=\; f(0). \tag{11.37}$$

Exercise 11.3.2. *Show that*

$$\int_{-\infty}^{\infty} f(x)\delta(x-a)dx \;=\; f(a). \tag{11.38}$$

A good analogy for the δ functional is a point mass. A point mass has no extension (no length, width or height) but £nite mass. Therefore, a point mass has in£nite density.

A probability density must integrate to one. This corresponds to $\int 1 \cdot \delta(x)dx = 1$. We often refer to $\delta(x)$ as a point mass at the origin, and $\delta(x-a)$ as a point mass at a.

11.4 De£nition of the Semi-Circle Density

Consider

$$P(x) = \begin{cases} \frac{2}{\pi}\sqrt{1-x^2} & \text{if } |x| \leq 1 \\ 0 & \text{otherwise} \end{cases} \tag{11.39}$$

Exercise 11.4.1. *Show that $P(x)$ is a probability density. IE, show that it is non-negative and integrates to 1. Graph $P(x)$.*

We call $P(x)$ the semi-circle density.

11.5 Semi-Circle Rule: Preliminaries

Let λ_j be the eigenvalues of a real, symmetric $N \times N$ matrix A. We normalize the eigenvalues of A by dividing by $2\sqrt{N}$.

De£ne

$$\mu_{A,N}(x) \;=\; \frac{1}{N}\sum_{j=1}^{N} \delta\Big(x - \frac{\lambda_j(A)}{2\sqrt{N}}\Big). \tag{11.40}$$

$\delta\Big(x - \frac{\lambda_j(A)}{2\sqrt{N}}\Big)$ is a point mass at $\frac{\lambda_j(A)}{2\sqrt{N}}$. By summing these point masses and dividing by N, we have a probability distribution. For example,

$$\int_{-\infty}^{\infty} f(x)\mu_{A,N}(x)dx \;=\; \sum_{j=1}^{N} f\Big(\frac{\lambda_j(A)}{2\sqrt{N}}\Big). \tag{11.41}$$

We will show that, as $N \to \infty$, the above converges to the integral of f against the semi-circle density:

$$\int_{-\infty}^{\infty} f(x)P(x)dx. \tag{11.42}$$

What does this mean?

$$\sum_{j=1}^{N} f\Big(\frac{\lambda_j(A)}{2\sqrt{N}}\Big) \tag{11.43}$$

looks like a Riemann Sum. The statement that, for nice $f(x)$,

$$\sum_{j=1}^{N} f\Big(\frac{\lambda_j(A)}{2\sqrt{N}}\Big) \;\to\; \int_{-\infty}^{\infty} f(x)P(x)dx \tag{11.44}$$

means that as $N \to \infty$, the number of eigenvalues of a random A in $[a, b]$ equals

$$\int_{a}^{b} P(x)dx. \tag{11.45}$$

Theorem 11.5.1. *Choose the entries a_{ij} of a real, symmetric matrix independently from a £xed probability distribution p with mean zero, variance one, and £nite higher moments. For each A, form the probability measure $\mu_{A,N}$. As $N \to \infty$, with probability one the measures $\mu_{A,n}(x)dx$ converge to the semi-circle probability $P(x)dx$.*

This is not the most general version; however, it is rich enough for our purposes.

11.6 Sketch of Proof of the Semi-Circle Law

11.6.1 Calculation of Moments via Trace Formula

We will show that the expected value of the moments of the $\mu_{A,N}(x)$ equal the moments of the semi-circle.

De£nition 11.6.1. *$M_{A,N}(k)$ is the k^{th} moment of the probability measure attached to $\mu_{A,N}(x)dx$:*

$$M_{A,N}(k) \;=\; \int x^k \mu_{A,N}(x)dx \;=\; \frac{1}{N}\sum_{j=1}^{N}\Big(\frac{\lambda_j(A)}{2\sqrt{N}}\Big)^k. \tag{11.46}$$

Note that $\sum \lambda_j(A)^k = \text{Trace}(A^k)$. Thus, we have

$$M_{A,N}(k) \;=\; \frac{1}{2^k N^{1+\frac{k}{2}}}\text{Trace}(A^k). \tag{11.47}$$

We now calculate the expected values of the £rst few moments ($k = 0, 1$, 2 and 3).

Lemma 11.6.2. *The expected value of $M_{A,N}(0) = 1$.*

Proof:

$$E\Big[M_{A,N}(0)\Big] \;=\; \frac{1}{N}E\Big[\text{Trace}(I)\Big] \;=\; 1. \tag{11.48}$$

Note that summing the eigenvalues to the zeroth power is the same as taking the trace of the identity matrix. □

Lemma 11.6.3. *The expected value of $M_{A,N}(1) = 0$.*

Proof:

$$\begin{aligned} E\Big[M_{A,N}(1)\Big] \;&=\; \frac{1}{2N^{3/2}}E\Big[\text{Trace}(A)\Big] \\ &=\; \frac{1}{2N^{3/2}}E\Big[\sum_i a_{ii}\Big] \\ &=\; \frac{1}{2N^{3/2}}\sum_i E[a_{ii}] \;=\; 0, \end{aligned} \tag{11.49}$$

because we have assumed that each a_{ij} is drawn from a probability distribution with mean zero. □

Lemma 11.6.4. *The expected value of $M_{A,N}(2) = \frac{1}{4}$.*

Proof: Note that

$$\text{Trace}(A^2) = \sum_i \sum_j a_{ij} a_{ji}. \tag{11.50}$$

As our matrix is symmetric, $a_{ij} = a_{ji}$. Thus, the trace is $\sum_i \sum_j a_{ij}^2$. Now

$$\begin{aligned} E\Big[M_{A,N}(2)\Big] &= \frac{1}{4N^2} E\Big[\text{Trace}(A^2)\Big] \\ &= \frac{1}{4N^2} E\Big[\sum_i \sum_j a_{ij}^2\Big] \\ &= \frac{1}{4N^2} \sum_i \sum_j E[a_{ij}^2] \;=\; \frac{1}{4}, \end{aligned} \tag{11.51}$$

where the last line follows from each a_{ij} has variance 1. As their means are zero, the variance $E[a_{ij}^2] - E[a_{ij}]^2 = 1$ implies $E[a_{ij}] = 1$. There are N^2 pairs (i, j). Thus, we have $\frac{1}{4N^2} \cdot (N \cdot 1) = \frac{1}{4}$. $\Box$

Lemma 11.6.5. *The expected value of* $M_{A,N}(3) = 0$ *as* $N \to \infty$.

We need

$$\text{Trace}(A^3) \;=\; \sum_i \sum_j \sum_k a_{ij} a_{jk} a_{ki}. \tag{11.52}$$

We £nd

$$\begin{aligned} E\Big[M_{A,N}(3)\Big] &= \frac{1}{8N^{2.5}} E\Big[\text{Trace}(A^3)\Big] \\ &= \frac{1}{8N^{2.5}} E\Big[\sum_i \sum_j \sum_k a_{ij} a_{jk} a_{ki}\Big] \\ &= \frac{1}{8N^{2.5}} \sum_i \sum_j \sum_k E[a_{ij} a_{jk} a_{ki}]. \end{aligned} \tag{11.53}$$

There are three cases. If the subscripts i, j and k are all distinct, then a_{ij}, a_{jk}, and a_{ki} are three independent variables. Hence

$$E[a_{ij} a_{jk} a_{ki}] \;=\; E[a_{ij}] \cdot E[a_{jk}] \cdot E[a_{ki}] \;=\; 0. \tag{11.54}$$

If two of the subscripts are the same (say $i = j$) and the third is distinct, we have

$$E[a_{ii}a_{ik}a_{ki}] \;=\; E[a_{ii}] \cdot E[a_{ik}^2] \;=\; 0 \cdot 1 \;=\; 0. \qquad (11.55)$$

If all three subscripts are the same, we have

$$E[a_{ii}^3] \qquad (11.56)$$

This is the third moment of a_{ii}. It is the same for all variables a_{ii}, and is £nite by assumption. There are N triples where $i = j = k$.

Thus,

$$E\Big[M_{A,N}(3)\Big] \;=\; \frac{1}{8N^{2.5}} \cdot NE[a_{11}^3] \;=\; \frac{E[a_{11}^3]}{8} \cdot \frac{1}{N^{1.5}}. \qquad (11.57)$$

Thus, as $N \to \infty$, the expected value of the third moment is zero. $\Box$

To calculate the higher moments requires signi£cantly more delicate combinatorial arguments.

11.6.2 Calculation of Moments from the Semi-Circle

We now calculate the moments of the semi-circle. For $k \leq 3$, the k^{th} moment of the semi-circle $C(k)$ equals the expected k^{th} moment of $\mu_{A,N}(x)$ as $N \to \infty$.

$$C(k) \;=\; \int_{-\infty}^{\infty} x^k P(x)dx \;=\; \frac{2}{\pi}\int_{-1}^{1} x^k\sqrt{1-x^2}dx. \qquad (11.58)$$

We note that, by symmetry, $C(k) = 0$ for k odd, and $C(0) = 1$ as $P(x)$ is a probability density.

For $k = 2m$ even, we change variables $x = \sin\theta$.

$$C(2m) \;=\; \frac{2}{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin^{2m}\theta \cdot \cos^2\theta d\theta. \qquad (11.59)$$

Using $\sin^2\theta = 1 - \cos^2\theta$ gives

$$C(2m) \;=\; \frac{2}{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin^{2m}\theta d\theta \;-\; \frac{2}{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin^{2m+2}\theta d\theta. \qquad (11.60)$$

The above integrals can be evaluated exactly. We constantly use

$$\begin{aligned} \cos^2(\phi) &= \frac{1}{2} + \frac{1}{2}\cos(2\phi) \\ \sin^2(\phi) &= \frac{1}{2} - \frac{1}{2}\cos(2\phi). \end{aligned} \tag{11.61}$$

Repeated applications of the above allow us to write $\sin^{2m}(\theta)$ as a linear combination of 1, $\cos(2\theta), \dots, \cos(2m\theta)$.

Let

$$n!! = \begin{cases} n \cdot (n-2) \cdots 2 & \text{if } n \text{ is even} \\ n \cdot (n-2) \cdots 1 & \text{if } n \text{ is odd} \end{cases} \tag{11.62}$$

We £nd (either prove directly or by induction) that

$$\frac{2}{\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin^{2m} \theta d\theta \;=\; 2\frac{(2m-1)!!}{(2m)!!}. \tag{11.63}$$

Exercise 11.6.6. *Show the above gives*

$$C(2m) \;=\; 2\frac{(k-1)!!}{(k+2)!!}. \tag{11.64}$$

Also, show $C(2)$ agrees with our earlier calculation.

Chapter 12

More Graphs, Maps mod p, Fourier Series and n alpha

More on Graphs; Arithmetic maps mod p. Lecture by Peter Sarnak; notes by Steven J. Miller. Appendix: Introduction to Fourier Series and $\{n\alpha\}$, by Steven J. Miller.

12.1 Kesten's Measure

For a k-regular graph, de£ne

$$d\mu_k(t) = \begin{cases} \frac{c_k\sqrt{(k-1)-\frac{t^2}{4}}}{1-(\frac{t}{k})^2}dt & \text{if } |t| \leq 2\sqrt{k-1} \\ 0 & \text{otherwise} \end{cases} \tag{12.1}$$

Fix a and b, and consider the N eigenvalues λ_j. Count

$$\frac{\#\{j : \lambda_j \in [a,b]\}}{N}. \tag{12.2}$$

Then

Claim 12.1.1.

$$\lim_{N\to\infty} \frac{\#\{j : \lambda_j \in [a,b]\}}{N} \ = \ \mu_k\Big([a,b]\Big). \tag{12.3}$$

We have

$$\Big(\rho_k(2n)\Big)^{\frac{1}{2n}} \ \to \ 2\sqrt{k-1}. \tag{12.4}$$

12.2 Generating Functions on k-Regular Trees

12.2.1 $R(z)$

Fix k. The **generating function** $R(z)$ is

$$R(z) \;=\; \sum_{n=0}^{\infty} r_n z^n, \tag{12.5}$$

where

$$r_n \;=\; \frac{\rho(n)}{k^n} \;=\; \text{Probability that we return to } v \text{ in } n \text{ steps.} \tag{12.6}$$

For $|z| < 1$, the series expansion for $R(z)$ converges.

Let q_n be the probability of starting at v and ending at v for the £rst time (after n steps).

12.2.2 $Q(z)$

De£ne

$$Q(z) \;=\; \sum_{n=0}^{\infty} q_n z^n. \tag{12.7}$$

Exercise 12.2.1. *Prove*

$$R(z) \;=\; \frac{1}{1 - Q(z)}. \tag{12.8}$$

12.2.3 $T(z)$

De£ne

$$T(z) \;=\; \sum_{n=0}^{\infty} t_n z^n, \tag{12.9}$$

where for w adjacent to v, t_n is the probability of going from w to v in n-steps for the £rst time.

Further, let $t_{w,v}(n)$ be the probability of going from w to v in n-steps £rst time and $d(w,v) = m \geq 1$. Remember that $d(w,v)$ is the distance from w to v.

Exercise 12.2.2. *Prove*

1. $Q(z) = zT(z)$.
2. $\sum_{n=0}^{\infty} t_{w,v}(n) z^n = \Big(T(z)\Big)^n$.

Exercise 12.2.3. *Prove*

$$T(z) \;=\; \frac{z}{k} + \frac{k-1}{k} z T^2(z). \tag{12.10}$$

Note that this explicitly gives us $T(z)$ by application of the quadratic formula:

$$T(z) \;=\; \frac{1 \pm \sqrt{1 - 4\Big(\frac{k-1}{k}z\Big)\frac{z}{k}}}{2\frac{k-1}{k}z}. \tag{12.11}$$

Now that we have $T(z)$ we have $Q(z)$, from which we get $R(z)$. $T(z)$ will have a square-root – it will be an algebraic function of z.

12.3 Recovering the Measure $f(x)$ from $R(z)$

We have

$$R(z) \;=\; \sum_{n=0}^{\infty} r_n z^n. \tag{12.12}$$

As we know $T(z)$, we know $R(z)$, hence we know the numbers r_n.
Now,

$$r_n \;=\; \int_{\infty}^{\infty} x^n f(x) dx \;=\; \int_{-k}^{k} x^n f(x) dx. \tag{12.13}$$

How do we recover $f(x)$ given the numbers r_n? We've now normalized the eigenvalues to lie in $[-1, 1]$.

$$
\begin{aligned}
R(z) &= \sum_{n=0}^{\infty}\left(\int_{-1}^{1} x^n f(x)dx\right) \\
&= \int_{-1}^{1}\left(\sum_{n=0}^{\infty}(xz)^n\right)f(x)dx \\
&= \int_{-1}^{1}\frac{f(x)}{1-xz}dx \\
B(z) \;=\; \frac{1}{z}R\Big(\frac{1}{z}\Big) &= \int_{-1}^{1}\frac{f(x)}{z-x}dx. \qquad (12.14)
\end{aligned}
$$

Suppose we know the LHS. Can we recover $f(x)$? If $z \in [-1,1]$, the function will have a singularity. Thus, if $f(x)$ is a nice function, we do not expect to be able to make sense of the above relation if $z \in [-1,1]$. We will, however, consider z *close* to the interval $[-1,1]$.

Let $z = \xi + iy$, $\xi \in [-1,1]$, $y > 0$. Later we will take $z = \xi - iy$.

Look at

$$
\begin{aligned}
B(\xi+iy) - B(\xi-iy) &= \int_{-1}^{1} f(x)\left[\frac{1}{\xi+iy-x} - \frac{1}{\xi-iy-x}\right]dx \\
&= 2i\int_{-1}^{1}\frac{yf(x)}{(\xi-x)^2+y^2}dx. \qquad (12.15)
\end{aligned}
$$

We will study the above as $y \to 0$.

12.3.1 Poisson Kernel

Recall $\xi \in [-1,1]$, $f(x)$ £xed, we are integrating $f(x)$ against the **Poisson Kernel**

$$
\frac{y}{(\xi-x)^2+y^2}. \qquad (12.16)
$$

As $y \to 0$, the above looks singular at $x = \xi$.

At $x = \xi$, the kernel has height $\frac{1}{y}$, which is quite large.

If $x = \xi + \epsilon$, then as $y \to 0$, the kernel goes to 0 very rapidly.

Basically, as $y \to 0$, the kernel becomes a higher, thinner spike centered at ξ.

Now

$$\begin{aligned}
\int_{\infty}^{\infty} \frac{y}{(x-\xi)^2+y^2} dx &= \int_{-\infty}^{\infty} \frac{y}{t^2+y^2} dt \\
&= \int_{-\infty}^{\infty} \frac{y^2}{y^2\eta^2+y^2} d\eta, \text{ from } \frac{t}{y} = \eta \\
&= \int_{-\infty}^{\infty} \frac{1}{1+\eta^2} d\eta \\
&= \pi. \qquad (12.17)
\end{aligned}$$

This is an **approximation to the identity**.

Thus,

$$B(\xi + iy) - B(\xi - iy) \ \rightarrow \ 2\pi i f(\xi). \qquad (12.18)$$

12.3.2 Cauchy Integral Formula

If you have an analytic function $f(z)$ and γ is a curve enclosing z then

$$\frac{1}{2\pi i} \int_{\gamma} \frac{f(\zeta)}{z-\zeta} d\zeta \qquad (12.19)$$

In our case above, we cannot apply Cauchy's Integral Formula, as our function $f(x)$ is not analytic. It is compactly supported, and no non-zero analytic function is compactly supported.

Call this permutation ϕ:

$$\phi : \mathbb{F}_p^* \rightarrow \mathbb{F}_p^*, \qquad (12.20)$$

where ϕ^2 is the identity.

Question 12.3.1. *Does ϕ behave like a random permutation?*

12.4 Third Problem

12.4.1 Introduction

Let p be a large prime, and consider the map

$$x \mapsto x^{-1} \bmod p,\ x \neq 0. \tag{12.21}$$

This is a map from $\mathbb{F}_p \to \mathbb{F}_p$. The map is not completely random, as

$$\begin{aligned} 1 &\mapsto 1 \\ 2 &\mapsto \frac{p+1}{2} \\ p-1 &\mapsto p-1. \end{aligned} \tag{12.22}$$

The map which sends $x \to x^{-1}$ is a permutation of $\mathbb{F}_p^*$. It is not a completely arbitrary permutation, as it pairs x with x^{-1} (away from a few very special x's, such as $p-1$ and 1).

Thus, this permutation is a product of transpositions.

12.4.2 Character Sums

Let

$$1 \leq A \leq B \leq p,\ B - A \text{ large}. \tag{12.23}$$

Let

$$\overline{m} \text{ be the inverse of } m \bmod p \tag{12.24}$$

IE, $m\overline{m} \equiv 1 \bmod p$.

Let

$$e(z) \ = \ 2^{2\pi i z}. \tag{12.25}$$

For $\nu \in \mathbb{Z}/p\mathbb{Z}$, consider

$$S \ = \ \sum_{A \leq m \leq B} e\left(\frac{\overline{m}\nu}{p}\right). \tag{12.26}$$

These sums will measure how equidistributed or random the map $m \to \overline{m}$ is.

Exercise 12.4.1. *Prove the trivial bound for $|S|$:*

$$|S| \leq B - A. \tag{12.27}$$

Let $N = B - A + 1$. By the Central Limit Theorem, with high probability if we add N random numbers of modulus one we expect square-root cancellation. Thus, we expect (if the inverse map is random) that $|S| \approx \sqrt{N}$.

12.4.3 Completing the Square

Let $a, b \in \mathbb{Z}$, and consider the **Kloosterman Sum**

$$\mathrm{Kl}(a, b, p) \;=\; \sum_{\substack{x \mod p \\ x \neq 0}} e\Big(\frac{ax + b\overline{x}}{p}\Big). \tag{12.28}$$

How large can the Kloosterman Sum be? If $a = b = 0$, then trivially $\mathrm{Kl}(0, 0, p) = p - 1$.

If $b = 0$ and $a \neq 0$ (or, by symmetry, the other way around) then

$$\begin{aligned}\mathrm{Kl}(a, b, p) \;&=\; \sum_{\substack{x \mod p \\ x \neq 0}} e\Big(\frac{ax}{p}\Big) \\ &=\; \sum_{\substack{y \mod p \\ y \neq 0}} e\Big(\frac{y}{p}\Big) \\ &=\; \sum_{y=0}^{p-1} e\Big(\frac{y}{p}\Big) \;-\; 1 \;=\; -1.\end{aligned} \tag{12.29}$$

Exercise 12.4.2. *Prove* $\sum_{y=0}^{p-1} e\Big(\frac{y}{p}\Big) = 0$. *Hint: Let* T *be this sum. Then show* $e\Big(\frac{1}{p}\Big)T = T$*; thus* $T = 0$.

12.4.4 Weil's Bound

Let $a \not\equiv 0 \bmod p$. Then

$$|\mathrm{Kl}(a, b, p)| \;\leq\; 2\sqrt{p}. \tag{12.30}$$

This is a very deep result.

How big is

$$\sum_{a \mod p} |\mathrm{Kl}(a, 1, p)|^2 \tag{12.31}$$

If we believe Weil's bound, each term is of size at most $2\sqrt{p}$, we square, then sum p terms. Thus, we expect a size of at most $4p^2$. We will show on average that Weil's bound is correct.

$$\begin{aligned}
\sum_{a \mod p} |\mathrm{Kl}(a,1,p)|^2 &= \sum_{a(p)} \Bigg| \sum_{\substack{x \mod p \\ x \neq 0}} e\Big(\frac{\overline{x} + ax}{p}\Big) \Bigg| \\
&= \sum_{a(p)} \sum_{\substack{x_1, x_2 \mod p \\ x_1, x_2 \neq 0}} e\Big(\frac{\overline{x}_1 - \overline{x}_2 + a(x_1 - x_2)}{p}\Big) \\
&= \sum_{\substack{x_1, x_2 \mod p \\ x_1, x_2 \neq 0}} e\Big(\frac{\overline{x}_1 - \overline{x}_2}{p}\Big) \sum_{a(p)} e\Big(\frac{a(x_1 - x_2)}{p}\Big) \\
&= (p-1)p, \qquad (12.32)
\end{aligned}$$

where the last line follows from the a-sum vanishes unless $x_1 = x_2$, which then collapses the sums. There are $p-1$ ways $x_1 = x_2$, and when this occurs, the a-sum gives p.

Exercise 12.4.3. *Consider*

$$\sum_{a(p)} |Kl(a,1,p)|^4. \qquad (12.33)$$

Above there are p-terms, each term of size $(2\sqrt{p})^4 = 16p^2$. Thus, show the sum is at most $16p^3$. You will £nd cp^3 for some c independent of p.

By looking at one term, as every summand is positive, we £nd

$$|\mathrm{Kl}(a,1,p)|^4 \leq cp^3. \qquad (12.34)$$

Thus, taking the fourth-root yields

$$|\mathrm{Kl}(a,1,p)| \leq c^{\frac{1}{4}} p^{\frac{3}{4}}. \qquad (12.35)$$

12.4.5 Fourier Expansion of Sums

De£ne the indicator function

$$I(y) = \begin{cases} 1 & A \leq y \leq A+N \\ 0 & \text{otherwise} \end{cases} \qquad (12.36)$$

Consider

$$
\begin{aligned}
S &= \sum_{A \le x \le A+N} e\Big(\frac{\nu\overline{x}}{p}\Big) \\
&= \sum_{x(p)} e\Big(\frac{\nu\overline{x}}{p}\Big) I(x). \qquad (12.37)
\end{aligned}
$$

We want to write $I(x)$ in terms of its Fourier Coef£cients

$$\widehat{I}(m) = \int_0^1 e(-mt)I(t)dt. \qquad (12.38)$$

Then

$$I(y) = \sum_{m=-\infty}^{\infty} \widehat{I}(m)e\Big(\frac{my}{p}\Big). \qquad (12.39)$$

12.4.6 Brief Review of Fourier Series

Consider the unit interval $[0, 1]$. De£ne

$$\phi_m(x) = e(mx). \qquad (12.40)$$

Then (if our function is suf£ciently nice)

$$f(x) = \sum_{m \in \mathbb{Z}} \widehat{f}(m)e(mx), \qquad (12.41)$$

where

$$\widehat{f}(m) = \int_0^1 f(x)e(-mx)dx. \qquad (12.42)$$

12.5 Fourier Analysis and the Equi-Distribution of $\{n\alpha\}$

12.5.1 Inner Product of Functions

We de£ne the exponential function by means of the series

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}, \tag{12.43}$$

which converges everywhere. Given the Taylor series expansion of $\sin x$ and $\cos x$, we can verify the identity

$$e^{ix} = \cos x + i \sin x. \tag{12.44}$$

Exercise 12.5.1. *Prove e^x converges for all $x \in \mathbb{R}$ (even better, for all $x \in \mathbb{C}$. Show the series for e^x also equals*

$$\lim_{n \to \infty} \left(1 + \frac{x}{n}\right)^n, \tag{12.45}$$

which you may remember from compound interest problems.

Exercise 12.5.2. *Prove, using the series de£nition, that $e^{x+y} = e^x e^y$. Use this fact to calculate the derivative of e^x. If instead you try to differentiate the series directly, you must justify the derivative of the in£nite sum is the in£nite sum of the derivatives.*

Remember the de£nition of **inner or dot product**: for two vectors $\vec{v} = (v_1, \cdots, v_n)$, $\vec{w} = (w_1, \cdots, w_n)$, we take the *inner product* $\vec{v} \cdot \vec{w}$ (also denoted $\langle v, w \rangle$) to mean

$$\vec{v} \cdot \vec{w} \ = \ \langle v, w \rangle = \sum_i v_i \bar{w}_i. \tag{12.46}$$

Further, the length of a vector v is

$$|v| \ = \ \langle v, v \rangle. \tag{12.47}$$

We generalize this for functions. For de£niteness, assume f and g are functions from $[0, 1]$ to $\mathbb{C}$. Divide the interval $[0, 1]$ into n equal pieces. Then we can represent the functions by

$$f(x) \ \longleftrightarrow \ \left(f(0), f\left(\frac{1}{n}\right), \dots, f\left(\frac{n-1}{n}\right)\right), \tag{12.48}$$

and similarly for g. Call these vectors f_n and g_n. As before, we consider

$$\langle f_n, g_n \rangle \;=\; \sum_{i=0}^{n-1} f\Big(\frac{i}{n}\Big) \cdot \bar{g}\Big(\frac{i}{n}\Big). \tag{12.49}$$

In general, as we continue to divide the interval ($n \to \infty$), the above sum diverges. For example, if f and g are identically 1, the above sum is n.

