Discrete Math, Second Problem Set (June 24)

REU 2003

Instructor: Laszlo Babai
Scribe: D. Jeremy Copeland

1 Number Theory

Remark 1.1. For an arithmetic progression, $a_0, a_1 = a_0 + d, a_2 = a_0 + 2d, \cdots$ to have infinitely many primes, it is necessary that $\gcd(d, a_0) = 1$.

This is true because the gcd will divide all terms. The converse is also true:

Theorem 1.2. (Dirichlet) *Whenever $gcd(d, a_0) = 1$, the arithmetic progression, $a_0, a_1 = a_0 + d, a_2 = a_0 + 2d, \cdots$ will have infinitely many primes.*

This remarkable theorem is proved using complex analysis. We shall give elementary proofs of special cases. The case $d = 1$ is simply the infinitude of primes.

Theorem 1.3. (Euclid) *There are infinitely many primes.*

Proof: For a contradiction, assume there are finitely many, $p_1, \ldots, p_n$. Construct $N = \Pi_{i=1}^n p_i$. Then for any i, p_i does not divide $N + 1$, so $N + 1$ is not divisible by any prime. This is a contradiction, since all numbers ≥ 2 are divisible by some prime. (This can be easily proved by induction.) $\square$

Definition 1.4. We say that two numbers, a, and b are **congruent mod m**, or $a \equiv b \mod m$ if $m \mid b - a$.

Lemma 1.5. *If $n \equiv -1 \mod 4$, then it must have a prime divisor $\equiv -1 \mod 4$.*

Proof: n is an odd number. Thus it is the product of odd primes. If all of these primes are $\equiv 1 \mod 4$, then their product is also $\equiv 1 \mod 4$, which is a contradiction. $\square$

Theorem 1.6. *There are infinitely many primes $\equiv -1 \mod 4$.*

Proof: Assume that there are finitely many, $p_1, \cdots p_n$. Construct $N = \Pi_{i=1}^n p_i$. Then by the Lemma, $4N - 1$ must have a prime divisor congruent to $-1 \mod 4$. However, this is a contradiction, since no p_i divides it (why?). $\square$

The infinitude of primes $\equiv 1 \mod 4$ is not so straightforward. We need the following

Lemma 1.7. *If p is an odd prime, and $p \mid a^2 + 1$, then $p \equiv 1 \mod 4$.*

Theorem 1.8. *There are infinitely many primes congruent to $1 \mod 4$.*

Proof: Assume that there are finitely many, $p_1, \cdots p_n$. Construct $N = \Pi_{i=1}^{n} p_i$. Then, by the Lemma, all (odd) prime divisors of $4N^2 + 1$ must be congruent to $1 \mod 4$. This is a contradiction, since this number is not divisible by any of the p_i. (Why did we multiply by 4?) □

We need to prove the Lemma. This, in turn, requires Fermat's Little Theorem, which we now state.

Theorem 1.9 (Fermat's Little Theorem). *If p is a prime, then $p \,|\, a^p - a$.*

- *Equivalently: If p is a prime then $a^p \equiv a \mod p$.*
- *Equivalently: If p is a prime and $p \nmid a$, then $p \,|\, a^{p-1} - 1$.*
- *Equivalently: If p is a prime and $p \nmid a$, then $a^{p-1} \equiv 1 \mod p$.*

Exercise 1.10. Show that each of the last three statements is equivalent to the first.

Proof: We may, (by adding a multiple of p to a) assume that a is positive. Now consider all strings of length p of numbers between 1 and a. For example, if $p = 7$, $a = 4$, we could have $(1, 4, 4, 3, 4, 1, 1)$. There are a^p such strings. Call two strings equivalent if "they make the same necklace," i. e., they differ by cyclic rotation. That is, the above string is equivalent to $(4, 4, 3, 4, 1, 1, 1)$, $(4, 3, 4, 1, 1, 1, 4)$, $(3, 4, 1, 1, 1, 4, 4)$, etc. Now in most cases, there will be exactly p equivalent strings. The exceptions are the strings $(x, x, x, x, \cdots, x)$, which will have only one string (themselves) in their equivalence class. Therefore the number of equivalence classes ("necklaces") is $(a^p - a)/p + a$, so $(a^p - a)/p$ must be an integer. □

Definition 1.11. a is a **quadratic residue** modulo the prime p if $a \not\equiv 0 \mod p$ and there exists an x such that $x^2 \equiv a \mod p$.

Proof of Lemma 1.7. The condition is that $a^2 \equiv -1 \mod p$; and by Fermat's Little Theorem we have $a^{p-1} \equiv 1 \mod p$. Since p is odd, $1 \equiv a^{p-1} \equiv (a^2)^{(p-1)/2} \equiv (-1)^{(p-1)/2}$. Now since p is odd, $-1 \not\equiv 1 \mod p$, so $(p-1)/2$ must be even, or $4 \,|\, p - 1$. □

Exercise 1.12. Prove: every prime greater than 3 is congruent to $\pm 1 \mod 6$.

Exercise 1.13. There are infinitely many primes of the form $6k - 1$. *Hint.* Follow the $4k - 1$ case.