There is a natural rescaling: we multiply each term in the sum by $\frac{1}{n}$, the size of the sub-interval. Note for the constant function, the sum is now independent of n.

Thus, for good f and g we are led to

$$\langle f, g \rangle \;=\; \lim_{n\to\infty} \sum_{i=0}^{n-1} f\Big(\frac{i}{n}\Big) \cdot \bar{g}\Big(\frac{i}{n}\Big)\frac{1}{n} = \int_0^1 f(x)\overline{g(x)}dx. \tag{12.50}$$

The last result follows by Riemann Integration.

De£nition 12.5.3. *We say two continuous functions on* $[0,1]$ *are orthogonal (or perpendicular) if their dot product equals zero.*

Exercise 12.5.4. *Prove* x^n *and* x^m *are not perpendicular on* $[0,1]$ *for* $n \neq m$.

We will see that the exponential function behaves very nicely under the inner product. De£ne

$$e_n(x) = e^{2\pi i n x} \;\text{ for } n \in \mathbb{Z}. \tag{12.51}$$

Then a straightforward calculation shows

$$\langle e_n(x), e_m(x) \rangle = \begin{cases} 1 & \text{if } n = m \\ 0 & \text{otherwise.} \end{cases} \tag{12.52}$$

Thus $e_0(x), e_1(x), e_2(x), \cdots$ are an **orthogonal set** of functions, which means they are pairwise perpendicular. As each function has length 1, we say the functions $e_n(x)$ are an **orthonormal set** of functions.

Exercise 12.5.5. *Prove* $\langle e_n(x), e_m(x) \rangle$ *is* 1 *if* $n = m$ *and* 0 *otherwise.*

12.5.2 Fourier Series and $\{n\alpha\}$

Fourier Series

Let f be continuous and periodic on $\mathbb{R}$ with period one. De£ne the n**th Fourier coef£cient** $\hat{f}(n)$ of f to be

$$\hat{f}(n) = a_n = \langle f(x), e_n(x)\rangle = \int_0^1 f(x)e^{-2\pi inx}dx. \tag{12.53}$$

Returning to the intuition of $\mathbb{R}^m$, we can think of the $e_n(x)$'s as an in£nite set of perpendicular directions. The above is simply the projection of f in the direction of $e_n(x)$.

Exercise 12.5.6. *Show*

$$\langle f(x) - \hat{f}(n)e_n(x), e_n(x)\rangle = 0. \tag{12.54}$$

This agrees with our intuition, namely, that if you remove the projection in a certain direction, what is left is perpendicular to that direction.

The N^{th} **partial Fourier series** of f is

$$s_N(x) = \sum_{n=-N}^{N} \hat{f}(n)e_n(x). \tag{12.55}$$

Exercise 12.5.7. *Prove*

1. $\langle f(x) - s_N(x), e_n(x)\rangle = 0$ *if* $|n| \leq N$.
2. $|\hat{f}(n)| \leq \int_0^1 |f(x)|dx$.
3. *If* $\langle f, f\rangle < \infty$*, then* $\sum_{n=-\infty}^{\infty} |\hat{f}(n)|^2 \leq \langle f, f\rangle$.
4. *If* $\langle f, f\rangle < \infty$*, then* $\lim_{|n|\to\infty} \hat{f}(n) = 0$.

As $\langle f(x) - s_N(x), e_n(x)\rangle = 0$ if $|n| \leq N$, we might think that we just have to let N go to in£nity to obtain a series s_∞ such that

$$\langle f(x) - s_\infty(x), e_n(x)\rangle = 0. \tag{12.56}$$

Assume that for a periodic function $g(x)$ to be orthogonal to $e_n(x)$ for every n it must be zero for every x. Then $f(x) - s_\infty(x) = 0$, and hence $f = s_\infty$. Voilá – an expression for f as a sum of exponentials! Be careful, however. We have just glossed over the two central issues – completeness and, even worse, convergence. We will now see a way of avoiding some of our problems.

Weighted partial sums

De£ne

$$D_N(x) = \sum_{n=-N}^{N} e_n(x) = \frac{\sin((2N+1)\pi x)}{\sin \pi x},$$
$$F_N(x) = \frac{\sin^2(N\pi x)}{N \sin^2 \pi x} = \frac{1}{N}\sum_{n=0}^{N-1} D_n(x). \tag{12.57}$$

Here F stands for Féjer, D for Dirichlet. In general, functions which we are interested in taking their inner product against f are called **kernels**; thus, the Dirichlet kernel, the Féjer kernel, etc.

Note that, no matter what N is, $F_N(x)$ is positive for all x.

We say that a sequence $f_1(x), f_2(x), f_3(x), \dots$ of functions is an **approximation to the identity** if

1. $f_N(x) \geq 0$ for all x and every N;
2. $\int_0^1 f_N(x)dx = 1$;
3. $\lim_{N\to\infty} \int_\delta^{1-\delta} f_N(x)dx = 0$ if $0 < \delta < \frac{1}{2}$.

Theorem 12.5.8. *The Féjer kernels $F_1(x), F_2(x), F_3(x), \dots$ are an approximation to the identity.*

Proof: The £rst property is immediate. The second follows from the observation that $F_N(x)$ can be written as

$$F_N(x) = e_0(x) + \frac{N-1}{N}\Big(e_{-1}(x) + e_1(x)\Big) + \cdots, \tag{12.58}$$

and all integrals are zero but the £rst, which is 1.
To prove the third property, note that $F_N(x) \leq \frac{1}{N\sin^2 \pi\delta}$ for $\delta \leq x \leq 1-\delta$. □

Let f be a continuous, periodic function on $\mathbb{R}$ with period one. Thus, we can consider f as a function on just $[0,1]$, with $f(0) = f(1)$. De£ne

$$T_N(x) = \int_0^1 f(y) F_N(x-y)dy. \tag{12.59}$$

Recall the following de£nition and theorem:

De£nition 12.5.9 (Uniform Continuity). *A continuous function is uniformly continuous if given an $\epsilon > 0$, there exists a $\delta > 0$ such that $|x - y| < \delta$ implies $|f(x) - f(y)| < \epsilon$. Note that the same δ works for all points.*

Theorem 12.5.10. *Any continuous function on a closed, compact interval is uniformly continuous.*

Exercise 12.5.11. *Show x^n is uniformly continuous on $[a, b]$ for $-\infty < a < b < \infty$.*

Theorem 12.5.12. *Given $\epsilon > 0$, there is an N such that*

$$|f(x) - T_N(x)| \leq \epsilon \tag{12.60}$$

for every $x \in [0, 1]$.

Proof. For any positive N,

$$\begin{aligned} T_N(x) - f(x) &= \int_0^1 f(x - y) F_N(y) dy - f(x) \cdot 1 \\ &= \int_0^1 f(x - y) F_N(y) dy - \int_0^1 f(x) F_n(y) dy \text{ (property 2 of } F_N\text{)} \\ &= \int_0^\delta \Big(f(x - y) - f(x) \Big) F_N(y) dy \\ &+ \int_\delta^{1-\delta} \Big(f(x - y) - f(x) \Big) F_N(y) dy \\ &+ \int_{1-\delta}^1 \Big(f(x - y) - f(x) \Big) F_N(y) dy. \end{aligned} \tag{12.61}$$

Let $\delta \in (0, 1/2)$. Then, using the fact that the $F_N(x)$'s are an approximation to the identity, we £nd

$$\left| \int_\delta^{1-\delta} \Big(f(x - y) - f(x) \Big) F_N(y) dy \right| \leq 2 \max |f(x)| \cdot \int_\delta^{1-\delta} F_N(y) dy. \tag{12.62}$$

Since

$$\lim_{N \to \infty} \int_\delta^{1-\delta} F_N(y) dy = 0, \tag{12.63}$$

we obtain

$$\lim_{N\to\infty}\int_{\delta}^{1-\delta}(f(x-y)-f(x))F_N(y)dy=0. \tag{12.64}$$

Thus, by choosing N large enough (where large depends on δ), we can insure that this piece is at most $\frac{\epsilon}{3}$.

It remains to estimate what happens near zero. Since f is continuous and $[0,1]$ is compact, f is uniformly continuous. Thus, we can choose δ small enough that $|f(x-y)-f(x)|<\frac{\epsilon}{3}$ for any x and any positive $y<\delta$. Then

$$\left|\int_0^{\delta}\Big(f(x-y)-f(x)\Big)F_N(y)dy\right|\le\int_0^{\delta}\frac{\epsilon}{3}F_N(y)dy\le\frac{\epsilon}{3}\int_0^1F_N(y)dy\le\frac{\epsilon}{3}. \tag{12.65}$$

Similarly

$$\left|\int_{1-\delta}^{1}\Big(f(x-y)-f(x)\Big)F_N(y)dy\right|\le\frac{\epsilon}{3}. \tag{12.66}$$

Therefore

$$|T_N(x)-f(x)|\le\epsilon \tag{12.67}$$

for all N suf£ciently large. □

De£nition 12.5.13 (Trigonometric Polynomials). *Any £nite linear combination of the functions $e_n(x)$ is called a trigonometric polynomial.*

From Theorem 12.5.12 we immediately get the Stone-Weierstrass theorem:

Theorem 12.5.14 (Stone-Weierstrass). *Any continuous period function can be uniformly approximated by trigonometric polynomials.*

12.5.3 Equidistribution

We say that a sequence $\{x_n\}$, $x_n\in[0,1]$ is *equidistributed* if

$$\lim_{N\to\infty}\frac{1}{2N+1}\#\{n:|n|\le N,x_n\in(a,b)\}=b-a \tag{12.68}$$

for all $(a,b)\subset[0,1]$.

Theorem 12.5.15 (Weyl). *Let α be an irrational number in $[0,1]$. Let $x_n = \{n\alpha\}$, where $\{y\}$ denotes the fractional part of y. Then the sequence $\{x_n\}$ is equidistributed.*

Proof. We will estimate $\frac{1}{2N+1}\sum_{-N}^{N}\chi_{(a,b)}(x_n)$ as $N \to \infty$, where $\chi_{(a,b)}$ is the function taking the value 0 outside (a,b) and 1 inside (a,b). We call $\chi_{(a,b)}$ the **characteristic function** of the interval (a,b).

Thus, we must show

$$\lim_{N\to\infty}\frac{1}{2N+1}\sum_{n=-N}^{N}\chi_{(a,b)}(x_n) = b-a. \tag{12.69}$$

Consider $e_k(x) = e^{2\pi ikx}$. Since $x_n = \{n\alpha\} = n\alpha - [n\alpha]$ and $e_k(x) = e_k(x+m)$ for every integer m,

$$e_k(x_n) = e^{2\pi ikn\alpha}. \tag{12.70}$$

Hence

$$\begin{aligned}\frac{1}{2N+1}\sum_{n=-N}^{N}e_k(x_n) &= \frac{1}{2N+1}\sum_{n=-N}^{N}e_k(n\alpha)\\ &= \frac{1}{2N+1}\sum_{n=-N}^{N}(e^{2\pi ik\alpha})^n\\ &= \begin{cases}1 & \text{if } k=0\\ \frac{1}{2N+1}\frac{e_k(-N\alpha)-e_k((N+1)\alpha)}{1-e_k(\alpha)} & \text{if } k>0.\end{cases}\end{aligned} \tag{12.71}$$

Now for a £xed irrational α, $|1-e_k(\alpha)| > 0$. Therefore if $k \neq 0$:

$$\lim_{N\to\infty}\frac{1}{2N+1}\frac{e_k(-N\alpha)-e_k((N+1)\alpha)}{1-e_k(\alpha)} = 0. \tag{12.72}$$

Let $P(x) = \sum_k a_k e_k(x)$ be a £nite sum (ie, $P(x)$ is a trigonometric polynomial). By possibly adding some zero coef£cients, we can write $P(x)$ as a sum over a symmetric range: $P(x) = \sum_{k=-K}^{K} a_k a_k(x)$.

Exercise 12.5.16. *Show* $\int_0^1 P(x)dx = a_0$.

By the above arguments, we have shown that for any (£nite) trigonometric polynomial $P(x)$:

$$\lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} P(x_n) \;\to\; a_0 \;=\; \int_0^1 P(x)dx. \tag{12.73}$$

Consider two approximations to the characteristic function $\chi_{(a,b)}$:

1. f_{1m}: $f_{1m}(x) = 1$ if $a + \frac{1}{m} \leq x \leq b - \frac{1}{m}$, drops linearly to 0 at a and b, and is zero elsewhere.

2. f_{2m}: $f_{1m}(x) = 1$ if $a \leq x \leq b$, drops linearly to 0 at $a - \frac{1}{m}$ and $b + \frac{1}{m}$, and is zero elsewhere.

Note there are trivial modi£cations if $a = 0$ or $b = 1$. Clearly

$$f_{1m}(x) \;\leq\; \chi_{(a,b)}(x) \;\leq\; f_{2m}(x). \tag{12.74}$$

Therefore

$$\frac{1}{2N+1} \sum_{n=-N}^{N} f_{1m}(x_n) \;\leq\; \frac{1}{2N+1} \sum_{n=-N}^{N} \chi_{(a,b)}(x_n) \;\leq\; \frac{1}{2N+1} \sum_{n=-N}^{N} f_{2m}(x_n). \tag{12.75}$$

By Theorem 12.5.12, for each m, given $\epsilon > 0$ we can £nd trigonometric polynomials $P_{1m}(x)$ and $P_{2m}(x)$ such that $|P_{1m}(x) - f_{1m}(x)| < \epsilon$ and $|P_{2m}(x) - f_{2m}(x)| < \epsilon$.

As f_{1m} and f_{2m} are continuous functions, we can replace

$$\frac{1}{2N+1} \sum_{n=-N}^{N} f_{im}(x_n) \;\text{ with }\; \frac{1}{2N+1} \sum_{n=-N}^{N} P_{im}(x_n) \tag{12.76}$$

at a cost of at most ϵ.

As $N \to \infty$,

$$\frac{1}{2N+1} \sum_{n=-N}^{N} P_{im}(x_n) \;\to\; \int_0^1 P_{im}(x)dx. \tag{12.77}$$

But $\int_0^1 P_{1m}(x)dx = (b-a) - \frac{1}{m}$ and $\int_0^1 P_{2m}(x)dx = (b-a) + \frac{1}{m}$. Therefore, given m and ϵ, we can choose N large enough so that

$$(b-a) - \frac{1}{m} - \epsilon \;\leq\; \frac{1}{2N+1} \sum_{n=-N}^{N} \chi_{(a,b)}(x_n) \;\leq\; (b-a) + \frac{1}{m} + \epsilon. \tag{12.78}$$

Letting m tend to ∞ and ϵ tend to 0, we see $\frac{1}{2N+1}\sum_{n=-N}^{N} \chi_{(a,b)}(x_n) \to b - a$. □

Exercise 12.5.17. *Rigorously do the necessary book-keeping to prove the previous theorem.*

Exercise 12.5.18. *Prove*

1. *If $\alpha \in \mathbb{Q}$, then $\{n\alpha\}$ is periodic.*
2. *If $\alpha \notin \mathbb{Q}$, then no two $\{n\alpha\}$ are equal.*

Chapter 13

Liouville's Theorem Constructing Transcendentals

1. We prove Liouville's Theorem for the order of approximation by rationals of real algebraic numbers.
2. We construct several transcendental numbers.
3. We de£ne Poissonian Behaviour, and study the spacings between the ordered fractional parts of $\{n^k\alpha\}$.

Lecture by Steven J. Miller; notes for the £rst two by Steven J. Miller and Florin Spinu; notes for the third by Steven J. Miller.

13.1 Review of Approximating by Rationals

De£nition 13.1.1 (Approximated by rationals to order n). *A real number x is approximated by rationals to order n if there exist a constant $k(x)$ (possibly depending on x) such that there are in£nitely many rational $\frac{p}{q}$ with*

$$\left| x - \frac{p}{q} \right| < \frac{k(x)}{q^n}. \tag{13.1}$$

Recall that Dirichlet's Box Principle gaves us:

$$\left| x - \frac{p}{q} \right| < \frac{1}{q^2} \tag{13.2}$$

for in£ntely many fractions $\frac{p}{q}$. This was proved by choosing a large parameter Q, and considering the $Q+1$ fractionary parts $\{qx\} \in [0,1)$ for $q \in \{0,\ldots,Q\}$. The box principle ensures us that there must be two different q's, say:

$$0 \leq q_1 < q_2 \leq Q \tag{13.3}$$

such that both $\{q_1x\}$ and $\{q_2x\}$ belong to the same interval $[\frac{a}{Q}, \frac{a+1}{Q})$, for some $0 \leq a \leq Q-1$. Note that there are exactly Q such intervals partitioning $[0,1)$, and $Q+1$ fractionary parts! Now, the length of such an interval is $\frac{1}{Q}$ so we get

$$|\{q_2x\} - \{q_1x\}| < \frac{1}{Q}. \tag{13.4}$$

There exist integers p_1 and p_2 such that

$$\{q_1x\} = q_1x - p,\ \{q_2x\} = q_2x - p. \tag{13.5}$$

Letting $p = p_2 - p_1$ we £nd

$$|(q_2 - q_1)x - p| \leq \frac{1}{Q} \tag{13.6}$$

Let $q = q_2 - q_1$, so $1 \leq q \leq Q$, and the previous equation can be rewriten as

$$\left|x - \frac{p}{q}\right| < \frac{1}{qQ} \leq \frac{1}{q^2} \tag{13.7}$$

Now, letting $Q \to \infty$, we get an in£nite collection of rational fractions $\frac{p}{q}$ satisfying the above equation. If this collection contains only £nitely many distinct fractions, then one of these fractions, say $\frac{p_0}{q_0}$, would occur for in£ntely many choices Q_k of Q, thus giving us:

$$\left|x - \frac{p_0}{q_0}\right| < \frac{1}{qQ_k} \to 0, \tag{13.8}$$

as $k \to \infty$. This implies that $x = \frac{p_0}{q_0} \in \mathbb{Q}$. So, unles x is a rational number, we can £nd in£nitely many *distinct* rational numbers $\frac{p}{q}$ satisfying Equation 13.7. This means that any real, irrational number can be approximated to order $n = 2$ by rational numbers.

13.2 Liouville's Theorem

Theorem 13.2.1 (Liouville's Theorem). *Let x be a real algebraic number of degree n. Then x is approximated by rationals to order at most n.*

Proof. Let

$$f(X) \ = \ a_n X^n + \cdots a_1 X + a_0 \tag{13.9}$$

be the polynomial with integer coef£cients of smallest degree (minimal polynomial) such that x satis£es

$$f(x) \ = \ 0. \tag{13.10}$$

Note that $\deg x = \deg f$ and the condition of minimality implies that $f(X)$ is irreducible over $\mathbb{Z}$. Further, a well known result from algebra states that a polynomial irreducible over $\mathbb{Z}$ is also irreducible over $\mathbb{Q}$.

In particular, as $f(X)$ is irreducible over $\mathbb{Q}$, $f(X)$ does not have any rational roots. If it did, then $f(X)$ would be divisible by a linear polynomial $(X - \frac{a}{b})$. Let $G(X) = \frac{f(X)}{X - \frac{a}{b}}$. Clear denominators (multiply throughout by b), and let $g(X) = bG(X)$. Then $\deg g = \deg f - 1$, and $g(x) = 0$. This contradicts the minimality of f (we choose f to be a polynomial of smallest degree such that $f(x) = 0$). Therefore, f is non-zero at every rational.

Let

$$M \ = \ \sup_{|z-x|<1} |f'(z)|. \tag{13.11}$$

Let now $\frac{p}{q}$ be a rational such that $\left|x - \frac{p}{q}\right| < 1$. The Mean Value Theorem gives us that

$$\left|f\left(\frac{p}{q}\right) - f(x)\right| = \left|f'(c)\left(x - \frac{p}{q}\right)\right| \leq M\left|x - \frac{p}{q}\right| \tag{13.12}$$

where c is some real number between x and $\frac{p}{q}$; $|c - x| < 1$ for $\frac{p}{q}$ moderately close to x.

Now we use the fact that $f(X)$ does not have any rational roots:

$$0 \neq f\left(\frac{p}{q}\right) = a_n\left(\frac{p}{q}\right)^n + \cdots + a_0 = \frac{a_n p^n + \cdots a_1 p^{n-1} q + a_0 q^n}{q^n} \tag{13.13}$$

The numerator of the last term is a nonzero integer, hence it has absolute value at least 1. Since we also know that $f(x) = 0$ it follows that

$$\left| f\left(\frac{p}{q}\right) - f(x) \right| = \left| f\left(\frac{p}{q}\right) \right| = \frac{|a_n p^n + \cdots a_1 p^{n-1} q + a_0 q^n|}{q^n} \geq \frac{1}{q^n}. \quad (13.14)$$

Combining the equations 13.12 and 13.14, we get:

$$\frac{1}{q^n} \leq M \left| x - \frac{p}{q} \right| \Rightarrow \frac{1}{Mq^n} \leq \left| x - \frac{p}{q} \right| \quad (13.15)$$

whenever $|x - \frac{p}{q}| < 1$. This last equation shows us that x can be approximated by rationals to order at most n. For assume it was otherwise, namely that x can be approximated to order $n + \epsilon$. Then we would have an in£nite sequence of distinct rational numbers $\{\frac{p_i}{q_i}\}_{i\geq 1}$ and a constant $k(x)$ depending only on x such that

$$\left| x - \frac{p_i}{q_i} \right| < \frac{k(x)}{q_i^{n+\epsilon}}. \quad (13.16)$$

Since the numbers $\frac{p_i}{q_i}$ converge to x we can assume that they already are in the interval $(x - 1, x + 1)$. Hence they also satisfy Equation 13.15:

$$\frac{1}{q_i^n} \leq M \left| x - \frac{p_i}{q_i} \right|. \quad (13.17)$$

Combining the last two equations we get

$$\frac{1}{Mq_i^n} \leq \left| x - \frac{p_i}{q_i} \right| < \frac{k(x)}{q_i^{n+\epsilon}}, \quad (13.18)$$

hence

$$q_i^\epsilon < M \quad (13.19)$$

and this is clearly impossible for arbitrarily large q since $\epsilon > 0$ and $q_i \to \infty$.

□

Exercise 13.2.2. *Justify the fact that if $\{\frac{p_i}{q_i}\}_{i\geq 1}$ is a rational approximation to order $n \geq 1$ of x, then $q_i \to \infty$.*

Remark 13.2.3. *So far we have seen that the order to which an algebraic number can be approximated by rationals is bounded by its degree. Hence if a real, irrational number $\alpha \notin \mathbb{Q}$ can be approximated by rationals to an arbitrary large order, then α must be transcendental! This provides us with a recipe for constructing transcendental numbers.*

13.3 Constructing Transcendental Numbers

13.3.1 $\sum_m 10^{-m!}$

The following construction of transcendental numbers is due to Liouville.

Theorem 13.3.1. *The number*

$$x = \sum_{m=1}^{\infty} \frac{1}{10^{m!}} \tag{13.20}$$

is transcendental.

Proof. The series de£ning x is convergent, since it is dominated by the geometric series $\sum \frac{1}{10^m}$. In fact, the series converges very rapidly and it is this high rate of convergence that will yield x is transcendental.

Fix N large, and let $n > N$. Write

$$\frac{p_n}{q_n} = \sum_{m=1}^{n} \frac{1}{10^{m!}} \tag{13.21}$$

with $p_n, q_n > 0$ and $(p_n, q_n) = 1$. Then $\{\frac{p_n}{q_n}\}_{n\geq 1}$ is a monotone increasing sequence converging to x. In particular, all these rational numbers are distinct. Not also that q_n must divide $10^{n!}$, which implies

$$q_n \leq 10^{n!}. \tag{13.22}$$

Using this, we get

$$
\begin{aligned}
0 < x - \frac{p_n}{q_n} &= \sum_{m>n} \frac{1}{10^{m!}} = \frac{1}{10^{(n+1)!}} \left(1 + \frac{1}{10^{n+2}} + \frac{1}{10^{(n+2)(n+3)}} + \cdots\right) \\
&< \frac{2}{10^{(n+1)!}} = \frac{2}{(10^{n!})^{n+1}} \\
&< \frac{2}{q_n^{n+1}} \leq \frac{2}{q_n^N}. \qquad (13.23)
\end{aligned}
$$

This gives an approximation by rationals of order N of x. Since N can be chosen arbitrarily large, this implies that x can be approximated by rationals to arbitrary order. We can conclude, in view of our precious remark 13.2.3 that x is transcendental. □

13.3.2 $[10^{1!}, 10^{2!}, \dots]$

Theorem 13.3.2. *The number*

$$
y = [10^{1!}, 10^{2!}, \dots] \qquad (13.24)
$$

is transcendental.

Proof. Let $\frac{p_n}{q_n}$ be the continued fraction of $[10^{1!} \cdots 10^{n!}]$. Then

$$
\begin{aligned}
\left| y - \frac{p_n}{q_n} \right| &= \frac{1}{q_n q'_{n+1}} = \frac{1}{q_n(a'_{n+1} q_n + q_{n-1})} \\
&< \frac{1}{a_{n+1}} = \frac{1}{10^{(n+1)!}}. \qquad (13.25)
\end{aligned}
$$

Since $q_k = a_n q_{k-1} + q_{n-2}$, it implies that $q_k > q_{k-1}$ Also, $q_{k+1} = a_{k+1} q_n + q_{k-1}$, so we get

$$
\frac{q_{k+1}}{q_k} = a_{k+1} + \frac{q_{k-1}}{q_k} < a_{k+1} + 1. \qquad (13.26)
$$

Hence writing this inequality for $k = 1, \cdots, n-1$ we obtain

$$
\begin{aligned}
q_n = q_1 \frac{q_2}{q_1}\frac{q_3}{q_2}\cdots\frac{q_n}{q_{n-1}} &< (a_1+1)(a_2+1)\cdots(a_n+1) \\
&= (1+\frac{1}{a_1})\cdots(1+\frac{1}{a_n})a_1\cdots a_n \\
&< 2^n a_1\cdots a_n = 2^n 10^{1!+\cdots+n!} \\
&< 10^{2n!} = a_n^2
\end{aligned} \tag{13.27}
$$

Combining equations 13.25 and 13.27 we get:

$$
\begin{aligned}
\left| y - \frac{p_n}{q_n} \right| &< \frac{1}{a_{n+1}} = \frac{1}{a_n^{n+1}} \\
&< \left(\frac{1}{a_n^2}\right)^{\frac{n}{2}} < \left(\frac{1}{q_n^2}\right)^{\frac{n}{2}} \\
&= \frac{1}{q_n^{n/2}}.
\end{aligned} \tag{13.28}
$$

In this way we get, just as in the previous theorem, an approximation of y by rationals to arbitrary order. This proves that y is transcendental.

□

13.3.3 Buffon's Needle and π

Consider a collection of in£nitely long parallel lines in the plane, where the spacing between any two adjacent lines is d. Let the lines be located at $x = 0, \pm d, \pm 2d, \dots$. Consider a rod of length l, where for convenience we assume $l < d$.

If we were to *randomly* throw the rod on the plane, what is the probability it hits a line? This question was £rst asked by Buffon in 1733.

Because of the vertical symmetry, we may assume the center of the rod lies on the line $x = 0$, as shifting the rod (without rotating it) up or down will not alter the number of intersections. By the horizontal symmetry, we may assume $-\frac{d}{2} \leq x < \frac{d}{2}$. We posit that all values of x are equally likely. As x is continuous distributed, we may add in $x = \frac{d}{2}$ without changing the probability. The probability density function of x is $\frac{dx}{d}$.

Let θ be the angle the rod makes with the x-axis. As each angle is equally likely, the probability density function of θ is $\frac{d\theta}{2\pi}$.

We assume that x and θ are chosen independently. Thus, the probability density for (x, θ) is $\frac{dxd\theta}{d \cdot 2\pi}$.

The projection of the rod (making an angle of θ with the x-axis) along the x-axis is $l \cdot |\cos\theta|$. If $|x| \leq l \cdot |\cos\theta|$, then the rod hits exactly one vertical line exactly once; if $x > l \cdot |\cos\theta|$, the rod does not hit a vertical line. Note that if $l > d$, a rod could hit multiple lines, making the arguments more involved.

Thus, the probability a rod hits a line is

$$\begin{aligned} p &= \int_{\theta=0}^{2\pi} \int_{x=-l \cdot |\cos\theta|}^{l \cdot |\cos\theta|} \frac{dxd\theta}{d \cdot 2\pi} \\ &= \int_{\theta=0}^{2\pi} \frac{l \cdot |\cos\theta|}{d} \frac{d\theta}{2\pi} \\ &= \frac{2l}{\pi d}. \end{aligned} \tag{13.29}$$

Exercise 13.3.3. *Show*

$$\frac{1}{2\pi} \int_0^{2\pi} |\cos\theta| d\theta \ = \ \frac{2}{\pi}. \tag{13.30}$$

Let A be the random variable which is the number of intersections of a rod of length l thrown against parallel vertical lines separated by $d > l$ units. Then

$$A = \begin{cases} 1 & \text{with probability } \frac{2l}{\pi d} \\ 0 & \text{with probability } 1 - \frac{2l}{\pi d} \end{cases}. \tag{13.31}$$

If we were to throw N rods independently, since the expected value of a sum is the sum of the expected values (Lemma 6.3.8), we expect to observe

$$N \cdot \frac{2l}{\pi d} \tag{13.32}$$

intersections.

Turning this around, let us throw N rods, and let I be the number of observed intersections of the rods with the vertical lines. Then

$$I \ \approx \ N \cdot \frac{2l}{\pi d} \quad \to \quad \pi \ \approx \ \frac{N}{I} \cdot \frac{2l}{d}. \tag{13.33}$$

The above is an *experimental* formula for π!

Chapter 14

Poissonian Behavior and $\{n^k\alpha\}$

We now de£ne Poissonian Bevahior, and investigate the normalized spacings of the fractional parts of $n^2\alpha$. Lecture and notes by Steven J. Miller.

14.1 Equidistribution

We say a sequence of number $x_n \in [0,1)$ is equidistributed if

$$\lim_{N\to\infty} \frac{\#\{n : 1 \leq n \leq N \text{ and } x_n \in [a,b]\}}{N} = b - a \tag{14.1}$$

for any subinterval $[a,b]$ of $[0,1]$.

Recall Weyl's Result: If $\alpha \notin \mathbb{Q}$, then the fractional parts $\{n\alpha\}$ are equidistributed. Equivalently, $n\alpha$ mod 1 is equidistributed.

Similarly, one can show that for any integer k, $\{n^k\alpha\}$ is equidistributed. See Robert Lipshitz's paper for more details.

14.2 Point Masses and Induced Probability Measures

Recall from physics the concept of a unit point mass located at $x = a$. Such a point mass has no length (or, in higher dimensions, width or height), but £nite mass. As mass is the integral of the density over space, a £nite mass in zero volume (or zero length on the line) implies an in£nite density.

We can make this more precise by the notion of an Approximation to the Identity.

Definition 14.2.1 (Approximation to the Identity). *A sequence of functions $g_n(x)$ is an approximation to the identity (at the origin) if*

1. $g_n(x) \geq 0$.
2. $\int g_n(x)dx = 1$.
3. *Given $\epsilon, \delta > 0$ there exists $N > 0$ such that for all $n > N$, $\int_{|x|>\delta} g_n(x)dx < \epsilon$.*

We represent the limit of any such family of $g_n(x)$s by $\delta(x)$.

If $f(x)$ is a nice function (say near the origin its Taylor Series converges) then

$$\int f(x)\delta(x)dx \;=\; \lim_{n\to\infty}\int f(x)g_n(x) \;=\; f(0). \tag{14.2}$$

Exercise 14.2.2. *Prove Equation 14.2.*

Thus, in the limit the functions g_n are acting like point masses. We can consider the probability densities $g_n(x)dx$ and $\delta(x)dx$. For $g_n(x)dx$, as $n \to \infty$, almost all the probability is concentrated in a narrower and narrower band about the origin; $\delta(x)dx$ is the limit with all the mass at one point. It is a discrete (as opposed to continuous) probability measure.

Note that $\delta(x-a)$ acts like a point mass; however, instead of having its mass concentrated at the origin, it is now concentrated at a.

Exercise 14.2.3. *Let*

$$g_n(x) \;=\; \begin{cases} n & \text{if } |x| \leq \frac{1}{2n} \\ 0 & \text{otherwise} \end{cases} \tag{14.3}$$

Prove $g_n(x)$ is an approximation to the identity at the origin.

Exercise 14.2.4. *Let*

$$g_n(x) \;=\; c\frac{\frac{1}{n}}{\frac{1}{n^2}+x^2}. \tag{14.4}$$

Find c such that the above is an approximation to the identity at the origin.

Given N point masses located at x_1, $x_2, \ldots, x_N$, we can form a probability measure

$$\mu_N(x)dx = \frac{1}{N}\sum_{n=1}^{N}\delta(x - x_n)dx. \tag{14.5}$$

Note $\int \mu_N(x)dx = 1$, and if $f(x)$ is a nice function,

$$\int f(x)\mu_N(x)dx \;=\; \frac{1}{N}\sum_{n=1}^{N} f(x_n). \tag{14.6}$$

Exercise 14.2.5. *Prove Equation 14.6 for nice $f(x)$.*

Note the right hand side of Equation 14.6 looks like a Riemann sum. Or it *would* look like a Riemann sum if the x_ns were equidistributed. In general the x_ns will not be equidistributed, but assume for any interval $[a, b]$ that as $N \to \infty$, the fraction of x_ns $(1 \leq n \leq N)$ in $[a, b]$ goes to $\int_a^b p(x)dx$ for some nice function $p(x)$:

$$\lim_{N\to\infty} \frac{\#\{n : 1 \leq n \leq N \text{ and } x_n \in [a, b]\}}{N} \;\to\; \int_a^b p(x)dx. \tag{14.7}$$

In this case, if $f(x)$ is nice (say twice differentiable, with £rst derivative uniformly bounded), then

$$\begin{aligned}
\int f(x)\mu_N(x)dx &= \frac{1}{N}\sum_{n=1}^{N} f(x_n) \\
&\approx \sum_{k=-\infty}^{\infty} f\Big(\frac{k}{N}\Big)\frac{\#\{n : 1 \leq n \leq N \text{ and } x_n \in \left[\frac{k}{N}, \frac{k+1}{N}\right]\}}{N} \\
&\to \int f(x)p(x)dx.
\end{aligned} \tag{14.8}$$

De£nition 14.2.6 (Convergence to $p(x)$). *If the sequence of points x_n satis£es Equation 14.7 for some nice function $p(x)$, we say the probability measures $\mu_N(x)dx$ converge to $p(x)dx$.*

14.3 Neighbor Spacings

We now consider £ner questions. Let α_n be a collection of points in $[0,1)$. We order them by size:

$$0 \leq \alpha_{\sigma(1)} \leq \alpha_{\sigma(2)} \leq \cdots \leq \alpha_{\sigma(N)}, \tag{14.9}$$

where σ is a permutation of $123\cdots N$. Note the ordering depends crucially on N. Let $\beta_j = \alpha_{\sigma(j)}$.

We consider how the differences $\beta_{j+1} - \beta_j$ are distributed. We will use a slightly different de£nition of distance, however.

Recall $[0,1)$ is equivalent to the unit circle under the map $x \to e^{2\pi i x}$. Thus, the numbers $.999$ and $.001$ are actually very close; however, if we used the standard de£nition of distance, then $|.999 - .001| = .998$, which is quite large. Wrapping $[0,1)$ on itself (identifying 0 and 1), we see that $.999$ and $.001$ are separated by $.002$.

De£nition 14.3.1 (mod 1 distance). *Let $x, y \in [0,1)$. We de£ne the mod 1 distance from x to y, $||x-y||$, by*

$$||x-y|| \ = \ \min\Big\{|x-y|,\ 1-|x-y|\Big\}. \tag{14.10}$$

Exercise 14.3.2. *Show that the mod 1 distance between any two numbers in $[0,1)$ is at most $\frac{1}{2}$.*

In looking at spacings between the β_js, we have $N-1$ pairs of neighbors:

$$(\beta_2,\beta_1),\ (\beta_3,\beta_2),\ \ldots,\ (\beta_N,\beta_{N-1}). \tag{14.11}$$

These pairs give rise to spacings $\beta_{j+1} - \beta_j \in [0,1)$.

We can also consider the pair (β_1, β_N). This gives rise to the spacing $\beta_1 - \beta_N \in [-1,0)$; however, as we are studying this sequence mod 1, this is equivalent to $\beta_1 - \beta_N + 1 \in [0,1)$.

Henceforth, whenever we perform any arithmetic operation, we always mean mod 1; thus, our answers always live in $[0,1)$

De£nition 14.3.3 (Neighbor Spacings). *Given a sequence of numbers α_n in $[0,1)$, £x an N and arrange the numbers α_n ($n \leq N$) in increasing order. Label the new sequence β_j; note the ordering will depend on N. Let $\beta_{-j} = \beta_{N-j}$ and $\beta_{N+j} = \beta_j$.*

1. *The nearest neighbor spacings are the numbers $\beta_{j+1} - \beta_j$, $j = 1$ to N.*

2. *The k^{th}-neighbor spacings are the numbers $\beta_{j+k} - \beta_j$, $j = 1$ to N.*

Remember to take the differences $\beta_{j+k} - \beta_j$ mod 1.

Exercise 14.3.4. *Let $\alpha = \sqrt{2}$, and let $\alpha_n = \{n\alpha\}$ or $\{n^2\alpha\}$. Calculate the nearest neighbor and the next-nearest neighbor spacings in each case for $N = 10$.*

De£nition 14.3.5 (wrapped unit interval). *We call $[0, 1)$, when all arithmetic operations are done mod 1, the wrapped unit interval.*

14.4 Poissonian Behavior

Let $\alpha \notin \mathbb{Q}$. Fix a positive integer k, and let $\alpha_n = \{n^k\alpha\}$. As $N \to \infty$, look at the ordered α_ns, denoted by β_n. How are the nearest neighbor spacings of β_n distributed? How does this depend on k? On α? On N?

Before discussing this problem, we consider a simpler case. Fix N, and consider N independent random variables x_n. Each random variable is chosen from the uniform distribution on $[0, 1)$; thus, the probability that $x_n \in [a, b)$ is $b - a$.

Let y_n be the x_ns arranged in increasing order. How do the neighbor spacings behave?

First, we need to decide what is the correct scale to use for our investigations. As we have N objects on the wrapped unit interval, we have N nearest neighbor spacings. Thus, we expect the average spacing to be $\frac{1}{N}$.

De£nition 14.4.1 (Unfolding). *Let $z_n = Ny_n$. The numbers $z_n = Ny_n$ have unit mean spacing. Thus, while we expect the average spacing between adjacent y_ns to be $\frac{1}{N}$ units, we expect the average spacing between adjacent z_ns to be 1 unit.*

So, the probability of observing a spacing as large as $\frac{1}{2}$ between adjacent y_ns becomes negligible as $N \to \infty$. What we should ask is what is the probability of observing a nearest neighbor spacing of adjacent y_ns that is *half* the average spacing. In terms of the z_ns, this will correspond to a spacing between adjacent z_ns of $\frac{1}{2}$ a unit.

14.4.1 Nearest Neighbor Spacings

By symmetry, on the wrapped unit interval the expected nearest neighbor spacing is independent of j. Explicitly, we expect $\beta_{j+1} - \beta_j$ to have the same distribution as $\beta_{i+1} - \beta_i$.

What is the probability that, when we order the x_ns in increasing order, the next x_n after x_1 is located between $\frac{t}{N}$ and $\frac{t+\Delta t}{N}$? Let the x_ns in increasing order be labeled $y_1 \leq y_2 \leq \cdots \leq y_N$, $y_n = x_{\sigma(n)}$.

As we are choosing the x_ns independently, there are $\binom{N-1}{1}$ choices of subscript n such that x_n is nearest to x_1. This can also be seen by symmetry, as each x_n is equally likely to be the £rst to the *right* of x_1 (where, of course, .001 is just a little to the right of .999), and we have $N-1$ choices left for x_n.

The probability that $x_n \in \left[\frac{t}{N}, \frac{t+\Delta t}{N}\right]$ is $\frac{\Delta t}{N}$.

For the remaining $N-2$ of the x_ns, each must be further than $\frac{t+\Delta t}{N}$ from x_n. Thus, they must *all* lie in an interval (or possibly two intervals if we wrap around) of length $1 - \frac{t+\Delta t}{N}$. The probability that they all lie in this region is $\left(1 - \frac{t+\Delta t}{N}\right)^{N-2}$.

Thus, if $x_1 = y_l$, we want to calculate the probability that $||y_{l+1} - y_l|| \in \left[\frac{t}{N}, \frac{t+\Delta t}{N}\right]$. This is

$$\begin{aligned}
\text{Prob}\left(||y_{l+1} - y_l|| \in \left[\frac{t}{N}, \frac{t+\Delta t}{N}\right]\right) &= \binom{N-1}{1} \cdot \frac{\Delta t}{N} \cdot \left(1 - \frac{t+\Delta t}{N}\right)^{N-2} \\
&= \left(1 - \frac{1}{N}\right) \cdot \left(1 - \frac{t+\Delta t}{N}\right)^{N-2} \Delta t.
\end{aligned} \tag{14.12}$$

For N enormous and Δt small,

$$\begin{aligned}
\left(1 - \frac{1}{N}\right) &\approx 1 \\
\left(1 - \frac{t+\Delta t}{N}\right)^{N-2} &\approx e^{-(t+\Delta t)} \approx e^{-t}.
\end{aligned} \tag{14.13}$$

Thus

$$\text{Prob}\left(||y_{l+1} - y_l|| \in \left[\frac{t}{N}, \frac{t+\Delta t}{N}\right]\right) \to e^{-t} \Delta t. \tag{14.14}$$

Remark 14.4.2. *The above argument is in£nitesimally wrong. Once we've located y_{l+1}, the remaining x_ns do not need to be more than $\frac{t+\Delta t}{N}$ units to the right of $x_1 = y_l$; they only need to be further to the right than y_{l+1}. As the incremental gain in probabilities for the locations of the remaining x_ns is of order Δt, these contributions will not in¤uence the large N, small Δt limits. Thus, we ignore these effects.*

To rigorously derive the limiting behavior of the nearest neighbor spacings using the above arguments, one would integrate over x_m ranging from $\frac{t}{N}$ to $\frac{t+\Delta t}{N}$, and the remaining events x_n would be in the a segment of length $1 - x_m$. As

$$\left|\Big(1 - x_m\Big) - \Big(1 - \frac{t + \Delta t}{N}\Big)\right| \ \le\ \frac{\Delta t}{N}, \tag{14.15}$$

this will lead to corrections of higher order in Δt, hence negligible.
We can rigorously avoid this by instead considering the following:

1. Calculate the probability that all the other x_ns are at least $\frac{t}{N}$ units to the right of x_1. This is

$$p_t \ =\ \Big(1 - \frac{t}{N}\Big)^{N-1} \ \to\ e^{-t}. \tag{14.16}$$

2. Calculate the probability that all the other x_ns are at least $\frac{t+\Delta t}{N}$ units to the right of x_1. This is

$$p_{t+\Delta t} \ =\ \Big(1 - \frac{t + \Delta t}{N}\Big)^{N-1} \ \to\ e^{-(t+\Delta t)}. \tag{14.17}$$

3. The probability that no x_ns are within $\frac{t}{N}$ units to the right of x_1 but at least one x_n is between $\frac{t}{N}$ and $\frac{t+\Delta t}{N}$ units to the right is $p_{t+\Delta t} - p_t$:

$$\begin{aligned} p_t - p_{t+\Delta t} &\ \to\ e^{-t} - e^{-(t+\Delta t)} \\ &\ =\ e^{-t}\Big(1 - e^{-\Delta t}\Big) \\ &\ =\ e^{-t}\left(1 - 1 + \Delta t + O\Big((\Delta t)^2\Big)\right) \\ &\ \to\ e^{-t}\Delta t. \end{aligned} \tag{14.18}$$

Definition 14.4.3 (Unfolding Spacings). *If $y_{l+1} - y_l \in \left[\frac{t}{N}, \frac{t+\Delta t}{N}\right]$, then $N(y_{l+1} - y_l) \in [t, t+\Delta t]$. The new spacings $z_{l+1} - z_l$ have unit mean spacing. Thus, while we expect the average spacing between adjacent y_ns to be $\frac{1}{N}$ units, we expect the average spacing between adjacent z_ns to be 1 unit.*

14.4.2 k^{th} Neighbor Spacings

Similarly, one can easily analyze the distribution of the k^{th} neighbor spacings when each x_n is chosen independently from the uniform distribution on $[0, 1)$.

Again, consider $x_1 = y_l$. Now we want to calculate the probability that y_{l+k} is between $\frac{t}{N}$ and $\frac{t+\Delta t}{N}$ units to the *right* of y_l.

Therefore, we need exactly $k-1$ of the x_ns to lie between 0 and $\frac{t}{N}$ units to the right of x_1, exactly one x_n (which will be y_{l+k}) to lie between $\frac{t}{N}$ and $\frac{t+\Delta t}{N}$ units to the right of x_1, and the remaining x_ns to lie at least $\frac{t+\Delta t}{N}$ units to the right of y_{l+k}.

Remark 14.4.4. *We face the same problem discussed in Remark 14.4.2; a similar argument will show that ignoring these affects will not alter the limiting behavior. Therefore, we will make these simplifications.*

There are $\binom{N-1}{k-1}$ ways to choose the x_ns that are at most $\frac{t}{N}$ units to the right of x_1; there is then $\binom{(N-1)-(k-1)}{1}$ ways to choose the x_n between $\frac{t}{N}$ and $\frac{t+\Delta t}{N}$ units to the right of x_1.

Thus,

$$\begin{aligned}
&\text{Prob}\left(||y_{l+k} - y_l|| \in \left[\frac{t}{N}, \frac{t+\Delta t}{N}\right]\right) \quad = \\
= \quad & \binom{N-1}{k-1}\left(\frac{t}{N}\right)^{k-1} \cdot \binom{(N-1)-(k-1)}{1}\frac{\Delta t}{N} \cdot \left(1 - \frac{t+\Delta t}{N}\right)^{N-(k+1)} \\
= \quad & \frac{(N-1)\cdots(N-1-(k-2))}{N^{k-1}} \frac{(N-1)-(k-1)}{N} \frac{t^{k-1}}{(k-1)!}\left(1 - \frac{t+\Delta t}{N}\right)^{N-(k+1)} \Delta t \\
\to \quad & \frac{t^{k-1}}{(k-1)!} e^{-t} \Delta t. \qquad (14.19)
\end{aligned}$$

Again, one way to avoid the complications is to integrate over x_m ranging from $\frac{t}{N}$ to $\frac{t+\Delta t}{N}$.

Or, similar to before, we can proceed more rigorously as follows:

1. Calculate the probability that exactly $k-1$ of the other x_ns are at most $\frac{t}{N}$ units to the right of x_1, and the remaining $(N-1)-(k-1)$ of the x_ns are at least $\frac{t}{N}$ units to the right of x_1. As there are $\binom{N-1}{k-1}$ ways to choose $k-1$ of the x_ns to be at most $\frac{t}{N}$ units to the right of x_1, this probability is

$$\begin{aligned} p_t &= \binom{N-1}{k-1}\left(\frac{t}{N}\right)^{k-1}\left(1-\frac{t}{N}\right)^{(N-1)-(k-1)} \\ &\to \frac{N^{k-1}}{(k-1)!}\frac{t^{k-1}}{N^{k-1}}e^{-t} \\ &\to \frac{t^{k-1}}{k-1!}e^{-t}. \end{aligned} \tag{14.20}$$

2. Calculate the probability that exactly $k-1$ of the other x_ns are at most $\frac{t}{N}$ units to the right of x_1, and the remaining $(N-1)-(k-1)$ of the x_ns are at least $\frac{t+\Delta t}{N}$ units to the right of x_1. Similar to the above, this gives

$$\begin{aligned} p_t &= \binom{N-1}{k-1}\left(\frac{t}{N}\right)^{k-1}\left(1-\frac{t+\Delta t}{N}\right)^{(N-1)-(k-1)} \\ &\to \frac{N^{k-1}}{(k-1)!}\frac{t^{k-1}}{N^{k-1}}e^{-(t+\Delta t)} \\ &\to \frac{t^{k-1}}{(k-1)!}e^{-(t+\Delta t)}. \end{aligned} \tag{14.21}$$

3. The probability that exactly $k-1$ of the x_ns are within $\frac{t}{N}$ units to the right of x_1 and at least one x_n is between $\frac{t}{N}$ and $\frac{t+\Delta t}{N}$ units to the right is $p_{t+\Delta t} - p_t$:

$$p_t - p_{t+\Delta t} \;\to\; \frac{t^{k-1}}{(k-1)!}e^{-t} - \frac{t^{k-1}}{(k-1)!}e^{-(t+\Delta t)} \;\to\; \frac{t^{k-1}}{(k-1)!}e^{-t}\Delta t. \tag{14.22}$$

Note that when $k=1$, we recover the nearest neighbor spacings.

14.5 Induced Probability Measures

We have proven the following:

Theorem 14.5.1. *Consider N independent random variables x_n chosen from the uniform distribution on the wrapped unit interval $[0,1)$. For £xed N, arrange the x_ns in increase order, labeled $y_1 \leq y_2 \leq \cdots \leq y_N$.*

Form the induced probability measure $\mu_{N,1}$ from the nearest neighbor spacings. Then as $N \to \infty$ we have

$$\mu_{N,1}(t)dt \ = \ \frac{1}{N}\sum_{n=1}^{N}\delta\Big(t - N(y_n - y_{n-1})\Big)dt \ \to \ e^{-t}dt. \tag{14.23}$$

Equivalently, using $z_n = Ny_n$:

$$\mu_{N,1}(t)dt \ = \ \frac{1}{N}\sum_{n=1}^{N}\delta\Big(t - (z_n - z_{n-1})\Big)dt \ \to \ e^{-t}dt. \tag{14.24}$$

More generally, form the probability measure from the k^{th} nearest neighbor spacings. Then as $N \to \infty$ we have

$$\mu_{N,k}(t)dt \ = \ \frac{1}{N}\sum_{n=1}^{N}\delta\Big(t - N(y_n - y_{n-k})\Big)dt \ \to \ \frac{t^{k-1}}{(k-1)!}e^{-t}dt. \tag{14.25}$$

Equivalently, using $z_n = Ny_n$:

$$\mu_{N,k}(t)dt \ = \ \frac{1}{N}\sum_{n=1}^{N}\delta\Big(t - (z_n - z_{n-k})\Big)dt \ \to \ \frac{t^{k-1}}{(k-1)!}e^{-t}dt. \tag{14.26}$$

De£nition 14.5.2 (Poissonian Behavior). *We say a sequence of points x_n has Poissonian Behavior if in the limit as $N \to \infty$ the induced probability measures $\mu_{N,k}(t)dt$ converge to $\frac{t^{k-1}}{(k-1)!}e^{-t}dt$.*

Exercise 14.5.3. *Let $\alpha \in \mathbb{Q}$, and de£ne $\alpha_n = \{n^m\alpha\}$ for some positive integer m. Show the sequence of points α_n does not have Poissonian Behavior.*

Exercise 14.5.4. *Let $\alpha \notin \mathbb{Q}$, and de£ne $\alpha_n = \{n\alpha\}$. Show the sequence of points α_n does not have Poissonian Behavior. Hint: for each N, show the nearest neighbor spacings take on at most three distinct values (the three values depend on N). As only three values are ever assumed for a £xed N, $\mu_{N,1}(t)dt$ cannot converge to $e^{-t}dt$.*

14.6 Non-Poissonian Behavior

Conjecture 14.6.1. *With probability one (with respect to Lebesgue Measure), if $\alpha \notin \mathbb{Q}$, if $\alpha_n = \{n^2\alpha\}$ then the sequence of points α_n is Poissonian.*

There are constructions which show certain irrationals give rise to non-Poissonian behavior.