Exercise 1.14. There are infinitely many primes of the form $6k + 1$.

This is harder; you need a new lemma:

Exercise 1.15. If p is a prime and $p \mid a^2+a+1$, then $p=3$ or $p \equiv 1 \mod 6$. *Hint.* Prove: $a^3 \equiv 1 \mod p$.

Definition 1.16. If $p \nmid a$, the **order of** $a \mod p$ is the smallest positive integer k, such that $a^k \equiv 1 \mod p$. We say $k = \mathrm{ord}_p(a)$.

Exercise 1.17. Prove that the following statement is equivalent to Fermat's Little Theorem, 1.9.

- If p is a prime and $p \nmid a$, then $\mathrm{ord}_p(a) \mid p-1$.

Exercise 1.18. If p is a prime number other than 2 or 5 then the decimal expansion of $1/p$ is periodic with period $\mathrm{ord}_p(10)$.

Definition 1.19. We say that a is a **primitive root modulo** p if $\mathrm{ord}_p(a) = p-1$.

Exercise 1.20. 10 is a primitive root mod 7 if and only if 3 is a primitive root mod 7.

Exercise 1.21. Verify: $\{1,3,3^2,3^3,3^4,3^5,3^6\} \equiv \{1,3,2,6,4,5\} \pmod 7$, so that all numbers from 1 to 6 appear before repetition occurs in the sequence $3^k \mod 7$.

Exercise 1.22. Prove: a is a primitive root mod p if and only if the set $\{1,a,a^2,\cdots a^{p-2}\}$ represents all numbers numbers modulo p that are not divisible by p.

Theorem 1.23. *For every prime, p, there exists a primitive root mod p.*

(We shall prove this important result later and will not use it in this set of exercises.)

Our next goal is to prove the following remarkable result:

Theorem 1.24 (Gauss). *A prime number p can be written as a sum of two squares if and only if $p=2$ or $p \equiv 1 \mod 4$.*

The "only if" part is straightforward:

Exercise 1.25. If $p \equiv -1 \mod 4$, then p cannot be written as the sum of two squares. *Hint.* Observe that $(\forall x)(x^2 \equiv 0 \text{ or } 1 \mod 4)$.

To prove the converse, we need to learn about polynomials, quadratic residues, and lattices.

Theorem 1.26. *Assume F is a field. If $f(x)$ is a polynomial over F then $x-a$ is a divisor of $f(x)-f(a)$.*

Proof: Let $f(x) = \sum_{k=0}^{n} c_k x^k$. Since $x - a \mid x^k - a^k$, and $f(x) - f(a) = \sum c_k(x^k - a^k)$, we see that $(x-a) \mid f(x) - f(a)$, $\square$

Notation 1.27. $F[x]$ is the ring of all polynomials over F.

Corollary 1.28. *If $f \in F[x]$, and $f(a) = 0$, then $f(x) = (x-a)g(x)$ for some $g \in F[x]$.*

Corollary 1.29. *A nonzero polynomial of degree n has at most n roots.*

Proof: By induction on n. $\square$

Theorem 1.30. *For an odd prime p, the following are equivalent:*

(a) $p \equiv 1 \mod 4$.

(b) -1 *is a quadratic residue* $\bmod p$.

Proof: We already proved that (b) $\Rightarrow$ (a) (Lemma 1.7). For the other direction, assume now that $p \equiv 1 \mod 4$. Let $f(x) = x^p - x \in \mathbb{F}_p[x]$. By Fermat's Little Theorem, every element of $\mathbb{F}_p$ is a root of f. We factor f as $x^p - x = x(x^{p-1} - 1) = x(x^{(p-1)/2} - 1)(x^{(p-1)/2} + 1)$. Since none of the factors has more roots than its degree and the total number of roots is p, each factor must have exactly as many roots as its degree. In particular, there must exist a root a of $(x^{(p-1)/2} + 1)$. However, since $p \equiv 1 \mod 4$, $-1 \equiv a^{(p-1)/2} \equiv (a^{(p-1)/4})^2 \mod p$. Therefore -1 is a square mod p. $\square$

Definition 1.31. A lattice in $\mathbb{R}^n$ is a set $\mathbb{Z}\mathbf{u}_1 + \cdots + \mathbb{Z}\mathbf{u}_n = \{a_1\mathbf{u_1} + \cdots + a_n\mathbf{u_n} : a_1, \ldots, a_n \in \mathbb{Z}\}$, where $\{\mathbf{u_1}, \ldots, \mathbf{u_n}\}$ is linearly independent over $\mathbb{R}$.

Definition 1.32. The set $\{\sum_{i=1}^{n} a_i \mathbf{u}_i : 0 \le a_i \le 1\}$ is a **fundamental parallelepiped** of the lattice L.

Our main interest in this section will be the two-dimensional case, so instead of the n-dimensional parallelepipeds the reader may think of the familiar *parallelograms* in the plane.

Exercise 1.33. The translates of the fundamental parallelepiped by vectors in L tile $\mathbb{R}^n$, i.e., they cover $\mathbb{R}^n$ without overlapping interiors.