Theorem 14.6.2. *Let $\alpha \in \mathbb{Q}$ such that $\left|\alpha - \frac{p_n}{q_n}\right| < \frac{a_n}{q_n^3}$ holds in£nitely often, with $a_n \to 0$. Then there exist integers $N_j \to \infty$ such that $\mu_{N_j,1}(t)$ does not converge to $e^{-t}dt$.*

As $a_n \to 0$, eventually $a_n < \frac{1}{10}$ for all n large. Let $N_n = q_n$, where $\frac{p_n}{q_n}$ is a good rational approximation to α:

$$\left|\alpha - \frac{p_n}{q_n}\right| < \frac{a_n}{q_n^3}. \tag{14.27}$$

Remember that all subtractions are performed on the wrapped unit interval. Thus, $||.999 - .001|| = .002$.

We look at $\alpha_k = \{k^2\alpha\}$, $1 \leq k \leq N_n = q_n$. Let the β_ks be the α_ks arranged in increasing order, and let the γ_ks be the numbers $\{k^2\frac{p_n}{q_n}\}$ arranged in increasing order:

$$\begin{aligned} \beta_1 &\leq \beta_2 \leq \cdots \leq \beta_N \\ \gamma_1 &\leq \gamma_2 \leq \cdots \leq \gamma_N. \end{aligned} \tag{14.28}$$

14.6.1 Preliminaries

Lemma 14.6.3. *If $\beta_l = \alpha_k = \{k^2\alpha\}$, then $\gamma_l = \{k^2\frac{p_n}{q_n}\}$. Thus, the same permutation orders both the α_ks and the γ_ks.*

Proof. Multiplying both sides of Equation 14.27 by $k^2 \leq q_n^2$ yields

$$\left|k^2\alpha - k^2\frac{p_n}{q_n}\right| \ < \ k^2\frac{a_n}{q_n^2} \ \leq \ \frac{a_n}{q_n} \ < \ \frac{1}{2q_n}. \tag{14.29}$$

Thus, $k^2\alpha$ and $k^2\frac{p_n}{q_n}$ differ by at most $\frac{1}{2q_n}$. Therefore

$$\left|\left|\left\{k^2\alpha\right\} - \left\{k^2\frac{p_n}{q_n}\right\}\right|\right| \ < \ \frac{1}{2q_n}. \tag{14.30}$$

As the numbers $\{m^2\frac{p_n}{q_n}\}$ all have denominators of size at most $\frac{1}{q_n}$, we see that $\{k^2\frac{p_n}{q_n}\}$ is the closest of the $\{m^2\frac{p_n}{q_n}\}$ to $\{k^2\alpha\}$.

This implies that if $\beta_l = \{k^2\alpha\}$, then $\gamma_l = \{k^2\frac{p_n}{q_n}\}$, completing the proof. □

Exercise 14.6.4. *Prove the ordering is as claimed. Hint: about each $\beta_l = \{k^2\alpha\}$, the closest number of the form $\{c^2\frac{p_n}{q_n}\}$ is $\{k^2\frac{p_n}{q_n}\}$.*

14.6.2 Proof of Theorem 14.6.2

Exercise 14.6.5. *Assume $||a-b||, ||c-d|| < \frac{1}{10}$. Show*

$$||(a-b)-(c-d)|| \ < \ ||a-b|| + ||c-d||. \tag{14.31}$$

Proof of Theorem 14.6.2: We have shown

$$||\beta_l - \gamma_l|| \ < \ \frac{a_n}{q_n}. \tag{14.32}$$

Thus, as $N_n = q_n$:

$$\left|\left|N_n(\beta_l - \gamma_l)\right|\right| \ < \ a_n, \tag{14.33}$$

and the same result holds with l replaced by $l-1$.

By Exercise 14.6.5,

$$\left|\left|N_n(\beta_l - \gamma_l) - N_n(\beta_{l-1} - \gamma_{l-1})\right|\right| \ < \ 2a_n. \tag{14.34}$$

Rearranging gives

$$\left|\left|N_n(\beta_l - \beta_{l-1}) - N_n(\gamma_l - \gamma_{l-1})\right|\right| \ < \ 2a_n. \tag{14.35}$$

As $a_n \to 0$, this implies the difference between $\left|\left|N_n(\beta_l - \beta_{l-1})\right|\right|$ and $\left|\left|N_n(\gamma_l - \gamma_{l-1})\right|\right|$ goes to zero.

The above distance calculations were done mod 1. The actual differences will differ by an integer. Thus,

$$\mu^{\alpha}_{N_n,1}(t)dt \;=\; \frac{1}{N_n}\sum_{l=1}^{N_n}\delta\Big(t - N_n(\beta_l - \beta_{l-1})\Big) \tag{14.36}$$

and

$$\mu^{\frac{p_n}{q_n}}_{N_n,1}(t)dt \;=\; \frac{1}{N_n}\sum_{l=1}^{N_n}\delta\Big(t - N_n(\gamma_l - \gamma_{l-1})\Big) \tag{14.37}$$

are extremely close to one another; each point mass from the difference between adjacent β_ls is either within a_n units of a point mass from the difference between adjacent γ_ls, or is within a_n units of a point mass an integer number of units from a point mass from the difference between adjacent γ_ls. Further, $a_n \to 0$.

Note, however, that if $\gamma_l = \{k^2\frac{p_n}{q_n}\}$, then

$$N_n\gamma_l \;=\; q_n\Big\{k^2\frac{p_n}{q_n}\Big\} \;\in\; \mathbb{N}. \tag{14.38}$$

Thus, the induced probability measure $\mu^{\frac{p_n}{q_n}}_{N_n,1}(t)dt$ formed from the γ_ls is supported on the integers! Thus, it is impossible for $\mu^{\frac{p_n}{q_n}}_{N_n,1}(t)dt$ to converge to $e^{-t}dt$.

As $\mu^{\alpha}_{N_n,1}(t)dt$, modulo some possible integer shifts, is arbitrarily close to $\mu^{\frac{p_n}{q_n}}_{N_n,1}(t)dt$, the sequence $\{k^2\alpha\}$ is *not* Poissonian along the subsequence of Ns given by N_n, where $N_n = q_n$, q_n is a denominator in a good rational approximation to α. $\Box$

14.6.3 Measure of $\alpha \notin \mathbb{Q}$ with Non-Poissonian Behavior along a sequence N_n

What is the (Lebesgue) measure of $\alpha \notin \mathbb{Q}$ such that there are in£nitely many n with

$$\left|\alpha - \frac{p_n}{q_n}\right| \;<\; \frac{a_n}{q_n},\; a_n \to 0. \tag{14.39}$$

If the above holds, then for any constant $k(\alpha)$, for n large (large depends on both α and $k(\alpha)$) we have

$$\left|\alpha - \frac{p_n}{q_n}\right| \; < \; \frac{k(\alpha)}{q_n^{2+\epsilon}}. \tag{14.40}$$

Exercise 14.6.6. *Show this set has (Lebesgue) measure or size* 0.

Thus, almost no irrational numbers satisfy the conditions of Theorem 14.6.2, where *almost no* is relative to the (Lebesgue) measure.

Exercise 14.6.7. *In a topological sense, how many algebraic numbers satisfy the conditions of Theorem 14.6.2? How many transcendental numbers satisfy the conditions?*

Exercise 14.6.8. *Let α satisfy the conditions of Theorem 14.6.2. Consider the sequence N_n, where $N_n = q_n$, q_n the denominator of a good approximation to α. We know the induced probability measures $\mu_{N_n,1}^{\frac{p_n}{q_n}}(t)dt$ and $\mu_{N_n,1}^{\alpha}(t)dt$ do not converge to $e^{-t}dt$. Do these measures converge to anything?*

Remark 14.6.9. *In* The Distribution of Spacings Between the Fractional Parts of $\{n^2\alpha\}$ *(Z. Rudnick, P. Sarnak, A. Zaharescu), it is shown that for most α satisfying the conditions of Theorem 14.6.2, there* is *a sequence N_j along which $\mu_{N_n,1}^{\alpha}(t)dt$* does *converge to $e^{-t}dt$.*

Chapter 15

More Graphs, Kloosterman, Randomness of $x \to \overline{x}$ mod p

More on Graphs, Kloosterman, and the Third Problem on the Randomness of $x \mapsto \overline{x}$ mod p. Review of projective geometry and fractional linear transformations. Lecture by Peter Sarnak; notes by Steven J. Miller.

15.1 Kloosterman Sums

Recall

$$S(a,b,p) \;=\; \sum_{x \bmod p}^{*} e\Big(\frac{ax+b\overline{x}}{p}\Big), \tag{15.1}$$

where $\sum^*$ means sum over all x relatively prime to p, and $\overline{x} \equiv x^{-1}$ mod p.

Theorem 15.1.1 (Weil). *For p an odd prime and $a, b \in \mathbb{Z}$,*

$$|S(a,b,p)| \;\leq\; 2\sqrt{p}. \tag{15.2}$$

The above captures the randomness. We add $p-1$ numbers of modulus 1, and we see square-root cancellation. Weil's Theorem says the cancellation is "like" random numbers. Recall when we added ± 1, we expected to observe a sum around $\sqrt{N}$ if we had N summands.

15.2 Projective Geometry

Consider the map

$$\begin{aligned} \mathbb{C} &\rightarrow \mathbb{C} \cup \{\infty\} \\ z &\mapsto \frac{az+b}{cz+d}. \end{aligned} \tag{15.3}$$

We de£ne the above as the action of the matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{15.4}$$

on z.

$\mathbb{P}^1(\mathbb{R})$ is identi£ed with the perimeter of a circle, with antipodal points (points on a diagonal, separated by π radians) identi£ed.

15.3 Example

De£nition 15.3.1 **($\mathbb{P}^1(\mathbb{F}_p)$).** $\mathbb{P}^1(\mathbb{F}_p)$ *is the projective line,*

$$\mathbb{P}^1(\mathbb{F}_p) = \{0, 1, \ldots, \infty\}. \tag{15.5}$$

We construct a 3-regular graph G_p on $p + 1$ vertices as follows:

1. Join x to $x + 1$ mod p.
2. Join x to $x - 1$ mod p.
3. Join x to $-\overline{x}$ mod p,

where $\frac{1}{\infty} = 0$.

Form the adjacency matrix of the above graph. Is there a spectral graph?

Theorem 15.3.2. *There is a spectral gap!*

$$\lambda_1(G_p) \leq 2.99. \tag{15.6}$$

These graphs are not *Ramanujan in general.*

We can look at the three maps as matrices

$$T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad S = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad T^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \tag{15.7}$$

Exercise 15.3.3. $(ST)^3 = \pm I$.

Exercise 15.3.4. *Show the three maps we used to create G_p can be given by*

1. $x \to Tx$ *(corresponding to* $x \to x + 1$*),*
2. $x \to T^{-1}x$ *(corresponding to* $x \to x - 1$*),*
3. $x \to Sx$ *(corresponding to* $x \to -\overline{x}$*.*

15.4 Stereographic Projections and Fractional Linear Transformations

What are the analytic, $1-1$ invertible maps from $\mathbb{C} \to \mathbb{C}$? What if we include ∞.

First, one might ask what is ∞?

Take a sphere, call the north pole N. Consider the in£nite plane $z = 0$.

To each point P on the sphere, draw the line from N to P, and write down the point of intersection on the plane $z = 0$. Call this map S (stereographic projection; preserves angles), and call the sphere $S^2 = \mathbb{P}^1$.

Thus,

$$S(P) \in \mathbb{C}, \ S(\infty) \leftrightarrow N, \ S^2 \cong \mathbb{C} \cup \{\infty\}. \tag{15.8}$$

Recall GL_2 are the 2×2 matrices with non-zero determinant. $SL_2(\mathbb{C})$ is the group of 2×2 matrices with determinant 1.

Fractional Linear Transformation: $z \mapsto \frac{az+b}{cz+d}$.

If we have two linear transformations

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \ \delta = \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} \tag{15.9}$$

then

$$\gamma(\delta z) \ = \ (\gamma\delta)z, \tag{15.10}$$

where $(\gamma\delta)$ is usual matrix multiplication.

15.5 More Kesten

Let Γ be a group generated by

$$A_1,\ A_1^{-1},\ A_2,\ A_2^{-1},\ \ldots\ ,\ A_k,\ A_k^{-1}. \tag{15.11}$$

We make a $2k$-regular graph by joining x and y with an edge ($x \sim y$) if and only if $x = A_j^{\pm 1} y$ for some j.

De£ne, for $f : \Gamma \to \mathbb{C}$,

$$Bf(x) \;=\; \sum_{x \sim y} f(y). \tag{15.12}$$

To make the sum make sense if we have in£nitely many vertices we require $f \in l^2(\Gamma)$, the space of functions g where $\sum |g(\gamma)|^2 < \infty$.

One could ask about the spectrum of B on this space. The HW problem was on a k-regular tree. Make words using the generators. Have the notion of a free group: no relation (ie, no word is the identity word) except the trivial ones ($A_i A_i^{-1} = I$).

$$\Gamma \;=\; \langle A_1, A_2, \ldots, A_k \rangle. \tag{15.13}$$

De£nition 15.5.1 (Free Group). *Γ is a free group if the only relations in Γ giving the identity are the obvious ones (ie, the only word in $A_1, \ldots, A_k$ that is the identity word is words of the form $A_k^{-1} A_i A_i^{-1} A_j^{-1} A_j A_k$, and so on).*

The graph we just spoke about, $G(\Gamma)$ on A_1, A_1^{-1} up to A_k, A_k^{-1} is a $2k$-regular tree if and only if Γ is a free group on $A_1, \ldots, A_k$. This is called a Cayley Graph (relative to the given generators).

Theorem 15.5.2 (Kesten). *The spectrum of B when Γ is free on k generators is*

$$\textit{spectrum}(B) \;=\; \Big[-2\sqrt{2k-1}, 2\sqrt{2k-1} \Big]. \tag{15.14}$$

Further, Γ is free on $A_1, \ldots, A_k$ if and only if

$$\textit{spectrum}(B) \;\subset\; \Big[-2\sqrt{2k-1}, 2\sqrt{2k-1} \Big]. \tag{15.15}$$

Thus, the spectrum contains a point outside *this interval if and only if Γ is not free.*

Finally, $2k$ is in the spectrum if and only if Γ is amenable (for example, abelian).

Corollary 15.5.3. *Our graphs G_p cannot be Ramanujan, as there* is *a relation among the generators, namely* $(ST)^3 = \pm I$.

15.6 $\mathbf{SL}_2(\mathbb{Z})$

De£nition 15.6.1 ($\mathbf{SL}_2(\mathbb{Z})$). *$SL_2(\mathbb{Z})$ is the group of 2×2 matrices with unit determinant and integer coef£cients.*

Exercise 15.6.2. *Consider the matrices*

$$T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad S = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad T^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \tag{15.16}$$

Show these three matrices generate $SL_2(\mathbb{Z})$; as $(ST)^3 = \pm I$, this shows $SL_2(\mathbb{Z})$ is not a free group.

Where should $\mathrm{SL}_2(\mathbb{Z})$ act? Lubotsky: you don't understand a group until it acts on something you know. Galois had groups acting on roots of polynomials (permuting roots).

What does $\mathrm{SL}_2(\mathbb{Z})$ act on? It does act on the sphere, but it is too big a space. We want to study the smallest space where it acts reasonably.

Look at $\mathrm{SL}_2(\mathbb{R})$, the group of 2×2 matrices with determinant 1 and real entries. Let z be in the upper half plane, so $z = x + iy$, $y > 0$. Then

$$z \;\mapsto\; \frac{az+b}{cz+d}, \ \mathrm{Im}\left(\frac{az+b}{cz+d}\right) \;>\; 0. \tag{15.17}$$

A similar statement holds for z in the lower half plane. Thus, $\mathrm{SL}_2(\mathbb{R})$ maps the upper (lower) half plane to itself.

By shifting by an integer, we can bring any $x \in \mathbb{R}$ to $x' \in [0, 1)$.

Gauss was the £rst to draw the fundamental domain for $\mathrm{SL}_2(\mathbb{Z})$ acting on the upper half plane:

Draw a circle of radius 1 with center 0; draw vertical lines at $x = \pm\frac{1}{2}$, going from the point on the circle to in£nity. The region formed is called the fundamental domain for $\mathrm{SL}_2(\mathbb{Z})$ on the upper half plane. This means that any z in the upper half plane can be brought into this region.

15.7 Is $x \to \overline{x}$ mod p Random?

15.7.1 First Test

Question 15.7.1. *To what extent is / how random is $x \to \overline{x}$ mod p?*

There are p numbers.

If look at all numbers $1 \le x \le p-1$, we get all the numbers back (in some new order). We can look at a long segment (on what scale?), say

$$\sqrt{p} \ \le x \le 2\sqrt{p}. \tag{15.18}$$

Now $\overline{x}$ will be all over $[1, p-1]$. This generates $\sqrt{p}$ numbers between 1 and $p-1$. The average spacing is is approximately $\frac{p-1}{\sqrt{p} \approx \sqrt{p}}$. How are numbers spaced?

First, we need to write them in increasing order

$$1 \le a_1 \ < \ a_2 \ < \ \cdots \ < a_l \ \le p-1. \tag{15.19}$$

Let us denote the nearest neighbor spacings by $\Delta_1 = a_2 - a_1$, $\Delta_2 = a_3 - a_2$ and so on. Let

$$\delta_j = \frac{\Delta_j}{\sqrt{p}}. \tag{15.20}$$

Note the δ_js have unit mean spacing.

Conjecture 15.7.2 (Naive Conjecture). *We expect the spacings to follow Poissonian Statistics. Explicitly, the distribution of spacings we see here should be the same as that from choosing $\sqrt{p}$ numbers independently from the uniform distribution on $[0, 1)$.*

15.7.2 Second Test

Question 15.7.3 (Jim Propp). *Does $x \to \overline{x}$ behave like a random permutation of order 2?*

What is the distribution of the longest increasing sub-sequence?

Given a permutation of $1, 2, \ldots, N$, we have $i \mapsto \sigma(i)$. Permutations are often denoted by

$$\begin{pmatrix} 1 & 2 & \cdots & N \\ \sigma(1) & \sigma(2) & & \sigma(N) \end{pmatrix} \tag{15.21}$$

Look at the distribution of the longest increasing sub-sequence about the mean. Normalized appropriately, what does it look like?

15.7.3 Third Test: Hooley's R^*

$$\sum_{x=A}^{A+N} e\left(\frac{\overline{x}}{p}\right), \tag{15.22}$$

where $1 \leq A \leq A+N \leq p-1$.
When we add numbers of modulus one, we expect square-root cancellation.
Then

Conjecture 15.7.4 (Hooley's R^*). *For every $\epsilon > 0$,*

$$\left|\sum_{x=A}^{A+N} e\left(\frac{\overline{x}}{p}\right)\right| \ll_\epsilon N^{\frac{1}{2}} p^\epsilon. \tag{15.23}$$

Note $\ll_\epsilon$ means the left hand side is less than c_ϵ times the right hand side (for some c_ϵ > 0).

Note there is no dependence on A – the only dependence is on the size of summation N and the prime p.

If $N > \sqrt{p}$, the above can be proven. by Weil's bound.

15.8 Note on Non-trivial Bound of Fourth Powers of Kloosterman Sums

Note on conditions arising in non-trivial bound on sum of fourth powers of Kloosterman sums (Heath-Brown review). Supplemental notes by Alex Barnett.

Please refer to Professor Sarnak's lecture of 10/16/02, and Heath-Brown's review article on Kloosterman Sums.

There are six summations inherent in the desired sum

$$\sum_{a=0}^{p-1}\sum_{b=0}^{p-1} \left|\mathrm{Kl}(a,b,p)\right|^4, \tag{15.24}$$

namely the two sums shown and one internal sum in each of the Kl's. Using the result

$$\sum_{a=0}^{p-1} e\left(\frac{an}{p}\right) = \begin{cases} p-1, & n=0 \ (\mathrm{mod} p) \\ 0, & \text{otherwise} \end{cases} \tag{15.25}$$

twice, turns the two $e(\cdot)$ functions into counting conditions on the variables internal to the four Kl's. Calling these variables w, x, y, z, we want the total number of ways that, with $1 \leq w, x, y, z \leq p-1$ (the four surviving sums), we satisfy both

$$w + x - y - z \ = 0 \ \ (p) \tag{15.26}$$

and

$$\bar{w} + \bar{x} - \bar{y} - \bar{z} \ = 0 \ \ (p). \tag{15.27}$$

Multiplying Eq. 15.27 by $wxyz$ gives

$$(w+x)yz - (y+z)wx \ = \ 0 \ \ (p). \tag{15.28}$$

Substituting Eq. 15.26 gives

$$(w+x)(yz - wx) \ = \ 0 \ \ (p). \tag{15.29}$$

So either $w + x = 0 \ \ (p)$ or $yz - wx = 0 \ \ (p)$, or both.

The £rst set of cases has $x = -w$ from (15.26) giving $z = -y$, so there are $(p-1)^2$ choices of w and y. For each choice x and z are £xed uniquely. Therefore these cases contribute $(p-1)^2$ ways.

For the second set, we have two equations

$$y + z \ = \ w + x \ \ (p) \tag{15.30}$$

$$yz \ = \ wx \ \ (p) \tag{15.31}$$

for two unknowns y, z, for any of the arbitrary choices of w, x. You could combine these equations into the single quadratic

$$y^2 - y(w+x) + wx \ = 0 \ \ (p). \tag{15.32}$$

Two solutions for y are $y = w$ and $y = x$ (check by substitution). Since it is a quadratic, these are the only two solutions. Therefore the number of ways contributed is at most 2 times the $(p-1)^2$ ways of choosing w, x.

Over-counting due to the and/or is at least of order p or smaller, but also can only reduce the number of ways. Therefore the total number of ways $\leq 3(p-1)^2$, which is $O(p^2)$. From this follows the bound on the sum given in the article and lecture.

Chapter 16

Introduction to the Hardy-Littlewood Circle Method

Introduction to the Hardy-Littlewood Circle Method. Lecture and notes by Steven J. Miller.

16.1 Problems where the Circle Method is Useful

For each N, let A_N be a set of non-negative integers such that

1. $A_N \subset A_{N+1}$,
2. $|A_N| \to \infty$ as $N \to \infty$.

Let $A = \lim_{N\to\infty} A_N$.

Question 16.1.1. *Let s be a £xed positive integer. What can one say about $a_1 + \cdots + a_s$? Ie, what numbers n are representable as a sum of s summands from A?*

We consider three problems; we will mention later why we are considering sets A_N.

16.1.1 Waring's Problem

Let A be the set of k^{th} powers of non-negative numbers, and let

$$A_N = \{0^k, 1^k, \dots, N^k\}. \tag{16.1}$$

Question 16.1.2. *Fix a positive integer k. For what positive integers s can every integer be written as a sum of s numbers, each number a k^{th} power?*

Thus, in this case, we are trying to solve

$$n \;=\; a_1^k + \cdots + a_N^k. \tag{16.2}$$

16.1.2 Goldbach's Problem

Let A be the set of all prime numbers, and let A_N be the set of all primes at most N.

Question 16.1.3. *Can every even number be written as the sum of two primes?*

In this example, we are trying to solve

$$2n \;=\; a_1 + a_2, \tag{16.3}$$

or, in more suggestive notation,

$$2n \;=\; p_1 + p_2. \tag{16.4}$$

16.1.3 Sum of Three Primes

Again, let A be the set of all primes, and A_N all primes up to N.

Question 16.1.4. *Can every odd number be written as the sum of three primes?*

Again, we are studying

$$2n + 1 \;=\; p_1 + p_2 + p_3. \tag{16.5}$$

16.2 Idea of the Circle Method

16.2.1 Introduction

De£nition 16.2.1 ($e(z)$). *We de£ne $e(z) = e^{2\pi i z}$.*

Exercise 16.2.2. *Let $m, n \in \mathbb{Z}$. Prove*

$$\int_0^1 e(nx)e(-mx)dx = \begin{cases} 1 & \text{if } n = m \\ 0 & \text{otherwise} \end{cases} \tag{16.6}$$

Let A, A_N be as in any of the three problems above. Consider

$$f_N(x) \;=\; \sum_{a \in A_N} e(ax). \tag{16.7}$$

We investigate $\Big(f_N(x)\Big)^s$:

$$\begin{aligned} \Big(f_N(x)\Big)^s \;&=\; \prod_{j=1}^{s} \sum_{a_j \in A_N} e(a_j x) \\ &=\; \sum_m r_N(m) e(mx). \end{aligned} \tag{16.8}$$

The last result follows by collecting terms. When you multiply two exponentials, you add the exponents.

Thus, when we multiply the s products, how can we get a product which gives $e(mx)$?

We have s products, say $e(a_1 x)$ through $e(a_N x)$. Thus,

$$e(a_1 x) \cdots e(a_N x) \;=\; e\Big((a_1 + \cdots + a_N)x\Big) \;=\; e(mx). \tag{16.9}$$

Thus, the coef£cient $r_N(m)$ in $\Big(f_N(x)\Big)^s$ is the number of ways of writing

$$m \;=\; a_1 + \cdots + a_N, \tag{16.10}$$

with each $a_j \in A_N$.

As the elements of A_N are non-negative, if N is suf£ciently large $r_N(m)$ is equal to the number of ways of writing m as the sum of s elements of A.

The problem is, if m is larger than the largest term in A_N, then there may be other ways to write m as a sum of s elements of A.

Lemma 16.2.3.

$$r_N(m) \;=\; \int_0^1 \Big(f_N(x)\Big)^s e(-mx)dx. \tag{16.11}$$

Proof: direct calculation.

Note that, just because we have a closed form expression for $r_N(m)$, this does not mean we can actually *evaluate* the above integral. Recall, for example, the inclusion - exclusion formula for the number of primes at most N. This is an exact formula, but very hard to evaluate.

16.2.2 Useful Number Theory Results

We will use the following statements freely:

Theorem 16.2.4 (Prime Number Theorem). *Let $\pi(x)$ denote the number of primes at most x. Then*

$$\pi(x) \;=\; \sum_{p\le x} 1 \;=\; \frac{x}{\log x} \;+\; \textit{smaller}. \tag{16.12}$$

Upon applying Partial Summation, we may rewrite the above as

$$\sum_{p\le x} \log p \;=\; x \;+\; \textit{smaller}. \tag{16.13}$$

Theorem 16.2.5 (Siegel-Wal£sz). *Let $C, B > 0$, and let a and q be relatively prime. Then*

$$\sum_{\substack{p\le x\\ p\equiv a(q)}} \log p \;=\; \frac{x}{\phi(q)} + O\Big(\frac{x}{\log^C x}\Big) \tag{16.14}$$

for $q \le \log^B x$, and the constant above does not depend on x, q or a (ie, it only depends on C and B).

For completeness, we include a review of partial summation as an appendix to these notes.

16.2.3 Average Sizes of $\Big(f_N(x)\Big)^s$

Henceforth we will consider $f_N(x)$ arising from the three prime case. Thus, $s = 3$.

For analytic reasons, it is more convenient to instead analyze the function

$$F_N(x) \;=\; \sum_{p\le N} \log p \cdot e(px). \tag{16.15}$$

Working analogously as before, we are led to

$$R_N(m) \;=\; \int_0^1 \Big(F_N(x)\Big)^3 e(-mx)dx. \tag{16.16}$$

By partial summation, it is very easy to go from $R_N(m)$ to $r_N(m)$.

Exercise 16.2.6. *Prove the trivial bound for $|F_N(x)|$ is N. Take absolute values and use the Prime Number Theorem.*

We can, however, show that the average square of $F_N(x)$ is signi£cantly smaller:

Lemma 16.2.7. *The average value of $|F_N(x)|^2$ is $N \log N$.*

Proof: The following trivial observation will be extremely useful in our arguments. Let $g(x)$ be a complex-valued function, and let $\overline{g}(x)$ be its complex conjugate. Then $|g(x)|^2 = g(x)\overline{g}(x)$.

In our case, as $\overline{F}_N(x) = F_N(-x)$ we have

$$\begin{aligned}
\int_0^1 |F_N(x)|^2 \;&=\; \int_0^1 F_N(x)F_N(-x)dx \\
&=\; \int_0^1 \sum_{p\le N} \log p \cdot e(px) \sum_{q \le N} \log q \cdot e(-qx)dx \\
&=\; \sum_{p\le N}\sum_{q\le N} \log p \log q \int_0^1 e\Big((p-q)x\Big)dx \\
&=\; \sum_{p\le N} \log^2 p.
\end{aligned} \tag{16.17}$$

Using $\sum_{p\le N} \log p = N \;+\;$ small and Partial Summation, we can show

$$\sum_{p\le N} \log^p \;=\; N\log N \;+\; \text{smaller}. \tag{16.18}$$

Thus,

$$\int_0^1 |F_N(x)|^2 \;=\; N\log N \;+\; \text{smaller}. \tag{16.19}$$

Thus, taking square-roots, we see on average $|\Big(F_N(x)\Big)|^2$ is $N\log N$, signi£cantly smaller than the maximum possible value (N^2). Thus, we see we are getting almost square-root cancellation on average. □

16.2.4 De£nition of the Major and Minor Arcs

We split the unit interval $[0, 1)$ into two disjoint parts, the Major and the Minor arcs.

Roughly, the Major arcs will be a union of very small intervals centered at rationals with small denominator (relative to N). Near these rationals, we will be able to approximate $F_N(x)$ very well, and $F_N(x)$ will be of size N.

The minor arcs will be the rest of $[0, 1)$; we will show that $F_N(x)$ is signi£cantly smaller than N here.

Major Arcs

Let $B > 0$, and let $Q = (\log N)^B \ll N$.

For each $q \in \{1, 2, \dots, Q\}$ and $a \in \{1, 2, \dots, q\}$ with a and q relatively prime, consider the set

$$\mathcal{M}_{a,q} \;=\; \left\{x \in [0,1) : \;\left|x - \frac{a}{q}\right| \;<\; \frac{Q}{N}\right\}. \tag{16.20}$$

We also add in one interval centered at either 0 or 1, ie, the "interval" (or wrapped-around interval)

$$\left[0, \frac{Q}{N}\right] \;\cup\; \left[1 - \frac{Q}{N}, 1\right]. \tag{16.21}$$

Exercise 16.2.8. *Show, if N is large, that the major arcs $\mathcal{M}_{a,q}$ are disjoint for $q \le Q$ and $a \le q$, a and q relatively prime.*

We de£ne the Major Arcs to be the union of each arc $\mathcal{M}_{a,q}$:

$$\mathcal{M} \;=\; \bigcup_{q=1}^{Q} \bigcup_{\substack{a=1 \\ (a,q)=1}} \mathcal{M}_{a,q}, \tag{16.22}$$

where (a, q) is the greatest common divisor of a and q.

Exercise 16.2.9. *Show $|\mathcal{M}| < \frac{2Q^3}{N}$. As $Q = \log^B N$, this implies as $N \to \infty$, the major arcs are zero percent of the unit interval.*

Minor Arcs

The Minor Arcs, m, are whatever is *not* in the Major Arcs. Thus,

$$\mathrm{m} \;=\; [0,1) - \mathcal{M}. \tag{16.23}$$

Clearly, as $N \to \infty$, almost all of $[0,1)$ is in the Minor Arcs.

16.3 Contributions from the Major and Minor Arcs

16.3.1 Contribution from the Minor Arcs

We bound the contribution from the minor arcs to $r_N(m)$:

$$\begin{aligned}
\Big|\int_{\mathrm{m}} F_N^3(x)e(-mx)dx\Big| &\leq \int_{\mathrm{m}} |F_N(x)|^3 dx \\
&\leq \Big(\max_{x\in\mathrm{m}} |F_N(x)|\Big) \int_{\mathrm{m}} |F_N(x)|^2 dx \\
&\leq \Big(\max_{x\in\mathrm{m}} |F_N(x)|\Big) \int_0^1 F_N(x)F_N(-x)dx \\
&\leq \Big(\max_{x\in\mathrm{m}} |F_N(x)|\Big) N\log N. \qquad (16.24)
\end{aligned}$$

As the minor arcs are most of the unit interval, replacing $\int_{\mathrm{m}}$ with $\int_0^1$ doesn't introduce much of an over-estimation.

In order for the Circle Method to succeed, we need a non-trivial, good bound for

$$\max_{x\in\mathrm{m}} |F_N(x)| \tag{16.25}$$

This is where most of the difficulty arises, showing that there is good cancellation in $F_N(x)$ if we stay away from rationals with small denominator.

We will show that the contribution to the major arcs is

$$\mathfrak{S}(N)\frac{N^2}{2} \;+\; \text{smaller}, \tag{16.26}$$

where $\exists c_1, c_2 > 0$ such that, for all N, $c_1 < \mathfrak{S}(N) < c_2$.

Thus, we need the estimate that

$$\max_{x\in \mathfrak{m}} |F_N(x)| \ \leq \ \frac{N}{\log^{1+\epsilon} N}. \tag{16.27}$$

Relative to the average size of $|F_N(x)|^2$, this is signi£cantly smaller; however, as we are showing that the maximum value of $|F_N(x)|$ is bounded, this is a significantly more delicate question. We know such a bound cannot be true for all $x \in [0,1)$ (see below, and not that $F_N(0) = N$). The hope is that if x is not near a rational with small denominator, we will get moderate cancellation.

While this is very reasonable to expect, it is not easy to prove.

16.3.2 Contribution from the Major Arcs

Fix a $q \leq Q$ and an $a \leq q$ with a and q relatively prime. We evaluate $F\left(\frac{a}{q}\right)$.

$$\begin{aligned} F\left(\frac{a}{q}\right) &= \sum_{p\leq N} \log p \cdot e^{2\pi i p \frac{a}{q}} \\ &= \sum_{r=1}^{q} \sum_{\substack{p\equiv r(q)\\ p\leq N}} \log p \cdot e^{2\pi i \frac{ap}{q}} \\ &= \sum_{r=1}^{q} \sum_{\substack{p\equiv r(q)\\ p\leq N}} \log p \cdot e^{2\pi i \frac{ar}{q}} \\ &= \sum_{r=1}^{q} e^{2\pi i \frac{ar}{q}} \sum_{\substack{p\equiv r(q)\\ p\leq N}} \log p \end{aligned} \tag{16.28}$$

Note the beauty of the above. The dependence on p in the original sums is very weak – there is a $\log p$ factor, and there is $e\left(\frac{ap}{q}\right)$. In the exponential, we only need to know p mod q. Now, p runs from 2 to N, and q is at most $\log^B N$. Thus, in general $p \gg q$.

We use the Siegel-Wal£sz Theorem. We £rst remark that we may assume r and q are relatively prime. Why? If $p \equiv r \bmod q$, this means $p = \alpha q + r$ for some $\alpha \in \mathbb{N}$. If r and q have a common factor, there can be at most one prime p (namely r) such that $p \equiv r \bmod q$, and this can easily be shown to give a negligible contribution.

For any $C > 0$

$$\sum_{\substack{p\equiv r(q)\\ p\leq N}} \log p \;=\; \frac{N}{\phi(q)} \;+\; O\Big(\frac{N}{\log^C N}\Big). \tag{16.29}$$

Now, as $\phi(q)$ is at most q which is at most $\log^B N$, we see that the main term is signi£cantly greater than the error term (choose C much greater than B).

Note the Siegel-Wal£sz Theorem would be useless if $q \approx N^\epsilon$. Then the main term would be like $N^{1-\epsilon}$, which would be smaller than the error term.

This is one reason why, in constructing the major arcs, we take the denominators to be small.

Thus, we £nd

$$\begin{aligned} F\Big(\frac{a}{q}\Big) &= \sum_{\substack{r=1\\ (r,q)=1}}^{q} e^{2\pi i \frac{ar}{q}} \frac{N}{\phi(q)} \;+\; \text{smaller} \\ &= \frac{N}{\phi(q)} \sum_{\substack{r=1\\ (r,q)=1}}^{q} e^{2\pi i \frac{ar}{q}}. \end{aligned} \tag{16.30}$$

We merely sketch what happens now.

First, one shows that for $x \in \mathcal{M}_{a,q}$ that $F_N(x)$ is very close to $F\Big(\frac{a}{q}\Big)$. This is a standard analysis (Taylor Series Expansion – the constant term is a good approximation if you are suf£ciently close).

Thus, as the major arcs are distinct,

$$\int_{\mathcal{M}} F_N^3(x) e(-mx) dx \;=\; \sum_{q=1}^{Q} \sum_{\substack{a=1\\ (a,q)=1}} \int_{\mathcal{M}_{a,q}} F_N^3(x) e(-mx) dx. \tag{16.31}$$

We can approximate $F_N^3(x)$ by $F\Big(\frac{a}{q}\Big)$; integrating a constant gives the constant times the length of the interval. Each of the major arcs has length $\frac{2Q^3}{N}$. Thus we £nd that, up to a smaller correction term, the contribution from the Major Arcs is

$$\begin{aligned}
\int_{\mathcal{M}} F_N^3(x)e(-mx)dx &= \frac{2Q^3}{N}\sum_{q=1}^{Q}\sum_{\substack{a=1\\(a,q)=1}}\left(\frac{N}{\phi(q)}\sum_{\substack{r=1\\(r,q)=1}}^{q} e^{2\pi i\frac{ar}{q}}\right)^3 e\Big(\frac{-2\pi ima}{q}\Big)\\
&= N^2\cdot 2Q^3\sum_{q=1}^{Q}\frac{1}{\phi(q)^3}\sum_{\substack{a=1\\(a,q)=1}}\left(\sum_{\substack{r=1\\(r,q)=1}}^{q} e^{2\pi i\frac{ar}{q}}\right)^3 e\Big(\frac{-2\pi ima}{q}\Big).
\end{aligned} \tag{16.32}$$

To complete the proof, we need to show that what is multiplying N^2 is non-negative, and not too small.

We will leave this for another day, as it is getting quite late here.

16.4 Why Goldbach is Hard

Using

$$F_N(x) = \sum_{p\le N}\log p\cdot e^{2\pi ipx}, \tag{16.33}$$

we ﬁnd we must study

$$\int_0^1 F_N^s(x)dx, \tag{16.34}$$

where $s = 3$ if we are looking at $p_1 + p_2 + p_3 = 2n + 1$ and $s = 2$ if we are looking at $p_1 + p_2 = 2n$. Why does the circle method work for $s = 3$ but fail for $s = 2$?

16.4.1 $s = 3$ Sketch

Let us recall *brieﬂy* the $s = 3$ case. Near rationals $\frac{a}{q}$ with *small* denominator (*small* means $q \le \log^B N$), we can evaluate $F_N(\frac{a}{q})$. Using Taylor, if x is very close to $\frac{a}{q}$, we expect $F_N(x)$ to be close to $F_N(\frac{a}{q})$.

The Major Arcs have size $\frac{\log^B N}{N}$. As $F_N(x)$ is around N near such rationals, we expect the integral of $F_N^3(x)e(-mx)$ to be N^2 times a power of $\log N$.

Doing a careful analysis of the singular series shows that the contribution is actually $\mathfrak{S}(N)N^2$, where there exist constants independent of N such that $0 < c_1 < \mathfrak{S}(N) < c_2 < \infty$.

A direct calculation shows that

$$\int_0^1 |F_N(x)|^2 dx \ = \ \int_0^1 F_N(x)F_N(-x)dx \ = \ N. \tag{16.35}$$

Thus, if m denotes the minor arcs,

$$\begin{aligned} \Big| \int_{\mathrm{m}} F_N^3(x)e(-mx)dx \Big| \ & \leq \ \max_{x \in \mathrm{m}} |F_N(x)| \int_0^1 |F_N(x)|^2 dx \\ & \leq \ N \max_{x \in \mathrm{m}} |F_N(x)|. \end{aligned} \tag{16.36}$$

As the major arcs contribute $\mathfrak{S}(N)N^2$, we need to show

$$\max_{x \in \mathrm{m}} |F_N(x)| \ \ll \ \frac{N}{\log^D N}. \tag{16.37}$$

Actually, we just need to show the above is $\ll o(N)$. This is the main dif£culty – the trivial bound is $|F_N(x)| \leq N$. As $F_N(0) = N$ plus lower order terms, we cannot do better in general.

Exercise 16.4.1. *Show $F_N(\frac{1}{2}) = N - 1$ plus lower order terms.*

The key observation is that, if we stay away from rationals with small denominator, we can prove there is cancellation in $F_N(x)$. While we don't go into details here (see, for example, Nathanson's Additive Number Theory: The Classical Bases, Chapter 7), the savings we obtain is small. We show

$$\max_{x \in \mathrm{m}} |F_N(x)| \ \ll \ \frac{N}{\log^D N}. \tag{16.38}$$

Note that Equation 16.35 gives us signi£cantly better cancellation on average, telling us that $|F_N(x)|^2$ is usually of size N.

Thus, it is our dream to be so lucky as to see $\left| \int_I |F_N(x)|^2 dx \right|$ for any $I \subset [0,1)$, as we can evaluate this extremely well.

16.4.2 $s = 2$ Sketch

What goes wrong when $s = 2$? As a £rst approximation, if $s = 3$ has the Major Arcs contributing a constant times N^2 (and $F_N(x)$ was of size N on the Major Arcs), one might guess that the Major Arcs for $s = 2$ will contribute a constant times N.

How should we estimate the contribution from the Minor Arcs? We have $F_N^2(x)$. If we just throw in absolute values we get

$$\Big| \int_{\mathrm{m}} F_N^2(x) e(-mx) dx \Big| \;\leq\; \int_0^1 |F_N(x)|^2 dx \;=\; N. \tag{16.39}$$

Note, unfortunately, that this is the same size as the expected contribution from the Major Arcs!

We could try pulling a $\max_{x \in \mathrm{m}} |F_N(x)|$ outside the integral, and hope to get a good savings. The problem is this leaves us with $\int_{\mathrm{m}} |F_N(x)| dx$.

Recall

Lemma 16.4.2.

$$\int_0^1 |f(x) g(x)| dx \;\leq\; \Big(\int_0^1 |f(x)|^2 dx \Big)^{\frac{1}{2}} \cdot \Big(\int_0^1 |g(x)|^2 dx \Big)^{\frac{1}{2}}. \tag{16.40}$$

For a proof, see Lemma 16.5.1.

Thus,

$$\begin{aligned} \Big| \int_{\mathrm{m}} F_N^2(x) e(-mx) dx \Big| \;&\leq\; \max_{x \in \mathrm{m}} |F_N(x)| \int_0^1 |F_N(x)| dx \\ &\leq\; \max_{x \in \mathrm{m}} |F_N(x)| \Big(\int_0^1 |F_N(x)|^2 dx \Big)^{\frac{1}{2}} \cdot \Big(\int_0^1 1^2 dx \Big)^{\frac{1}{2}} \\ &\leq\; \max_{x \in \mathrm{m}} |F_N(x)| \cdot N^{\frac{1}{2}} \cdot 1. \end{aligned} \tag{16.41}$$

As the Major Arcs contribute something of size N, we would need

$$\max_{x \in \mathrm{m}} |F_N(x)| \;\ll\; o(\sqrt{N}). \tag{16.42}$$

There is almost no chance of such cancellation. We know

$$\int_0^1 |F_N(x)|^2 dx \;=\; N \text{ plus lower order terms.} \tag{16.43}$$

Thus, the average size of $|F_N(x)|$ is N, so we expect $|F_N(x)|$ to be about $\sqrt{N}$. To get $o(N)$ would be unbelievably good fortune!

While the above sketch shows the Circle Method is not, at present, powerful enough to handle the Minor Arc contributions, all is not lost. The quantity we *need* to bound is

$$\Big| \int_{\mathrm{m}} F_N^2(x) e(-mx) dx \Big|. \tag{16.44}$$

However, we have instead been studying

$$\int_{\mathrm{m}} |F_N(x)|^2 dx \tag{16.45}$$

and

$$\max_{x \in \mathrm{m}} |F_N(x)| \int_0^1 |F_N(x)| dx. \tag{16.46}$$

Thus, we are ignoring the probable oscillation / cancellation in the integral $\int F_N(x) e(-mx) dx$. It is *this cancellation* that will lead to the Minor Arcs contributing signi£cantly less than the Major Arcs.

However, showing there is cancellation in the above integral is very dif£cult. It is a lot easier to work with absolute values.

16.5 Cauchy-Schwartz Inequality

Lemma 16.5.1. *[Cauchy-Schwarz]*

$$\int_0^1 |f(x)g(x)| dx \ \leq \ \Big(\int_0^1 |f(x)|^2 dx \Big)^{\frac{1}{2}} \cdot \Big(\int_0^1 |g(x)|^2 dx \Big)^{\frac{1}{2}}. \tag{16.47}$$

For notational simplicity, assume f and g are real-valued, positive functions. Working with $|f|$ and $|g|$ we see there is no harm in the above.

Let

$$h(x) = \ f(x) + \lambda g(x), \ \ \lambda \ = \ -\frac{\int_0^1 f(x)g(x)dx}{\int_0^1 g(x)^2 dx} \tag{16.48}$$

As $\int_0^1 h(x)^2 dx \geq 0$, we have

$$
\begin{aligned}
0 &\leq \int_0^1 \Big(f(x)+\lambda g(x)\Big)^2 dx \\
&= \int_0^1 f(x)^2 dx \;+\; 2\lambda \int_0^1 f(x)g(x)dx \;+\; \lambda^2 \int_0^1 g(x)^2 dx \\
&= \int_0^1 f(x)^2 dx \;-\; 2\frac{\Big(\int_0^1 f(x)g(x)dx\Big)^2}{\int_0^1 g(x)^2 dx} \;+\; \frac{\Big(\int_0^1 f(x)g(x)dx\Big)^2}{\int_0^1 g(x)^2 dx} \\
&= \int_0^1 f(x)^2 dx \;-\; \frac{\Big(\int_0^1 f(x)g(x)dx\Big)^2}{\int_0^1 g(x)^2 dx} \\
\frac{\Big(\int_0^1 f(x)g(x)dx\Big)^2}{\int_0^1 g(x)^2 dx} &\leq \int_0^1 f(x)^2 dx \\
\Big(\int_0^1 f(x)g(x)dx\Big)^2 &\leq \int_0^1 f(x)^2 dx \cdot \int_0^1 g(x)^2 dx \\
\int_0^1 f(x)g(x)dx &\leq \Big(\int_0^1 f(x)^2 dx\Big)^{\frac{1}{2}} \cdot \Big(\int_0^1 g(x)^2 dx\Big)^{\frac{1}{2}}. \qquad (16.49)
\end{aligned}
$$

Again, for general f and g, replace $f(x)$ with $|f(x)|$ and $g(x)$ with $|g(x)|$ above. Note there is nothing special about $\int_0^1$. $\square$

The Cauchy-Schwarz Inequality is often useful when $g(x) = 1$. In this special case, it is important that we integrate over a £nite interval.

Exercise 16.5.2. *For what f and g is the Cauchy-Schwarz Inequality an equality?*

16.6 Partial Summation

Lemma 16.6.1 (Partial Summation: Discrete Version). *Let $A_N = \sum_{n=1}^N a_n$. then*

$$\sum_{n=M}^{N} a_n b_n = A_N b_N - A_{M-1} b_M + \sum_{n=M}^{N-1} A_n (b_n - b_{n+1}) \qquad (16.50)$$

Proof. Since $A_n - A_{n_1} = a_n$,

$$\begin{aligned}
\sum_{n=M}^{N} a_n b_n &= \sum_{n=M}^{N} (A_n - A_{n-1}) b_n \\
&= (A_N - A_{N-1}) b_N + (A_{N-1} - A_{N-2}) b_{N-1} + \cdots + (A_M - A_{M-1}) b_M \\
&= A_N b_N + (-A_{N-1} b_N + A_{N-1} b_{N-1}) + \cdots + (-A_M b_{M+1} + A_M b_M) - a_{M-1} b_M \\
&= A_N b_N - a_{M_1} b_M + \sum_{n=M}^{N-1} A_n (b_n - b_{n+1}). \qquad (16.51)
\end{aligned}$$

□

Lemma 16.6.2 (Abel's Summation Formula - Integral Version). *Let $h(x)$ be a continuously differentiable function. Let $A(x) = \sum_{n \le x} a_n$. Then*

$$\sum_{n \le x} a_n h(n) = A(x) h(x) - \int_1^x A(u) h'(u) du \qquad (16.52)$$

See, for example, W. Rudin, *Principles of Mathematical Analysis*, page 70.

Partial Summation allows us to take knowledge of one quantity and convert that to knowledge of another.

For example, suppose we know that

$$\sum_{p \le x} \log p \ = \ x + O(x^{\frac{1}{2}+\epsilon}). \qquad (16.53)$$

We use this to glean information about $\sum_{p \le x} 1$.

De£ne

$$h(n) = \frac{1}{\log n} \quad \text{and} \quad a_n = \begin{cases} \log n & \text{if } n \text{ is prime} \\ 0 & \text{otherwise.} \end{cases} \qquad (16.54)$$

Applying partial summation to $\sum_{p \le x} a_n h(n)$ will give us knowledge about $\sum_{p \le x} 1$. Note as long as $h(n) = \frac{1}{\log n}$ for n prime, it doesn't matter how we de£ne $h(n)$ elsewhere; however, to use the integral version of Partial Summation, we need h to be a differentiable function.

Thus

$$\begin{aligned}
\sum_{p \le x} 1 &= \sum_{p \le x} a_n h(n) \\
&= \left(x + O(x^{\frac{1}{2}+\epsilon})\right)\frac{1}{\log x} - \int_2^x \left(u + O(u^{\frac{1}{2}+\epsilon})\right) h'(u) du. \quad (16.55)
\end{aligned}$$

The main term $(A(x)h(x))$ equals $\frac{x}{\log x}$ plus a signi£cantly smaller error.

We now calculate the integral, noting $h'(u) = -\frac{1}{u \log^2 u}$. The error piece in the integral gives a constant multiple of

$$\int_2^x \frac{u^{\frac{1}{2}+\epsilon}}{u \log^2 u} du. \quad (16.56)$$

As $\frac{1}{\log^2 u} \le \frac{1}{\log^2 2}$ for $2 \le u \le x$, the integral is bounded by

$$\frac{1}{\log^2 2} \int_2^x u^{-\frac{1}{2}+\epsilon} < \frac{1}{\log^2 2} \frac{1}{\frac{1}{2}+\epsilon} x^{\frac{1}{2}+\epsilon}, \quad (16.57)$$

which is signi£cantly less than $A(x)h(x) = \frac{x}{\log x}$.

We now need to handle the other integral:

$$\int_2^x \frac{u}{u \log^2 u} du = \int_2^x \frac{1}{\log^2 u} du. \quad (16.58)$$

The obvious approximation to try is $\frac{1}{\log^2 u} \le \frac{1}{\log^2 2}$. Unfortunately, plugging this in bounds the integral by $\frac{x}{\log^2 2}$. This is larger than the expected main term, $A(x)h(x)$!

As a rule of thumb, whenever you are trying to bound something, try the simplest, most trivial bounds £rst. Only if they fail should you try to be clever.

Here, we need to be clever, as we are bounding the integral by something larger than the observed terms.

We split the integral into two pieces:

$$\int_2^x = \int_2^{\sqrt{x}} + \int_{\sqrt{x}}^x \quad (16.59)$$

For the £rst piece, we use the trivial bound for $\frac{1}{\log^2 u}$. Note the interval has length $\sqrt{x} - 2 < \sqrt{x}$. Thus, the £rst piece contributes at most $\frac{x^{\frac{1}{2}}}{\log^2 2}$, signi£cantly less than $A(x)h(x)$.

The reason trivial bounds failed for the entire integral is the length was too large (of size x); there wasn't enough decay in the function.

The advantage of splitting the integral in two is that in the second piece, even though most of the length of the original interval is here (it is of length $x - \sqrt{x} \approx x$), the function $\frac{1}{\log^2 u}$ is small here. Instead of bounding it by a constant, we now bound it by substituting in the smallest value of u on this interval, $\sqrt{x}$. Thus, the contribution from this integral is at most $\frac{x-\sqrt{x}}{\log^2 \sqrt{x}} < \frac{4x}{\log^2 x}$. Note that this is significantly less than the main term $A(x)h(x) = \frac{x}{\log x}$.

Chapter 17

Multiplicative Functions, Kloosterman, p-adic Numbers, and Review of the Three Problems: Germain Primes, $\lambda_1(G)$ for Random Graphs, Randomness of $x \to \overline{x}$ mod p

Multiplicative Functions, Kloosterman Sums and Bounds, p-adic Numbers. Review of the Three Problems (Germain Primes, Randomness of $x \to \overline{x}$ mod p, Random Graphs). Lecture by Peter Sarnak; notes by Steven Miller.

17.1 Multiplicative Functions, Kloosterman and p-adic Numbers

17.1.1 Multiplicative Functions

De£nition 17.1.1 (Multiplicative Functions). *Let f be de£ned on the positive integers $\mathbb{N}$. f is multiplicative if*

$$f(mn) \ = \ f(m)f(n) \ \text{ if } m \text{ and } n \text{ are relatively prime.} \qquad (17.1)$$

This is the same as

$$\sum_{n=1}^{\infty} \frac{f(n)}{n^s} = \prod_p \Big(1 + f(p)p^{-s} + f(p^2)p^{-2s} + \cdots\Big). \tag{17.2}$$

We call the above an **Euler Product**. The standard example is the **Riemann Zeta Function**,

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_p \frac{1}{1-p^{-s}}. \tag{17.3}$$

There are many multiplicative functions (some, like Dirichlet Characters, do not even require m and n to be relatively prime).

The Kloosterman sums are *not* multiplicative.

17.1.2 Kloosterman Sums

If c_1 and c_2 are relatively prime,

$$K(a, b, c_1c_2) = K(*, *, c_1) \cdot K(*, *, c_2), \tag{17.4}$$

where the $*$s are functions of a, b, c_1 and c_2.

Say we show $|K(a, b, p)| \ll p^{\nu}$, where ν does not depend on a or b or p.

Then, if $c = \prod_i p_i^{r_i}$, we have

$$K\Big(a, b, \prod_i p_i^{r_i}\Big) = \prod_i K(*, *, p_i^{r_i}). \tag{17.5}$$

Thus, we just need to get bounds over prime powers to bound a general Kloosterman sum.

Theorem 17.1.2 (Salie). *If* $\alpha \geq 2$*,*

$$K(a, b, p^{\alpha}) \leq 2p^{\frac{\alpha}{2}}. \tag{17.6}$$

Proof: elementary.

$$\sum_{x \bmod p^2} = \sum_{x_1=0}^{p-1} \sum_{x_2=0}^{p-1}, \tag{17.7}$$

where $x = x_1 + x_2 p$ and $x_1, x_2 \in \{0, 1, \ldots, p-1\}$.
Thus, when we encounter terms like $e\left(\frac{x_1+px_2}{p^2}\right)$, we need the inverse of $\overline{x_1 + px_2}$.
Let x_1^{-1} be the inverse of x_1 mod p. Then

$$(x_1 + px_2)^{-1} \;=\; x_1^{-1}(1 + x_1^{-1}x_2 p)^{-1}. \tag{17.8}$$

Note that

$$(1 - pb)^{-1} \;=\; 1 + pb + O(p^2). \tag{17.9}$$

Say we want the inverse mod p^2 of $(1 - bp)$. Try multiplying by $(1 + bp)$. We get

$$(1 - pb)(1 + pb) \;=\; 1 - b^2p^2 \;\equiv\; 1 \bmod p^2. \tag{17.10}$$

The above arguments is Hensel's Lemma.

17.1.3 p-adic numbers

We de£ne the p-adic norm of a rational $\alpha = \frac{a}{b}$, a and b relatively prime, by

$$||\alpha||_p \;=\; p^{-m}, \text{ where } \frac{a}{b} = p^m l,\ (p, l) = 1. \tag{17.11}$$

Note that numbers that are highly divisible by p are small p-adically.

We have the rationals $\mathbb{Q}$ and the p-adic norm $|| * ||_p$. Similar to completing the rationals $\mathbb{Q}$ with the usual norm to get $\mathbb{R}$, we can complete the rationals with respect to this norm. The resulting £eld is called the p-adic numbers, $\mathbb{Q}_p$.

$\mathbb{Q} \subset \mathbb{R}$, and $\mathbb{Q}$ is dense in $\mathbb{R}$. Similarly $\mathbb{Q} \subset \mathbb{Q}_p$ and $\mathbb{Q}$ is dense in $\mathbb{Q}_p$.

Let $x \in \mathbb{Q}_p$. Then

$$x \;=\; \frac{a_{-m}}{p^m} + \frac{a_{-m+1}}{p^{m-1}} + \cdots + a_0 + a_1 p + a_2 p^2 + \cdots, \tag{17.12}$$

where $0 \le a_i \le p - 1$.

Suppose we have a solution $f(x_0) \equiv 0 \bmod p$. We then try and £nd x_1 such that $f(x_0 + px_1) \equiv 0 \bmod p^2$. Hensel noted that all we need to £nd x_1 is some knowledge of the derivative at the previous stage.

17.2 Germain Primes

$p - 1 = 2q$, p and q prime. What are the statistics? How many are there up to x? Do they know about each other? What are their correlations? What about $p - 3$?

The Circle Method is a way of trying to solve additive equations (Waring's Problem, Goldbach's Problem $p_1+p_2 = 2n$, Vinogradov's Three Primes Theorem $p_1 + p_2 + p_3 = 2n + 1$, Twin Primes $p_2 - p_1 = 2$).

De£nition 17.2.1 (Germain Primes). *If p is prime and $p-1 = 2q$ for some prime q, we say p is a Germain Prime.*

De£nition 17.2.2 ($\pi_G(x)$). *Recall $\pi(x)$ is the number of primes at most x. Then $\pi(x) \sim \frac{x}{\log x}$. Let $\pi_G(x)$ be the number of Germain primes at most x. If the probability of getting a prime is* $\log x$*, then we might expect that*

$$\pi_G(x) \quad = \quad \sum_{\substack{p\leq x\\ p-1=2q}} 1 \quad \sim \quad \text{Const} \cdot \frac{x}{\log^2 x}. \tag{17.13}$$

Using the circle method, we will try and estimate the above constant, and hope the minor arcs *do not contribute to the main term.*

Major McMahan (from the army, friendly with Hardy and Littlewood) made tables of primes to the *millions.* He checked, and Hardy and Littlewood's constant (for twin primes) was correct and Sylvester (who made a probabilistic argument) was shown to be slightly off.

See Hardy and Littlewood, Acta Mathematica, v44, 1923, Partitio numerorum. III: *On the expression of a number as a sum of primes.*

We will then investigate the nearest neighbor and k^{th}-nearest neighbor spacings.

Also look at Robert Vaughn, *The Hardy Littlewood Method.*

17.3 Randomness of $x \to \overline{x}$

Given a prime p, look at $x \to \overline{x}$. How do we compute $\overline{x}$?

One can compute $\overline{x}$ by using the Euclidean Algorithm (very fast, $\log p$ steps). Recall the Euclidean Algorithm gives a and b such that $ax + bp = 1$. Thus, mod p, $ax \equiv 1$, or $a = \overline{x}$ mod p.

We now study the spacings between $\overline{x}$ as x ranges in some interval mod p. If the interval is very small, we don't expect randomness. What if we take an interval of length $\sqrt{p}$. Do we see Poissonian Behavior there for a £xed prime?

Now, £x a number a. Look at $\frac{a \bmod p}{p}$ as you vary p.

Theorem 17.3.1 (Duke-Iwaniec). *Suppose $p \equiv 1$ mod 4. There is a square root of -1 mod p; ie, $\exists x$ such that $x^2 \equiv 1$ mod p. Now we can take $1 \le x \le \frac{p-1}{2}$, so $\frac{x}{p} \in \left[0, \frac{1}{2}\right]$. Then $\frac{x}{p}$, as we vary p, is equi-distributed.*

One can also look at the longest increasing sub-sequence.

Knuth, volume 2 of the Art of Computer Programming. Look at the stuff on generating random numbers.

17.4 Random Graphs / Ramanujan Graphs

Bollobas, *Random Graphs*: he will have a model of the random 3-regular graphs (what it means, how to generate, how many are there). *Very* hard if you don't distinguish between isomorphic graphs (which have the same spectrum).

Look at the distribution of the second largest eigenvalue λ_1 of the random 3-regular graph. Find the mean and the variance, graph.

Professor Sarnak will give a lecture on the Tracy-Widom distribution (which is the distribution of the biggest eigenvalue in some random ensemble – will we see the same distribution here)?

What is the scale for normalizing?

Take an interval, see how many eigenvalues in it, slide the interval down, and see how the number varies.

Chapter 18

Random Graphs, Autocorrelations, Random Matrix Theory and the Mehta-Gaudin Theorem

Random Graphs (especially graphs with large girth and chromatic number), Autocorrelation, Random Matrix Theory (Vandermonde determinants, orthonormal polynomials) and the Mehta-Gaudin Theorem (large N limits of quantities related to the joint density function of the eigenvalues). Lecture by Peter Sarnak; notes by Steven Miller.

18.1 Random Graphs

De£nition 18.1.1 (Chromatic Number). *The chromatic number is the least number of colors such that each vertex has a different color than all of its neighbors.*

Example 18.1.2. *A bi-partite graph is* 2*-colorable, as is a tree (alternate colors as you go through the generations).*

What can force you to have a lot of colors? If a vertex is joined to n_v vertices, you need a lot of colors *if* the vertices it is joined to are joined to each other.

De£nition 18.1.3 (Girth). *The girth is the shortest closed cycle.*

If the girth is large, make a vertex blue, yellow next level, blue on next level, et cetera until you come back on yourself.

Question 18.1.4. *Can you make a graph with large girth and large chromatic number?*

The two £ght each other. Erdos solved this problem by showing that if you take a *Random Graph* (with suitable properties), then that graph will have large girth and large chromatic number with high probability.

Take a Random Graph with n vertices and basically $n^{1+\epsilon}$ edges placed at random among the $\binom{n}{2} = O(n^2)$ possible edges.

The random graph has short cycles, but the number of short cycles is small. Erdos removes certain graphs with small girth, and shows with high probability the graphs left have large girth and large chromatic number.

See Mckay's paper: he proves Kesten's measure holds for the random graph as the number of vertices goes to in£nity.

18.2 Baire Category

Given $\alpha \notin \mathbb{Q}$ and C_α, how often can we £nd $\frac{a}{q} \in \mathbb{Q}$ such that

$$\left|\alpha - \frac{a}{q}\right| \ \geq \ \frac{C_\alpha}{q^{2+\epsilon}}. \tag{18.1}$$

In Lebesgue Measure, almost all α satisfy the above in£nitely often.

In the Baire Category, this inequality does not hold in£nitely often.

18.3 Autocorrelation

Note: Alex Barnett lectured on this section.

x-axis is number of swaps n, y-axis is number of graphs with given λ_1, $\lambda_1(n)$. Say takes 100 swaps to randomize. Then the y-value at 101 swaps should be independent of the y-value at 1 swap.

But we don't know the number of swaps before we have moved far enough.

Let $\lambda_1'(n) = \lambda_1(n) - \overline{\lambda}_1$, where $\overline{\lambda}_1$ is the average value.

Autocorrelation: Say the x-axis runs to m.

$$A(c) \ = \ \frac{1}{m}\sum_n \lambda_1'(n)\lambda_1'(n+c). \tag{18.2}$$

The above is a function of c, symmetric in c. As c gets large, $A(c)$ dies to zero, and has largest value at $c = 0$.

Look for c such that say 10% of the area is from c onward.

18.4 Gaudin's Method

18.4.1 Introduction

From Random Matrix Theory: we have a probability distribution on $\mathbb{R}^n$:

$$p_\beta(x_1, \ldots, x_N) \;=\; c_N(\beta) \prod_{j<k} |x_j - x_k|^\beta e^{-\sum_{j=1}^N x_j^2} dx_1 \cdots dx_N. \tag{18.3}$$

Start off with a real $N \times N$ matrix, diagonalize with eigenvalues $x_1, \ldots, x_N$. If you choose the matrix at random, you get N numbers, and you have a probability distribution on the eigenvalues.

We've derived the probability above for $N \times N$ matrices. For convenience, we order the eigenvalues.

If $\beta = 1$ we call the ensemble GOE (Gaussian Orthogonal Ensembles); if $\beta = 2$ we have GUE (Unitary) and if $\beta = 4$ we have GSE (Symplectic).

What is the correlation between two eigenvalues? What is the probability of observing a given spacing between two eigenvalues? We've done this in the 2×2 case.

In the $N \times N$ case, we would need to integrate out most of the eigenvalues. The dif£culty is $\prod |x_j - x_k|^\beta$.

For $\beta = 1, 2$ or 4, we can evaluate these integrals; we are fortunate that these values are the ones that arise in practice.

In fact, even just determining $c_N(\beta)$ is dif£cult. This is called the *Selberg Integral*, which A. Selberg solved in high school!

We will only consider $\beta = 2$, and will be interested in the limit as $N \to \infty$ (under appropriate re-scaling).

$$R_N(x_1, \ldots, x_N) \;=\; \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} p_2(x_1, x_2, \ldots, x_n, x_{n+1}, \ldots, x_N) dx_{n+1} \cdots dx_N. \tag{18.4}$$

This will be a symmetric function of the £rst n variables. If we integrate all but 1 variable we get the density of eigenvalues; if we integrate all but two we get information on pairs of eigenvalues.

Remark 18.4.1. *$\beta = 0$ is Poissonian.*

18.4.2 Vandermonde Determinants

Notation: dp means

$$dp(\theta_1, \ldots, \theta_N) \;=\; c_N \prod_{j<k} \left| e^{i\theta_j} - e^{i\theta_k} \right|^2 d\theta_1 \cdots d\theta_N. \tag{18.5}$$

We are now working on the N-torus $[0, 2\pi] \times \cdots \times [0, 2\pi]$. This goes under the name CUE (Circular Unitary Ensemble).

Remember the group

$$U(N) = \{N \times N \text{ matrices } A \text{ with } AA^\star = I\}. \tag{18.6}$$

Similar to the diagonalization of symmetric matrices, for any unitary matrix U there is a unitary matrix V such that $V^{-1}UV$ is diagonal; further, the eigenvalues have absolute value 1, and hence can be written as $e^{i\theta}$.

Suppose we have $f_1, \ldots, f_N$. We form the Vandermonde of the N-variables

$$\text{Van}(f_1, \ldots, f_N) \;=\; \prod_{i<j} (f_i - f_j). \tag{18.7}$$

Today we will only use the square, so we don't worry about ordering so that $f_i < f_j$.

Exercise 18.4.2.

$$\text{Van}(f_1, \ldots, f_N) \;=\; \det \left(f_i^{j-1} \right)_{1 \le i,j \le N}. \tag{18.8}$$

Thus, we have

$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ f_1 & f_2 & \cdots & f_N \\ \vdots & \vdots & \ddots & \vdots \\ f_1^{N-1} & f_2^{N-1} & \cdots & f_N^{N-1} \end{vmatrix} \tag{18.9}$$

18.4.3 Orthonormal Polynomials

On the unit circle T, we have the measure

$$d\mu(t) \;=\; \frac{dt}{2\pi}. \tag{18.10}$$

Let $f(t)$ be a function such that

$$\int_T f(t)d\mu(t) \ = \ 0, \quad \int_T |f(t)|^2 d\mu(t) = 1. \tag{18.11}$$

De£ne a sequence of monic polynomials $P_n(x)$ for $n \in \mathbb{N}$ and $\phi_n(t)$ with

$$\phi_n(t) \ = \ P_n(f(t)), \quad \phi_0(t) \ = \ \frac{1}{\sqrt{\mu(T)}}, \quad \int_T \phi_i(t)\overline{\phi}_j(t)d\mu(t) = \delta_{ij}. \tag{18.12}$$

This is Gramm-Schmidt, where the inner product between two functions f and g is given by

$$\langle f, g \rangle \ = \ \int_T f(t)\overline{g}(t)d\mu(t), \tag{18.13}$$

and the 'Kronecker delta' symbol (the discrete analog of the continuous delta 'function' $\delta(\cdot)$) is de£ned by

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases} \tag{18.14}$$

We introduce orthogonal polynomials to handle the integral. The above process (constructing the P_ns and the ϕ_ns) gives an orthonormal sequence of polynomials.

18.4.4 Kernel $K_N(x, y)$

De£ne the kernel

$$K_N(x, y) \ = \ \sum_{j=0}^{N-1} \phi_j(x)\overline{\phi}_j(y). \tag{18.15}$$

Exercise 18.4.3. *Prove the following:*

1. $\int_T K_N(x, x)d\mu(x) \ = \ N.$
2. $\int_T K_N(x, y)K_N(y, z)d\mu(y) \ = \ K_N(x, z).$

Remark 18.4.4.

$$\int_T K_N(x,y)g(y)d\mu(y) \;=\; \sum_{j=0}^{N-1}\Big[\int_T \overline{\phi}_j(y)g(y)d\mu(y)\Big]\phi_j(x). \qquad (18.16)$$

Thus, integrating g against K_N projects g onto the £rst N vectors.

De£ne, for $1 \le n \le N$,

$$D_{n,N}(t_1,\dots,t_n) \;=\; \det\Big(K_N(t_j,t_k)\Big)_{1\le j,k\le n}. \qquad (18.17)$$

For example,

$$D_{1,N} \;=\; K_N(t_1,t_1) \qquad (18.18)$$

and

$$D_{2,N} \;=\; \begin{vmatrix} K_N(t_1,t_1) & K_N(t_1,t_2) \\ K_N(t_2,t_1) & K_N(t_2,t_2) \end{vmatrix} \qquad (18.19)$$

18.4.5 Gaudin-Mehta Theorem

Theorem 18.4.5 (Gaudin-Mehta). *We have*

1.

$$\frac{1}{\mu(T)}\mathrm{Van}\Big(f(t_1),\dots,f(t_N)\Big) \;=\; \mathrm{det}_{N\times N}\Big(\phi_{i-1}(t_j)\Big)_{1\le i,j\le N}. \qquad (18.20)$$

2.

$$\frac{1}{\mu(T)}\Big|\mathrm{Van}\Big(f(t_1),\dots,f(t_N)\Big)\Big|^2 \;=\; D_{N,N}(t_1,\dots,t_N). \qquad (18.21)$$

3. For $1 \le n \le N$,

$$\int_T D_{n,N}(t_1,\dots,t_n)d\mu(t_n) \;=\; (N+1-n)D_{n-1,N}(t_1,\dots,t_{n-1}). \qquad (18.22)$$

The third statement is the beef, allowing us to integrate out one variable at a time by induction.

Remember

$$D_{n,N}(t_1,\dots,t_n) \;=\; \det{}_{n\times n}\Big(K_N(t_j,t_k)\Big)_{1\le j,k\le n}. \tag{18.23}$$

Corollary 18.4.6. *Let F be a symmetric function of $t_1,\dots,t_n$, and de£ne*

$$\begin{aligned} F_{n,N}(t_1,\dots,t_N) &= \sum_{1\le i_1<i_2\cdots<i_n<N} F(t_{i_1},\dots,t_{i_n}) \\ d\mu_{n,N}(t_1,\dots,t_N) &= \frac{1}{n!}D_{n,N}(t_1,\dots,t_N)d\mu(t_1)\cdots d\mu(t_N). \end{aligned} \tag{18.24}$$

Then

$$\int_{T^N} F_{N,N}(t_1,\dots,t_N)d\mu_{n,N}(t_1,\dots,t_N) \;=\; \int_{T^n} F(t_1,\dots,t_n)d\mu_{n,N}(t_1,\dots,t_n). \tag{18.25}$$

How might we use the above? For example, consider for $1\le j,k\le N$, and consider for f even

$$\sum_{1\le j<k\le N} f(x_j-x_k). \tag{18.26}$$

What is the expectation of the above? In this case, F is a function of two variables, and $F(x_1,x_2)=f(|x_1-x_2|)$ and we now just need to integrate $f(|x_1-x_2|)$ against the determinant of a 2×2 matrix, and this is the only place where N will arise.

Suppose we had

$$dp(x_1,\dots,x_N) \;=\; e^{-\sum x_j^2}\prod_{j<k}|x_j-x_k|^2dx_1\cdots dx_N. \tag{18.27}$$

Consider the expectation of

$$\sum_{1\le j<k\le N} f(|x_j-x_k|). \tag{18.28}$$

According to the corollary, the answer is just

$$\int_{\mathbb{R}}\int_{\mathbb{R}} f(|x_1 - x_2|)\frac{1}{2!}\begin{vmatrix} K_N(x_1,x_1) & K_N(x_1,x_2) \\ K_N(x_2,x_1) & K_N(x_2,x_2) \end{vmatrix} e^{-x_1^2-x_2^2}dx_1dx_2. \tag{18.29}$$

This is *enormous* progress – we started with N variables; we now have 2 variables. We need to take the $N \to \infty$ limit of the determinant, a much easier question.

18.4.6 Example

$T = [0, 2\pi]$, $d\mu(x) = \frac{dx}{2\pi}$, $f(x) = e^{ix}$, $f^n(x) = e^{inx}$, and $P_n(x) = x^n$. These P_ns are monic, $\phi_n(x) = P_n(f(x))$ is orthonormal, which gives $\phi_n(x) = e^{inx}$, and clearly

$$\int_0^{2\pi} e^{inx}e^{-imx}dx \ = \ \delta_{ij}. \tag{18.30}$$

Finally, we obtain a geometric progression

$$\begin{aligned} K_N(x,y) &= \sum_{n=0}^{N-1} e^{in(x-y)} \\ &= \frac{1-e^{iN(x-y)}}{1-e^{i(x-y)}}. \end{aligned} \tag{18.31}$$

We will symmetrize (and go from $-N$ to N), and when we take the 2×2 determinant, we get something like

$$\frac{\sin\left(\frac{N(x-y)}{2}\right)}{\sin\left(\frac{x-y}{2}\right)}. \tag{18.32}$$

The most famous pair correlation: we have N eigenvalues so that the mean spacing is 1. The *pair correlation* is

$$1 - \left[\frac{\sin\Big(\pi(x-y)\Big)}{\pi(x-y)}\right]^2. \tag{18.33}$$

Chapter 19

Increasing Length Subsequences and Tracy-Widom

More on increasing length subsequence and the Tracy-Widom distribution. Lecture by Peter Sarnak, notes by Steven Miller.

19.1 Increasing Length Subsequences

Consider the set of permutations S_n on n numbers. Let $\sigma \in S_n$ be a random permutation, and let L_σ be the length of the longest increasing sub-sequence. What is the expected value of L_σ?

Conjecture 19.1.1 (Ulam).

$$E[L_\sigma] \sim 2\sqrt{n}. \tag{19.1}$$

Proved by several people (Schepp, Vircheck (?), ...).

At Bell Labs, many people (including Odlyzko) investigated. Monte-Carlo simulations for variance (beginning 1993). Dividing variance by $n^{\frac{1}{3}}$ was good. Looking at the expectation of

$$\frac{L_\sigma - 2\sqrt{n}}{n^{\frac{1}{6}}} \tag{19.2}$$

and investigated whether or not it went to a limit. Noticed this distribution is negative (shifted to the left). Prefers to be *less* than $2\sqrt{n}$.

19.2 Tracy-Widom

On $\mathbb{R}^n$, we have

$$P_N(x)dx \;=\; e^{-\sum_{j=1}^N x_j^2} \prod_{j<k} |x_j - x_i|^\beta \, dx_1 \cdots dx_n, \;\; \beta \in \{1,2,4\}. \tag{19.3}$$

Will use the weight $e^{-x^2}dx$, which will give rise to Hermite polynomials.

Definition 19.2.1 ($F_{N,\beta}(s)$). *$F_{N,\beta}(s)$ is the probability that there is no $x \in [s,\infty)$.*

We can write this as a determinant (see the papers by Mehta and Mehta-Gaudin). We will have again

$$K_N(x,y) \;=\; \sum_{0 \le j \le N-1} \phi_j(x)\overline{\phi}_j(y). \tag{19.4}$$

Remember the semi-circle rule, that (with some normalization) the eigenvalues lie in $[-2\sqrt{N}, 2\sqrt{N}]$.

What is the expected value of the largest eigenvalue? We know most are in $[-2\sqrt{N}, 2\sqrt{N}]$. We haven't discussed whether or not there are outliers. With probability one, we can show that there will be no such outliers.

We have N eigenvalues. Near 0 is called *the bulk*. As there are N numbers, we expect eigenvalues in an interval of size $\frac{1}{N}$ near the origin.

What about eigenvalues near the edges, $\pm 2\sqrt{N}$? Let $s \in (-\infty,\infty)$. Look at

$$F_N\Big(2\sqrt{N} + \frac{s}{N^{\frac{1}{6}}}\Big). \tag{19.5}$$

This is the scaling limit. Say $s = 0$. this would give us what is happening at the origin.

Theorem 19.2.2 (Tracy-Widom).

$$\lim_{N\to\infty} F_N\Big(2\sqrt{N} + \frac{s}{N^{\frac{1}{6}}}\Big) \;=\; F_\beta(s), \;\; F_\beta(s) = \frac{dF_\beta(s)}{ds}. \tag{19.6}$$

Here the $N^{\frac{1}{6}}$ arises from the particular problem we're interested in. Here, we are looking at eigenvalues from random matrices.

Chapter 20

Circle Method and Germain Primes

Using the Hardy-Littlewood Circle Method (and assuming no main term contribution from the Minor Arcs), we calculate the expected number of Germain primes. Calculations and notes by Steven Miller.

20.1 Preliminaries

20.1.1 De£nitions

Let

$$e(x) \quad = \quad e^{2\pi i x} \tag{20.1}$$

and

$$\lambda(n) \;=\; \begin{cases} \log p & \text{if } n = p \text{ is prime} \\ 0 & \text{otherwise} \end{cases} \tag{20.2}$$

Finally, de£ne

$$c_q(a) \;=\; \sum_{\substack{r=1 \\ (r,q)=1}}^{q} e\Big(r\frac{a}{q}\Big). \tag{20.3}$$

20.1.2 Partial Summation

Lemma 20.1.1 (Partial Summation: Discrete Version). *Let* $A_N = \sum_{n=1}^{N} a_n$*. then*

$$\sum_{n=M}^{N} a_n b_n = A_N b_N - A_{M-1} b_M + \sum_{n=M}^{N-1} A_n (b_n - b_{n+1}) \tag{20.4}$$

Proof. Since $A_n - A_{n-1} = a_n$,

$$\begin{aligned}
\sum_{n=M}^{N} a_n b_n &= \sum_{n=M}^{N} (A_n - A_{n-1}) b_n \\
&= (A_N - A_{N-1}) b_N + (A_{N-1} - A_{N-2}) b_{N-1} + \cdots + (A_M - A_{M-1}) b_M \\
&= A_N b_N + (-A_{N-1} b_N + A_{N-1} b_{N-1}) + \cdots + (-A_M b_{M+1} + A_M b_M) - a_{M-1} b_M \\
&= A_N b_N - a_{M_1} b_M + \sum_{n=M}^{N-1} A_n (b_n - b_{n+1}).
\end{aligned} \tag{20.5}$$

□

Lemma 20.1.2 (Abel's Summation Formula - Integral Version). *Let* $h(x)$ *be a continuously differentiable function. Let* $A(x) = \sum_{n \le x} a_n$*. Then*

$$\sum_{n \le x} a_n h(n) = A(x) h(x) - \int_1^x A(u) h'(u) du \tag{20.6}$$

See, for example, W. Rudin, *Principles of Mathematical Analysis*, page 70.

20.1.3 Siegel-Wal£sz

Theorem 20.1.3. *[Siegel-Wal£sz] Let* $C, B > 0$*, and let* a *and* q *be relatively prime. Then*

$$\sum_{\substack{p \le x \\ p \equiv a(q)}} \log p \; = \; \frac{x}{\phi(q)} + O\Big(\frac{x}{\log^C x}\Big) \tag{20.7}$$

for $q \le \log^B x$*, and the constant above does not depend on* x*,* q *or* a *(ie, it only depends on* C *and* B*).*

20.1.4 Germain Integral

De£ne

$$\begin{aligned} f_{1N}(x) &= \sum_{p_1 \leq N} \log p_1 \cdot e(p_1 x) \\ f_{2N}(x) &= \sum_{p_2 \leq N} \log p_2 \cdot e(-2p_2 x) \\ f_N(x) &= \sum_{p_1 \leq N} \sum_{p_2 \leq N} \log p_1 \log p_2 \cdot e\Big((p_1 - 2p_2)x\Big). \end{aligned} \tag{20.8}$$

Consider

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} f_N(x)e(-x)dx = \sum_{p_1 \leq N} \sum_{p_2 \leq N} \log p_1 \log p_2 \int_{-\frac{1}{2}}^{\frac{1}{2}} e\Big((p_1 - 2p_2 - 1)x\Big)dx. \tag{20.9}$$

Note

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} e\Big((p_1 - 2p_2 - 1)x\Big)dx = \begin{cases} 1 & \text{if } p_1 - 2p_2 - 1 = 0 \\ 0 & \text{if } p_1 - 2p_2 - 1 \neq 0 \end{cases} \tag{20.10}$$

Thus, we get a contribution of $\log p_1 \log p_2$ if p_1 and $p_2 = \frac{p_1 - 1}{2}$ are both primes. Thus,

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} f_N(x)e(-x)dx = \sum_{\substack{p_1 \leq N \\ p_2 = \frac{p_1 - 1}{2} \text{ prime}}} \log p_1 \log p_2. \tag{20.11}$$

The above is a weighted counting of Germain primes.

20.1.5 Major and Minor Arcs

Let B be a positive integer, $Q = \log^B N$, and de£ne the Major Arc $\mathcal{M}_{a,q}$

$$\mathcal{M}_{a,q} = \Big\{x \in [0,1) : \Big|x - \frac{a}{q}\Big| < \frac{Q}{N}\Big\}. \tag{20.12}$$

We also add in one interval centered at either 0 or 1, ie, the "interval" (or wrapped-around interval)

$$\left[0, \frac{Q}{N}\right] \ \cup \ \left[1 - \frac{Q}{N}, 1\right]. \tag{20.13}$$

For convenience, we often use the interval $[-\frac{1}{2}, \frac{1}{2}]$ instead of $[0, 1]$, in which case we would have

$$\left[-\frac{1}{2}, -\frac{1}{2} + \frac{Q}{N}\right] \bigcup \left[\frac{1}{2} - \frac{Q}{N}, \frac{1}{2}\right]. \tag{20.14}$$

For functions that are periodic of period one, we could instead consider

$$\left[\frac{1}{2} - \frac{Q}{N}, \frac{1}{2} + \frac{Q}{N}\right]. \tag{20.15}$$

The Major Arcs are de£ned by

$$\mathcal{M} \ = \ \bigcup_{q \le Q} \bigcup_{\substack{a=1 \\ (a,q)=1}}^{q} \mathcal{M}_{a,q}. \tag{20.16}$$

The Minor Arcs, m, are whatever is *not* in the Major Arcs.
Then

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} f_N(x) e(-x) dx \ = \ \int_{\mathcal{M}} f_N(x) e(-x) dx \ + \ \int_{\mathrm{m}} f_N(x) e(-x) dx. \tag{20.17}$$

We will assume that there is no net contribution over the minor arcs. Thus, in the sequel we investigate

$$\int_{\mathcal{M}} f_N(x) e(-x) dx. \tag{20.18}$$

20.1.6 Reformulation of Germain Integral

$$\begin{aligned}
f_{1N}(x) &= \sum_{m_1 \le N} \lambda(m_1) \cdot e(m_1 x) \\
f_{2N}(x) &= \sum_{m_2 \le N} \lambda(m_2) \cdot e(-2m_2 x) \\
f_N(x) &= \sum_{m_1 \le N} \sum_{m_2 \le N} \lambda(m_1)\lambda(m_2) \cdot e\Big((m_1 - 2m_2)x\Big). \qquad (20.19)
\end{aligned}$$

We investigate

$$\int_{\mathcal{M}} f_N(x) e(-x) dx. \qquad (20.20)$$

We will show the Major Arcs contribute, up to lower order terms, $T_2 N$, where T_2 is a constant independent of N. The length of the Major Arc $\mathcal{M}_{a,q}$ is $\frac{Q}{N}$. We sum over $(a, q) = 1$ and $q \le Q$. Thus, the total length is bounded by

$$\sum_{q \le Q} q \cdot \frac{Q}{N} \ll \frac{Q^3}{N} \ll \frac{\log^B}{N}. \qquad (20.21)$$

By choosing B suf£ciently large, we will be able to make all the errors from the Major Arc calculations less than the main term from the Major Arcs. Of course, we have absolutely no control over what happens on the minor arcs, and we will simply assume there is no contribution from the minor arcs.

Thus, on the Major Arc $\mathcal{M}_{a,q}$, success will be in £nding a function of size N^2 such that the error from this function to $f_N(x)$ on $\mathcal{M}_{a,q}$ is much smaller than N^2, say N^2 divided by a large power of $\log N$.

Similarly, when we integrate over the Major Arcs, we will £nd the main terms will be of size N; again, success will be in showing the errors in the approximations are much smaller than N, say N divided by a large power of $\log N$.

We are able to do this because of the Siegel-Wal£sz Theorem (Theorem 20.1.3). Given *any* $B > 0$, we can £nd a $C > 0$ such that, if $q \le \log^B N$, then

$$\sum_{\substack{p \le N \\ p \equiv r(q)}} \log p = \frac{N}{\phi(q)} + O\Big(\frac{N}{\log^C N}\Big), \qquad (20.22)$$

$(r, q) = 1$. Thus, we can take C enormous, large enough so that even when we multiply by the length of the Major Arcs (of size $\frac{\log^{3B} N}{N}$, we still have something small.

20.2 $f_N(x)$ and $u(x)$

20.2.1 $f\left(\frac{a}{q}\right)$

We now calculate $f_N\left(\frac{a}{q}\right)$ for $q \le \log^B N$.

Up to lower order terms,

$$\begin{aligned}
f_N\Big(\frac{a}{q}\Big) &= \sum_{p_1 \le N} \log p_1 \cdot e\Big(p_1\frac{a}{q}\Big) \sum_{p_2 \le N} \log p_2 \cdot e\Big(-2p_2\frac{a}{q}\Big) \\
&= \sum_{r_1=1}^{q} \sum_{\substack{p_1 \le N \\ p_1 \equiv r_1 (q)}} \log p_1 \cdot e\Big(p_1\frac{a}{q}\Big) \sum_{r_2=1}^{q} \sum_{\substack{p_2 \le N \\ p_2 \equiv r_1 (q)}} \log p_2 \cdot e\Big(-2p_2\frac{a}{q}\Big) \\
&= \sum_{r_1=1}^{q} e\Big(r_1\frac{a}{q}\Big) \sum_{r_2=1}^{q} e\Big(r_2\frac{-2a}{q}\Big) \sum_{\substack{p_1 \le N \\ p_1 \equiv r_1 (q)}} \log p_1 \sum_{\substack{p_2 \le N \\ p_2 \equiv r_2 (q)}} \log p_2 \\
&= \frac{N^2}{\phi^2(q)} \sum_{\substack{r_1=1 \\ (r_1,q)=1}}^{q} e\Big(r_1\frac{a}{q}\Big) \sum_{\substack{r_2=1 \\ (r_2,q)=1}}^{q} e\Big(r_2\frac{-2a}{q}\Big) \\
&= \frac{N^2}{\phi^2(q)} c_q(a) c_q(-2a), \qquad (20.23)
\end{aligned}$$

where the second to last line follows from the Siegel-Wal£sz Theorem (Theorem 20.1.3). We restrict to $(r_i, q) = 1$ because if $(r_i, q) > 1$, there is at most one prime $p_i \equiv r_i \bmod q$.

20.2.2 $u(x)$

Let

$$u(x) = \sum_{m_1 \le N} \sum_{m_2 \le N} e\Big((m_1 - 2m_2)x\Big). \qquad (20.24)$$

We will often look at

$$\frac{c_q(a)c_q(-2a)}{\phi^2(q)}u(x). \tag{20.25}$$

Note

$$u(0) \;=\; N^2. \tag{20.26}$$

20.3 $f_N(\alpha) - \frac{c_q(a)c_q(-2a)}{\phi^2(q)}u(\alpha - \frac{a}{q})$, $\alpha \in \mathcal{M}_{a,q}$

Let

$$C_q(a) \;=\; \frac{c_q(a)c_q(-2a)}{\phi^2(q)}. \tag{20.27}$$

We write α as $\beta + \frac{a}{q}$, $\beta \in \Big[-\frac{Q}{N}, \frac{Q}{N}\Big]$, $Q = \log^B N$. As always, we ignore lower order terms.

Note $f_N(x)$ is approximately $C_q(a)N^2$ for x near $\frac{a}{q}$. We now expand and show $f_N(\alpha)$ is $C_q(a)u\Big(\alpha - \frac{a}{q}\Big)$ plus errors of size $\frac{N^2}{\log^{C-2B} N}$ for $\alpha \in \mathcal{M}_{a,q}$.

20.3.1 Setup

$$\begin{aligned}
S_{a,q}(\alpha) &= f_N(\alpha) - C_q(a)u\Big(\alpha - \frac{a}{q}\Big) \\
&= \sum_{m_1,m_2 \le N} \lambda(m_1)\lambda(m_2)e\Big((m_1 - 2m_2)\alpha\Big) \;-\; C_q(a) \sum_{m_1,m_2 \le N} e\Big((m_1 - 2m_2)\beta\Big) \\
&= \sum_{m_1,m_2 \le N} \left[\lambda(m_1)\lambda(m_2)e\Big((m_1 - 2m_2)\frac{a}{q}\Big) - C_q(a)\right] e\Big((m_1 - 2m_2)\beta\Big) \\
&= \sum_{m_1 \le N} \left[\sum_{m_2 \le N} \Big[\lambda(m_1)\lambda(m_2)e\Big((m_1 - 2m_2)\frac{a}{q}\Big) - C_q(a)\Big] e(-2m_2\beta)\right] e(m_1\beta)
\end{aligned} \tag{20.28}$$

We now apply Partial Summation multiple times. First, we apply Partial Summation to the m_2-sum:

$$\begin{aligned}
S_{2;a,q} &= \sum_{m_2 \le N} \Big[\lambda(m_1)\lambda(m_2)e\Big((m_1 - 2m_2)\frac{a}{q}\Big) - C_q(a)\Big] e(-2m_2\beta) \\
&= \sum_{m_2 \le N} a_{m_2} b_{m_2} \\
&= A_2(N)e(-2N\beta) + 4\pi i\beta \int_0^N \sum_{m_2 \le u} a_{m_2} e(-u\beta) du. \qquad (20.29)
\end{aligned}$$

We hit the above with $e(m_1\beta)$, and sum from $m_1 = 1$ to N. We get two pieces:

$$\begin{aligned}
S_{1\sum;a,q} &= \sum_{m_1 \le N} A_2(N)e(-2N\beta) \cdot e(m_1\beta) \\
S_{1\int;a,q} &= \sum_{m_1 \le N} 4\pi i\beta \int_0^N \sum_{m_2 \le u} a_{m_2} e(-u\beta) du \cdot e(m_1\beta) \\
S_{a,q} &= S_{1\sum;a,q} + S_{1\int;a,q}. \qquad (20.30)
\end{aligned}$$

20.3.2 $S_{1\sum;a,q}$

$$\begin{aligned}
S_{1\sum;a,q} &= \sum_{m_1 \le N} A_2(N)e(-2N\beta) \cdot e(m_1\beta) \\
&= e(-2N\beta) \sum_{m_1 \le N} A_2(N)e(m_1\beta) \\
&= e(-2N\beta) \sum_{m_1 \le N} \sum_{m_2 \le N} \Big[\lambda(m_1)\lambda(m_2)e\Big((m_1 - 2m_2)\frac{a}{q}\Big) - C_q(a)\Big] e(m_1\beta) \\
&= e(-2N\beta) \Bigg[A_1(N)e(N\beta) \\
&\quad -2\pi i\beta \int_0^N \sum_{m_1 \le t} \sum_{m_2 \le N} \Big[\lambda(m_1)\lambda(m_2)e\Big((m_1 - 2m_2)\frac{a}{q}\Big) - C_q(a)\Big] e(t\beta) dt. \qquad (20.31)
\end{aligned}$$

First Piece

The £rst piece, the $A_1(N)e(N\beta)$ term, is small for $q \le Q$. Why? We have (up to lower order terms)

$$\begin{aligned} A_1(N)e(N\beta) &= \sum_{m_1,m_2\le N} \lambda(m_1)\lambda(m_2)e\Big((m_1-2m_2)\frac{a}{q}\Big) - \sum_{m_1,m_2\le N} C_q(a) \\ &= C_q(a)N^2 - N^2C_q(a) \;=\; 0. \end{aligned} \tag{20.32}$$

Thus, because of our choice of functions, the leading terms vanish, and the remaining term is small.

Second Piece

We now study the second piece. Note $|\beta| \le \frac{Q}{N} = \frac{\log^2 B}{N}$, and $C_q(a) = \frac{c_q(a)}{\phi^2(q)}\frac{c_q(-2a)}{\phi^2(q)}$.

Up to lower order terms, the m_2-sum will leave us with

$$\beta\frac{c_q(-2a)N}{\phi(q)}\int_0^N \sum_{m_1\le t}\left[\lambda(m_1)e\Big(m_1\frac{a}{q}\Big) - \frac{c_q(a)}{\phi(q)}\right]e(t\beta)dt. \tag{20.33}$$

Note $f_N(x)$ is a multiple of N^2 for x near $\frac{a}{q}$. Thus, we want to make sure the above is well dominated by N^2.

For $t \le \sqrt{N}$, this is immediate. For $t \ge \sqrt{N}$, using Siegel-Wal£sz (Theorem 20.1.3), we can make the bracketed quantity in the integrant dominated by $\frac{N}{\log^C N}$ for any C when $q \le \log^B N$. Thus, we integrate a quantity that is at most $\frac{N}{\log^C N}$ over an interval of length N, we multiply by $N\beta \ll Q = \log^B N$.

Thus, choosing C appropriately, the integral contributes $\frac{N^2}{\log^{C-B} N}$, and hence is negligible.

Remark 20.3.1. *Note, of course, that the contribution is only negligible while* $|\beta| \le \frac{Q}{N}$.

Lemma 20.3.2. $S_{1\sum;a,q}$ *is a lower order correction.*

20.3.3 $S_{1\int;a,q}$

We must evaluate

$$S_{1\int;a,q} \;=\; \sum_{m_1\le N} 4\pi i\beta \int_0^N \sum_{m_2\le u} a_{m_2} e(-u\beta)du \cdot e(m_1\beta), \qquad (20.34)$$

where

$$a_{m_2} \;=\; \Big[\lambda(m_1)\lambda(m_2)e\Big((m_1-2m_2)\frac{a}{q}\Big) - C_q(a)\Big]. \qquad (20.35)$$

We bring the sum over m_1 inside the integral and again use Partial Summation.

We will ignore the integration and β for now, as these will contribute $\beta N \ll Q = \log^B N$ times the maximum value of the integrand. We will leave the $e(-u\beta)du$ with this integration.

When $u \le \sqrt{N}$, we can immediately show the above is a lower order correction. Thus, below we always assume $u \ge \sqrt{N}$.

First Piece

We have

$$\begin{aligned}
S_{1\int\sum;a,q} &= \sum_{\substack{m_1\le N\\ m_2\le u}} \Big[\lambda(m_1)\lambda(m_2)e\Big((m_1-2m_2)\frac{a}{q}\Big) - C_q(a)\Big]e(N\beta) \\
&= e(N\beta)\Bigg[\sum_{\substack{m_1\le N\\ m_2\le u}} \lambda(m_1)\lambda(m_2)e\Big((m_1-2m_2)\frac{a}{q}\Big) - C_q(a)\sum_{\substack{m_1\le N\\ m_2\le u}} 1\Bigg]. \\
&= e(N\beta)\Bigg[C_q(a)uN - C_q(a)uN + \text{Lower Order Terms}\Bigg], \qquad (20.36)
\end{aligned}$$

where by the Siegel-Wal£sz Theorem (Theorem 20.1.3), the error in the bracketed quantity is of size $\frac{uN}{\log^C N}$.

We then integrate from $u = \sqrt{N}$ to N and multiply by β, giving a contribution bounded by

$$\beta N \cdot \frac{N^2}{\log^C N} \;\ll\; \frac{\log^B}{N}\frac{N^3}{\log^C N} \;\ll\; \frac{N^2}{\log^{C-B} N}, \qquad (20.37)$$

again getting a lower order correction to $f_N(x)$ for x near $\frac{a}{q}$ (remember $f_N(x)$ is of size N^2).

Second Piece

Again, $u \geq \sqrt{N}$, and we have

$$2\pi i\beta \int_0^N \sum_{m_1 \leq t} \left[\sum_{m_2 \leq u} \left[\lambda(m_1)\lambda(m_2) e\Big((m_1 - 2m_2)\frac{a}{q}\Big) - C_q(a) \right] \right] e(t\beta) dt. \tag{20.38}$$

Again, for $t \leq \sqrt{N}$, the contribution will be a lower order correction. For $t, u \geq \sqrt{N}$,

Again, executing the sum over m_1 and m_2 will give us

$$C_q(a)ut - C_q(a)ut + \text{Lower Order Terms}, \tag{20.39}$$

with the lower order terms of size $\frac{ut}{\log^C N}$.

Integrating over t (from $\sqrt{N}$ to N), then integrating over u (from $\sqrt{N}$ to N) and then multiplying by β^2 gives an error bounded by

$$\beta^2 N^2 \cdot \frac{N^2}{\log^C N} \ll \frac{\log^{2B} N}{N^2} \frac{N^4}{\log^C N} \ll \frac{N^2}{\log^{C-2B} N}, \tag{20.40}$$

again a lower order correction.

20.4 Integrals of $u(x)$

20.4.1 Formulations

Remember

$$u(x) = \sum_{m_1, m_2 \leq N} e\Big((m_1 - 2m_2)x\Big). \tag{20.41}$$

We need to study $\int_{-\frac{1}{2}}^{\frac{1}{2}} f_N(x) e(-x) dx$. We have shown that

$$f_N(\alpha) = C_q(a) u\Big(\alpha - \frac{a}{q}\Big) + O\Big(\frac{N^2}{\log^{C-2B} N}\Big), \ \ \alpha \in \mathcal{M}_{a,q}. \tag{20.42}$$

Thus, we must evaluate

$$\begin{aligned}
\int_{\mathcal{M}_{a,q}} u\Big(\alpha - \frac{a}{q}\Big) \cdot e(-\alpha) d\alpha &= \int_{\frac{a}{q}-\frac{Q}{N}}^{\frac{a}{q}+\frac{Q}{N}} u\Big(\alpha - \frac{a}{q}\Big) \cdot e(-\alpha) d\alpha \\
&= \int_{-\frac{Q}{N}}^{\frac{Q}{N}} u(\beta) \cdot e\Big(-\frac{q}{q} - \beta\Big) d\beta \\
&= e\Big(-\frac{a}{q}\Big) \int_{-\frac{Q}{N}}^{\frac{Q}{N}} u(\beta) e(-\beta) d\beta. \qquad (20.43)
\end{aligned}$$

20.4.2 $\int_{-\frac{1}{2}}^{\frac{1}{2}} u(x)e(-x)dx$

$$\begin{aligned}
\int_{-\frac{1}{2}}^{\frac{1}{2}} u(x)e(-x)dx &= \int_{-\frac{1}{2}}^{\frac{1}{2}} \sum_{m_1,m_2 \le N} e\Big((m_1 - 2m_2)x\Big) \cdot e(-x)dx \\
&= \sum_{m_1,m_2 \le N} \int_{-\frac{1}{2}}^{\frac{1}{2}} e\Big((m_1 - 2m_2 - 1)x\Big) dx. \qquad (20.44)
\end{aligned}$$

If $m_1 - 2m_2 - 1 = 0$, the integral gives 1. There are approximately $\frac{N}{2}$ ways to choose $m_1, m_2 \le N$ such that $m_1 - 2m_2 - 1 = 0$.

Assume now $m_1 - 2m_2 - 1 \neq 0$. Then the integral vanishes.

Hence,

Lemma 20.4.1.

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} u(x)e(-x)dx = \frac{N}{2} + O(1). \qquad (20.45)$$

20.4.3 $\int_{-\frac{1}{2}}^{-\frac{Q}{N}} + \int_{\frac{Q}{N}}^{\frac{1}{2}} u(x)e(-x)dx$

De£ne

$$\begin{aligned} I_1 &= \left[-\frac{1}{2}, -\frac{1}{2} + \frac{Q}{N}\right] \\ I_2 &= \left[-\frac{1}{2} + \frac{Q}{N}, -\frac{Q}{N}\right] \\ I_3 &= \left[\frac{Q}{N}, \frac{1}{2} - \frac{Q}{N}\right] \\ I_4 &= \left[\frac{1}{2} - \frac{Q}{N}, \frac{1}{2}\right] \\ I &= I_1 \cup I_2 \cup I_3 \cup I_4. \end{aligned} \tag{20.46}$$

20.4.4 Integral over I_2, I_3

We have

$$\begin{aligned} \int_{I_i} u(x)e(-x)dx &= \int_{I_i} \sum_{m_1, m_2 \leq N} e\Big((m_1 - 2m_2 - 1)x\Big)dx \\ &= \int_{I_i} \sum_{m_1 \leq N} e(m_1 x) \sum_{m_2 \leq N} e(-2m_2 x) \cdot e(-x)dx \\ &= \int_{I_i} \frac{e(x) - e((N+1)x)}{1 - e(x)} \frac{e(-2x) - e(-2(N+1)x)}{1 - e(-2x)} e(-x)dx. \end{aligned} \tag{20.47}$$

On I_2 and I_3, the integral is

$$\ll \int_{I_i} \frac{2}{x}\frac{2}{x} dx \ll \frac{N}{Q} = \frac{N}{\log^B N}, \tag{20.48}$$

see, for example, Nathanson (Additive Number Theory: The Classical Bases, Chapter 8).

20.4.5 Integral over I_1, I_4

Each of these intervals has length $\frac{Q}{N} = \frac{\log^B N}{N}$. There are $\frac{N}{2} + O(1)$ pairs such that $m_1 - 2m_2 - 1 = 0$. Each of these pairs will contribute (bound the integrand by 1) $\frac{Q}{N}$. As there are at most $\frac{N}{2}$ pairs, these contribute at most $\frac{N}{2}\frac{Q}{N} \ll \log^B N$.

Henceforth we assume $m_1 - 2m_2 - 1 \neq 0$. We write

$$I_1 \cup I_4 \;=\; \left[\frac{1}{2} - \frac{Q}{N}, \frac{1}{2} + \frac{Q}{N}\right] \;=\; I'. \tag{20.49}$$

We have

$$\begin{aligned} & \sum_{\substack{m_1, m_2 \le N \\ m_1 - 2m_2 - 1 \ne 0}} \int_{I'} e\Big((m_1 - 2m_2 - 1)x\Big) dx \\ = \quad & e\Big(-\frac{1}{2}\Big) \sum_{\substack{m_1, m_2 \le N \\ m_1 - 2m_2 - 1 \ne 0}} (-1)^{m_1} \int_{-\frac{Q}{N}}^{\frac{Q}{N}} e\Big((m_1 - 2m_2 - 1)x\Big) dx \\ = \quad & e\Big(-\frac{1}{2}\Big) \frac{1}{2\pi i} \sum_{\substack{m_1, m_2 \le N \\ m_1 - 2m_2 - 1 \ne 0}} (-1)^{m_1} \frac{2\sin\Big((m_1 - 2m_2 - 1)\frac{Q}{N}\Big)}{m_1 - 2m_2 - 1}, \end{aligned} \tag{20.50}$$

because, changing variables by sending x to $(x - \frac{1}{2}) + \frac{1}{2}$ gives factors of $e\Big((m_1 - 2m_2 - 1)\frac{1}{2}\Big) = e(-\frac{1}{2})e(\frac{m_1}{2})e(-m_2)$, and $e(\frac{m_1}{2}) = (-1)^{m_1}$.

$0 < |m_1 - 2m_2 - 1| \le N^{1-\epsilon}$

Let $w = m_1 - 2m_2 - 1$. We will do the case $0 < w \le N^{1-\epsilon}$, the case with $-N^{1-\epsilon} > w > 0$ being handled similarly.

For each w, there are at most N pairs of m_1, m_2 giving rise to such a w. For such w, $\frac{\sin(w\frac{Q}{N})}{w} \ll \frac{Q}{N}$ (because we are taking the sin of a quantity very close to zero).

Thus, these pairs contribute at most

$$\ll \; N \cdot \frac{Q}{N} \; \ll \; Q \;=\; \log^B N. \tag{20.51}$$

Inserting absolute values in Equation 20.50 gives a contribution of at most $\log^B N$ for such w, $0 < w < N^{1-\epsilon}$.

$N^{1-\epsilon} < |m_1 - 2m_2 - 1| \le N$

Again, let $w = m_1 - 2m_2 - 1$ and assume $N^{1-\epsilon} < |w| \le N$. We will only consider $w > 0$; $w < 0$ is handled similarly.

The cancellation is due to the presence of the factor $(-1)^{m_1}$; note that for the pair (m_1, m_2) we only care about the parity of m_1.

Consider w and $w-1$.

For $m_1 - 2m_2 - 1 = w$, the solutions are

$$\begin{aligned} m_1 &= w+3, \qquad m_2 = 1 \\ m_1 &= w+5, \qquad m_2 = 2 \\ m_1 &= w+7, \qquad m_2 = 3 \end{aligned} \tag{20.52}$$

and so on; thus there are about $\frac{N-w}{2}$ pairs, all with parity $-(-1)^w$.

For $m_1 - 2m_2 - 1 = w-1$, we again have about $\frac{N-w}{2}$ pairs, but now the parity is $(-1)^w$. Thus, each of the $\frac{N-w}{2}$ pairs with $m_1 - 2m_2 - 1 = w$ is matched with one of the $\frac{N-w}{2}$ pairs with $m_1 - 2m_2 - 1 = w - 1$, and we are off by at most $O(1)$ pairs, which will contribute

$$\ll \sum_{w=N^{1-\epsilon}}^{N} \frac{1}{w} \ll \log N. \tag{20.53}$$

For the remaining terms, we subtract in pairs, using the £rst order Taylor Expansion of $\sin(x)$. We have

$$\sum_{w=N^{1-\epsilon}}^{N} \left[\frac{\sin\left(w\frac{Q}{N}\right)}{w} - \frac{\sin\left(w\frac{Q}{N} - \frac{Q}{N}\right)}{w-1} \right]. \tag{20.54}$$

The Main Term of the Taylor Expansion gives $\ll \frac{1}{w^2}$, which when summed over w gives $\frac{1}{N^{1-\epsilon}}$. As we have about $\frac{N-w}{2} \ll N$ pairs for each w, this contributes at most $N \cdot \frac{1}{N^{1-\epsilon}} \ll N^{\epsilon}$.

We also have the £rst order term from the Taylor Expansion:

$$\sin\left(w\frac{Q}{N} - \frac{Q}{N}\right) = \sin\left(w\frac{Q}{N}\right) + O\left(\frac{Q}{N}\right). \tag{20.55}$$

This error leads to (remembering there are $\frac{N-w}{2} \ll N$ pairs for each w)

$$\ll N \sum_{w=N^{1-\epsilon}}^{N} \frac{\frac{Q}{N}}{w-1} \ll Q \log N^{\epsilon} \ll \log^{B+1} N. \tag{20.56}$$

20.4.6 Collecting the Pieces

We have shown

$$
\begin{aligned}
\int_{[-\frac{1}{2},\frac{1}{2}]} u(x)e(-x)dx &= \frac{N}{2} + O(1) \\
\int_{[-\frac{1}{2},\frac{1}{2}]-[-\frac{Q}{N},\frac{Q}{N}]} u(x)e(-x)dx &= O\Big(\frac{N}{\log^B N}\Big).
\end{aligned} \tag{20.57}
$$

Therefore

Lemma 20.4.2.

$$
\int_{-\frac{Q}{N}}^{\frac{Q}{N}} u(x)e(-x)dx = \frac{N}{2} + O\Big(\frac{N}{\log^B N}\Big). \tag{20.58}
$$

Remembering that we had

$$
\begin{aligned}
\int_{\mathcal{M}_{a,q}} u\Big(\alpha - \frac{a}{q}\Big) \cdot e(-\alpha)d\alpha &= \int_{\frac{a}{q}-\frac{Q}{N}}^{\frac{a}{q}+\frac{Q}{N}} u\Big(\alpha - \frac{a}{q}\Big) \cdot e(-\alpha)d\alpha \\
&= \int_{-\frac{Q}{N}}^{\frac{Q}{N}} u(\beta) \cdot e\Big(-\frac{q}{q} - \beta\Big)d\beta \\
&= e\Big(-\frac{a}{q}\Big) \int_{-\frac{Q}{N}}^{\frac{Q}{N}} u(\beta)e(-\beta)d\beta,
\end{aligned} \tag{20.59}
$$

we see that

Lemma 20.4.3.

$$
\int_{\mathcal{M}_{a,q}} u\Big(\alpha - \frac{a}{q}\Big) \cdot e(-\alpha)d\alpha = e\Big(-\frac{a}{q}\Big)\frac{N}{2}. \tag{20.60}
$$

20.5 Determination of the Main Term

We now calculate the contribution from the Major Arcs. Up to lower order terms,

$$
\begin{aligned}
\int_{\mathcal{M}} f_N(x)e(-x)dx &= \sum_{q\le Q}\sum_{\substack{a=1\\(a,q)=1}}^{q}\int_{\frac{a}{q}-\frac{Q}{N}}^{\frac{a}{q}+\frac{Q}{N}} f_N(\alpha)e(-\alpha)d\alpha \\
&= \sum_{q\le Q}\sum_{\substack{a=1\\(a,q)=1}}^{q}\int_{\frac{a}{q}-\frac{Q}{N}}^{\frac{a}{q}+\frac{Q}{N}} C_q(a)u\Big(\alpha-\frac{a}{q}\Big)e(-\alpha)d\alpha \\
&= \sum_{q\le Q}\sum_{\substack{a=1\\(a,q)=1}}^{q} e\Big(-\frac{a}{q}\Big)\int_{-\frac{Q}{N}}^{\frac{Q}{N}} C_q(a)u(\beta)e(-\beta) \\
&= \sum_{q\le Q}\sum_{\substack{a=1\\(a,q)=1}}^{q} C_q(a)e\Big(-\frac{a}{q}\Big)\frac{N}{2} \\
&= \frac{N}{2}\sum_{q\le Q}\sum_{\substack{a=1\\(a,q)=1}}^{q}\frac{c_q(a)c_q(-2a)}{\phi^2(q)}\cdot e\Big(-\frac{a}{q}\Big) \\
&= \frac{N}{2}\sum_{q=1}^{Q}\Bigg[\sum_{\substack{a=1\\(a,q)=1}}^{q} C_q(a)e\Big(-\frac{a}{q}\Big)\Bigg] \\
&= \frac{N}{2}\sum_{q=1}^{Q}\rho_q \\
&= \mathfrak{S}_N\frac{N}{2}, \qquad (20.61)
\end{aligned}
$$

where we have de£ned

$$\begin{aligned}
c_q(a) &= \sum_{\substack{r=1 \\ (r,q)=1}}^{q} e\Big(r\frac{a}{q}\Big) \\
C_q(a) &= \frac{c_q(a)c_q(-2a)}{\phi^2(q)} \\
\rho_q &= \sum_{\substack{a=1 \\ (a,q)=1}}^{q} C_q(a)e\Big(-\frac{a}{q}\Big) \\
\mathfrak{S}_N &= \sum_{\substack{a=1 \\ (a,q)=1}}^{q} \rho_q. \qquad (20.62)
\end{aligned}$$

20.5.1 Properties of $C_q(a)$ and ρ_q

We will follow the presentation of Nathanson (Additive Number Theory: The Classical Bases, Chapter 8 and Appendix A).

$c_q(a)$ is Multiplicative

We follow Nathanson, Pages $320-321$, Theorem $A.23$. Note that we are labeling by r what he labels a, and we are labeling by a what he labels n.

Lemma 20.5.1. *$c_q(a)$ is multiplicative; ie, if $(q,q')=1$, then $c_{qq'}(a)=c_q(a)c_{q'}(a)$.*

Proof: We have

$$\sum_{\substack{\widetilde{r}=1 \\ (\widetilde{r},qq')=1}}^{qq'} e\Big(\widetilde{r}\frac{a}{qq'}\Big). \qquad (20.63)$$

Exercise 20.5.2. *Show that we can write the $\widetilde{r}$s above as $\widetilde{r} \equiv rq'+r'q$ mod qq', where $1 \leq r \leq q$, $1 \leq r' \leq q'$, and $(r,q)=(r',q')=1$.*

Thus

$$
\begin{aligned}
c_q(a)c_{q'}(a) &= \sum_{\substack{r=1 \\ (r,q)=1}}^{q} e\Big(r\frac{a}{q}\Big) \sum_{\substack{r'=1 \\ (r',q')=1}}^{q'} e\Big(r'\frac{a}{q'}\Big) \\
&= \sum_{\substack{r=1 \\ (r,q)=1}}^{q} \sum_{\substack{r'=1 \\ (r',q')=1}}^{q'} e\Big(\frac{(rq'+r'q)a}{qq'}\Big) \\
&= \sum_{\substack{\widetilde{r}=1 \\ (\widetilde{r},qq')=1}}^{qq'} e\Big(\widetilde{r}\frac{a}{q}\Big) \;=\; c_{qq'}(a). \qquad (20.64)
\end{aligned}
$$

$c_q(a)$ **for** $(a, q) = 1$

Exercise 20.5.3. *Show that*

$$
h_d(a) \;=\; \sum_{r=1}^{d} e\Big(r\frac{a}{d}\Big) \;=\; \begin{cases} d & \text{if } d|a \\ 0 & \text{otherwise} \end{cases} \qquad (20.65)
$$

Recall the moebius function:

$$
\mu(d) \;=\; \begin{cases} (-1)^r & \text{if } d \text{ is the product of } r \text{ distinct primes} \\ 0 & \text{otherwise} \end{cases} \qquad (20.66)
$$

Exercise 20.5.4. *Prove*

$$
\sum_{d|(r,q)} \mu(d) \;=\; \begin{cases} 1 & \text{if } (r,q)=1 \\ 0 & \text{otherwise} \end{cases} \qquad (20.67)
$$

Then

$$
\begin{aligned}
c_q(a) &= \sum_{\substack{r=1\\(r,q)=1}}^{q} e\left(r\frac{a}{q}\right) \\
&= \sum_{r=1}^{q} e\left(r\frac{a}{q}\right) \sum_{d|(r,q)} \mu(d) \\
&= \sum_{d|q} \mu(d) \sum_{\substack{r=1\\d|r}}^{q} e\left(r\frac{a}{q}\right) \\
&= \sum_{d|q} \mu(d) \sum_{l=1}^{\frac{q}{d}} e\left(l\frac{a}{\frac{q}{d}}\right) \\
&= \sum_{d|q} \mu(d) h_{\frac{q}{d}}(a) \\
&= \sum_{d|q} \mu\left(\frac{q}{d}\right) h_d(a) \\
&= \sum_{\substack{d|q\\d|a}} \mu\left(\frac{q}{d}\right) \cdot d \\
&= \sum_{d|(a,q)} \mu\left(\frac{q}{d}\right) d. \qquad (20.68)
\end{aligned}
$$

Note that if $(a,q)=1$, then there is only one term above, namely $d=1$, which yields

$$c_q(a) \;=\; \mu(q) \;\text{ if } (a,q)=1. \qquad (20.69)$$

Corollary 20.5.5. *If* $q=p^k$, $k \geq 2$ *and* $(a,q)=1$, *then* $c_q(a)=0$.

$C_q(a)$ is Multiplicative

We have shown $c_{qq'}(a) = c_q(a)c_{q'}(a)$ if $(q,q')=1$. Recall the Euler phi-function, $\phi(q)$, is the number of numbers less than q which are relatively prime to q.

Exercise 20.5.6. *Prove that* $\phi(q)$ *is multiplicative; ie, if* $(q,q')=1$, *then* $\phi(qq') = \phi(q)\phi(q')$.

We now have

Lemma 20.5.7. *$C_q(a)$ is multiplicative.*

Proof: Assume $(q, q') = 1$. We have

$$\begin{aligned} C_{qq'}(a) &= \frac{c_{qq'}(a)c_{qq'}(-2a)}{\phi^2(qq')} \\ &= \frac{c_q(a)c_{q'}(a)c_q(-2a)c_{q'}(-2a)}{\phi^2(q)\phi^2(q')} \\ &= \frac{c_q(a)c_q(-2a)}{\phi^2(q)} \cdot \frac{c_{q'}(a)c_{q'}(-2a)}{\phi^2(q')} \\ &= C_q(a)C_{q'}(a). \end{aligned} \tag{20.70}$$

ρ_q is Multiplicative

We £rst prove a needed lemma.

Lemma 20.5.8. *Consider $C_{q_1}(a_1 q_2)$. Then*

$$C_{q_1}(a_1 q_2) = C_{q_1}(a_1) \tag{20.71}$$

if $(q_1, q_2) = 1$.

Proof:

$$\begin{aligned} C_{q_1}(a_1 q_2) &= \sum_{\substack{r_1=1 \\ (r_1,q_1)=1}}^{q_1} e\left(r_1 \frac{a_1 q_2}{q_1}\right) \\ &= \sum_{\substack{r_1=1 \\ (r_1,q_1)=1}}^{q_1} e\left(r_1 q_2 \frac{a_1}{q_1}\right) \\ &= \sum_{\substack{r=1 \\ (r,q_1)=1}}^{q_1} e\left(r \frac{a_1}{q_1}\right) = C_{q_1}(a), \end{aligned} \tag{20.72}$$

because $(q_1, q_2) = 1$ implies that as r_1 goes through all residue classes that are relatively prime to q_1, so too does $r = r_1 q_2$. $\Box$

Lemma 20.5.9. *ρ_q is multiplicative.*

Recall

$$\rho_q \ = \ \sum_{\substack{a=1 \\ (a,q)=1}}^{q} C_q(a) e\Big(-\frac{a}{q}\Big). \tag{20.73}$$

Assume $(q_1, q_2) = 1$. Then we can write the congruence classes mod $q_1 q_2$ as $a_1 q_2 + a_2 q_1$, with $1 \leq a_1 \leq q_1$, $1 \leq a_2 \leq q_2$ and $(a_1, q_1) = (a_2, q_2) = 1$.

$$\begin{aligned}
\rho_{q_1 q_2} &= \sum_{\substack{a=1 \\ (a,q_1q_2)=1}}^{q_1q_2} C_{q_1q_2}(a) e\Big(-\frac{a}{q_1q_2}\Big) \\
&= \sum_{\substack{a=1 \\ (a,q_1q_2)=1}}^{q_1q_2} C_{q_1}(a) C_{q_2}(a) e\Big(-\frac{a}{q_1q_2}\Big) \\
&= \sum_{\substack{a_1=1 \\ (a_1,q_1)=1}}^{q_1} \sum_{\substack{a_2=1 \\ (a_2,q_2)=1}}^{q_2} C_{q_1}(a_1q_2 + a_2q_1) C_{q_2}(a_1q_2 + a_2q_1) e\Big(-\frac{a_1q_2 + a_2q_1}{q_1q_2}\Big).
\end{aligned} \tag{20.74}$$

Exercise 20.5.10. *With a_1, a_2, q_1, q_2 as above,*

$$C_{q_1}(a_1q_2 + a_2q_1) \ = \ C_{q_1}(a_1q_2) \ \textit{and} \ C_{q_2}(a_1q_2 + a_2q_1) \ = \ C_{q_2}(a_2q_1). \tag{20.75}$$

Thus, we have

$$\begin{aligned}
\rho_{q_1 q_2} &= \sum_{\substack{a_1=1 \\ (a_1,q_1)=1}}^{q_1} \sum_{\substack{a_2=1 \\ (a_2,q_2)=1}}^{q_2} C_{q_1}(a_1q_2) C_{q_2}(a_2q_1) e\Big(-\frac{a_1q_2 + a_2q_1}{q_1q_2}\Big) \\
&= \sum_{\substack{a_1=1 \\ (a_1,q_1)=1}}^{q_1} C_{q_1}(a_1q_2) e\Big(-\frac{a_1}{q_1}\Big) \sum_{\substack{a_2=1 \\ (a_2,q_2)=1}}^{q_2} C_{q_2}(a_2q_1) e\Big(-\frac{a_2}{q_2}\Big) \\
&= \sum_{\substack{a_1=1 \\ (a_1,q_1)=1}}^{q_1} C_{q_1}(a_1) e\Big(-\frac{a_1}{q_1}\Big) \sum_{\substack{a_2=1 \\ (a_2,q_2)=1}}^{q_2} C_{q_2}(a_2) e\Big(-\frac{a_2}{q_2}\Big) \\
&= \rho_{q_1} \cdot \rho_{q_2}.
\end{aligned} \tag{20.76}$$

Thus, ρ_q is multiplicative. $\square$

Calculation of ρ_q

Lemma 20.5.11. *$\rho_{p^k} = 0$ if $k \geq 2$ and p is a prime.*

Proof: This follows immediately from $C_{p^k}(a) = 0$. $\Box$

Lemma 20.5.12. *If $p > 2$ is prime, $\rho_p = -\frac{1}{(p-1)^2}$.*

Proof:

$$\begin{aligned} \rho_p &= \sum_{\substack{a=1 \\ (a,p)=1}}^{p} C_p(a) e\Big(-\frac{a}{p}\Big) \\ &= \sum_{a=1}^{p-1} \frac{c_p(a)c_p(-2a)}{\phi^2(p)} e\Big(-\frac{a}{p}\Big). \end{aligned} \tag{20.77}$$

But as $p > 2$, $c_p(a) = c_p(-2a) = \mu(p)$ as $(a,p) = 1$. As $\mu^2(p) = 1$ and $\phi(p) = p - 1$ we have

$$\begin{aligned} \rho_p &= \sum_{a=1}^{p-1} \frac{1}{(p-1)^2} e\Big(-\frac{a}{p}\Big) \\ &= \frac{1}{(p-1)^2}\left[-e\Big(-\frac{0}{p}\Big) + \sum_{a=0}^{p-1} e\Big(-\frac{a}{p}\Big)\right] \\ &= -\frac{1}{(p-1)^2}. \end{aligned} \tag{20.78}$$

Lemma 20.5.13. *If $p = 2$, then $\rho_2 = 1$.*

Proof:

$$\begin{aligned} \rho_2 &= \sum_{\substack{a=1 \\ (a,2)=1}}^{2} C_2(a) e\Big(-\frac{a}{2}\Big) \\ &= C_2(1) e\Big(-\frac{1}{2}\Big) \\ &= \frac{c_2(1)c_2(-2)}{\phi^2(2)} \cdot e^{-\pi i} \\ &= \frac{e^{\pi i} e^{-2\pi i}}{1^2} \cdot e^{-\pi i} \;=\; 1, \end{aligned} \tag{20.79}$$

where we have used $c_2(1) = e^{\pi i}$ and $c_2(-2) = e^{-2\pi i}$.

Exercise 20.5.14. *Prove $c_2(1) = e^{\pi i}$ and $c_2(-2) = e^{-2\pi i}$.*

20.5.2 Determination of $\mathfrak{S}_N$ and $\mathfrak{S}$

Recall

$$\mathfrak{S}_N \;=\; \sum_{q \le Q} \rho_q. \tag{20.80}$$

We de£ne

$$\mathfrak{S} \;=\; \sum_{q} \rho_q. \tag{20.81}$$

Exercise 20.5.15. *Let h_q be any multiplicative sequence (with whatever growth conditions are necessary to ensure the convergence of all sums below). Then*

$$\sum_q h_q \;=\; \prod_{p \text{ prime}} \Big(1 + \sum_{k=1}^{\infty} h_{p^k}\Big). \tag{20.82}$$

$\mathfrak{S}$

We have

$$\begin{aligned} \mathfrak{S} &= \sum_q \rho_q \\ &= \prod_{p \text{ prime}} \Big(1 + \sum_{k=1}^{\infty} \rho_{p^k}\Big) \\ &= \prod_{p} \Big(1 + \rho_p\Big) \end{aligned} \tag{20.83}$$

because $\rho_{p^k} = 0$ for $k \ge 2$ and p prime by Lemma 20.5.11. We have previously shown (see Lemmas 20.5.12 and 20.5.13) that $\rho_2 = 1$ and $\rho_p = -\frac{1}{(p-1)}$ for $p > 2$ prime. Therefore

$$
\begin{aligned}
\mathfrak{S} &= \prod_p \Big(1+\rho_p\Big) \\
&= (1+\rho_2)\prod_{p>2}(1+\rho_p) \\
&= 2\prod_{p>2}\Big[1-\frac{1}{(p-1)^2}\Big] \\
&= 2T_2, \qquad (20.84)
\end{aligned}
$$

where

Deﬁnition 20.5.16 (Twin Prime Constant).

$$
T_2 = \prod_{p>2}\Big[1-\frac{1}{(p-1)^2}\Big] \approx .6601618158 \qquad (20.85)
$$

is the twin prime constant.

$\mathfrak{S}_N$

We need to estimate $|\mathfrak{S}-\mathfrak{S}_N|$*. As* ρ_q *is multiplicative and zero if* $q = p^k$ $(k \geq 2)$*, we see we need only look at sums of* ρ_p*. As* $\rho_p = -\frac{1}{(p-1)^2}$*, one can show that the difference between* $\mathfrak{S}$ *and* $\mathfrak{S}_N$ *tends to zero as* $N \to \infty$*.*

Thus,

Lemma 20.5.17.

$$
\mathfrak{S} = 2T_2. \qquad (20.86)
$$

20.5.3 Number of Germain Primes and Weighted Sums

Combining the above arguments, we have shown that, up to lower order terms,

$$
\begin{aligned}
\sum_{\substack{p\leq N \\ p,\,\frac{p-1}{2}\ \text{prime}}} \log(p)\cdot\log\Big(\frac{p-1}{2}\Big) &= \mathfrak{S}\frac{N}{2} \\
&= 2T_2\frac{N}{2} \\
&= T_2 N. \qquad (20.87)
\end{aligned}
$$

Note that we are counting Germain prime pairs by $\left(\frac{p-1}{2}, p\right)$ and not $(p, 2p+1)$. Such a difference in counting will introduce a factor of 2.

We can pass from this weighted sum to a count of the number of Germain prime pairs $\left(\frac{p-1}{2}, p\right)$ with $p \leq N$.

Again we follow Nathanson, Chapter 8. De£ne

$$\begin{aligned} \pi_G(N) &= \sum_{\substack{p \leq N \\ p, \frac{p-1}{2} \text{ prime}}} 1 \\ G(N) &= \sum_{\substack{p \leq N \\ p, \frac{p-1}{2} \text{ prime}}} \log(p) \cdot \log\left(\frac{p-1}{2}\right). \end{aligned} \tag{20.88}$$

Clearly

$$G(N) \leq \log^2 N \cdot \pi_G(N). \tag{20.89}$$

Therefore,

Lemma 20.5.18. *Up to lower order terms,*

$$\pi_G(N) \geq \frac{G(N)}{\log^2 N} = \frac{T_2 N}{\log^2 N}. \tag{20.90}$$

We now provide a bound in the opposite direction.

$$\pi_G(N^{1-\delta}) = \sum_{\substack{p \leq N^{1-\delta} \\ p, \frac{p-1}{2} \text{ prime}}} 1 \ll \frac{N^{1-\delta}}{\log N}. \tag{20.91}$$

Then

$$
\begin{aligned}
G(N) &\geq \sum_{\substack{p \geq N^{1-\delta} \\ p, \frac{p-1}{2} \text{ prime}}} \log p \cdot \log\left(\frac{p-1}{2}\right) \\
&= (1-\delta)^2 \log^2 N \sum_{\substack{p \geq N^{1-\delta} \\ p, \frac{p-1}{2} \text{ prime}}} 1 \\
&= (1-\delta)^2 \log^2 N \Big(\pi_G(N) - \pi_G(N^{1-\delta})\Big) \\
&\geq (1-\delta)^2 \log^2 N \pi_G(N) + O\left((1-\delta)^2 \log^2 N \cdot \frac{N^{1-\delta}}{\log N}\right)
\end{aligned} \tag{20.92}
$$

Therefore

$$
\begin{aligned}
\log^2 N \cdot \pi_G(N) &\leq (1-\delta)^{-2} \cdot G(N) + O\left(\log^2 N \cdot \frac{N^{1-\delta}}{\log N}\right) \\
0 \leq \log^2 N \cdot \pi_G(N) - G(N) &\leq \Big[(1-\delta)^{-2} - 1\Big] G(N) + O\Big(\log N \cdot N^{1-\delta}\Big)
\end{aligned} \tag{20.93}
$$

If $0 < \delta < \frac{1}{2}$, then $(1-\delta)^{-2} - 1 \ll \delta$. We thus have

$$
0 \leq \log^2 N \cdot \pi_G(N) - G(N) \ll N\left[\delta + O\Big(\frac{\log N}{N^\delta}\Big)\right]. \tag{20.94}
$$

Choose $\delta = \frac{2 \log \log N}{\log N}$. Then we get

$$
0 \leq \log^2 N \cdot \pi_G(N) - G(N) \leq O\Big(N \frac{\log \log N}{\log N}\Big). \tag{20.95}
$$

Recalling $G(N) \approx T_2 N$ gives

Lemma 20.5.19.

$$
\pi_G(N) \leq \frac{T_2 N}{\log^2 N}. \tag{20.96}
$$

Combining with the other bound we have £nally shown

Theorem 20.5.20. *Assuming there is no contribution to the main term from the Minor Arcs, up to lower order terms we have*

$$\pi_G(N) \ = \ \frac{T_2 N}{\log^2 N}, \tag{20.97}$$

where T_2 is the twin prime constant

$$T_2 \ = \ \prod_{p>2}\left[1 - \frac{1}{(p-1)^2}\right] \ \approx \ .6601618158. \tag{20.98}$$