Theorem 1.34. *A fundamental parallelepiped is characterized by the fact that it is a parallelepiped which intersects the lattice in exactly its 2^n corners.*

Recall that the area of a parallelogram spanned by $\mathbf{u} = \binom{u_1}{u_2}$ and $\mathbf{v} = \binom{v_1}{v_2}$ is *Area* $= |u_1v_2 - u_2v_1| = |\det \begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \end{pmatrix}|$. (The analogous determinant expression works in n-dimensions.)

Exercise 1.35. All fundamental parallelepipeds have the same area.

Definition 1.36. A subset S of a real vector space V, is called **convex** if for each $t \in \mathbb{R}$, $0 < t < 1$, and each $\mathbf{u}$, $\mathbf{v}$ in V, $(1-t)\mathbf{u} + t\mathbf{v} \in S$.

Theorem 1.37 (Minkowski). *Let L be a lattice in $\mathbb{R}^n$, and the area of a findamental parallelogram be A. If S is convex, centrally symmetric about the origin, and the area of S is greater than $2^n A$, then $L \cap S \neq \{0\}$.*

(The proof of this fundamental result in the "Geometry of Numbers" will be given later.) Next we use Minkowski's Theorem to prove Gauss' result, 1.24. The devilishly clever proof is due to Paul Turán.

Theorem 1.38. *If $p \equiv 1 \mod 4$, then there are integers, e, and f, such that $p = e^2 + f^2$.*

Proof: By Theorem 1.30, there exists a c such that $p \mid c^2 + 1$. Let $\mathbf{u} = \binom{c}{1}$ and $\mathbf{v} = \binom{p}{0}$ be vectors in $\mathbb{R}^2$. A general point in the lattice, L spanned by these is $x\mathbf{u} + y\mathbf{v} = \binom{xc+yp}{x}$. Note that $(xc+yp)^2 + x^2 \equiv x^2(c^2+1) \equiv 0 \mod p$, so if $\binom{e}{f}$ is in L, then $p \mid e^2 + f^2$. Look at the open disk: $S = \{\binom{e}{f} : e^2 + f^2 < \sqrt{2p}\}$. The area of S is $2\pi p > 4p$, and the area of the fundamental parallelogram is $|c0 - 1p| = p$, so by Minkowski's Theorem, 1.37, there is some nonzero point in the lattice and in the disk. Therefore, there exist e and f such that $0 < e^2 + f^2 < 2p$, and $p \mid e^2 + f^2$, so $e^2 + f^2 = p$. $\square$

Divisor game. Consider the following game, played by two people: Choose a number n. Players take turns naming a factor of n. Once a factor is named, it and none of its factors may be named again. The loser is the first player to say n.

Exercise 1.39. Find the winning strategies for the first player when $n = p^k$, $n = p^k q$, $n = p^k q^k$, $n = pqrs$.

Exercise 1.40. Show that the first player always has a winning strategy.
Hint. Do not try to find the winning strategy for all n.

Coin placing game. Given a rectangular table and an unlimited supply of quarters, each player takes turns placing quarters on the table (this is a "continuous" game). The quarters, once placed, may not be moved; they may not overlap, and may not hang over the edge.

Exercise 1.41. Find a winning strategy for the first player.

2 Linear Algebra

Let F be a field. Note that the ring $F[x]$ of univariate polynomials over F is an infinite-dminesional vector space over F with a nice (standard) basis: $\{1, x, x^2, x^3, x^4, \dots\}$.

Example 2.1. Let $F_n[x]$ denote the set of polynomials of degree $\leq n$. This is an $n+1$-dimensional subspace of $F[x]$. We calculate the matrix representing the map $\frac{d}{dx} : F_n[x] \to F_n[x]$ with respect to the standard basis $\mathbf{B} := \{1, x, \ldots, x^n\}$.

$$\left[\frac{d}{dx}\right]_{\mathbf{B}} = \begin{bmatrix} 0 & 1 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 2 & 0 & \ldots & 0 \\ 0 & 0 & 0 & 3 & \ldots & \vdots \\ 0 & 0 & 0 & 0 & \ddots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & n \\ 0 & 0 & 0 & 0 & \ldots & 0 \end{bmatrix}$$

Example 2.2. Consider the rotation of the Euclidean plane $\mathbb{E}^2$ by angle α about the origin. We denote this linear map by $R_\alpha : \mathbb{E}^2 \to \mathbb{E}^2$. Choose as the basis a pair of perpendicular unit vectors $\mathbf{e}$ and $\mathbf{f}$. Then we have that $R_\alpha(\mathbf{e}) = \cos(\alpha)\mathbf{e} + \sin(\alpha)\mathbf{f}$, and $R_\alpha(\mathbf{f}) = -\sin(\alpha)\mathbf{e} + \cos(\alpha)\mathbf{f}$. Therefore the matrix for R_α in the basis $\mathbf{e}, \mathbf{f}$, is

$$[R_\alpha]_{\mathbf{e},\mathbf{f}} = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix}.$$

Example 2.3. Consider the reflection of the Euclidean plane $\mathbf{E^2}$ through the line which makes an angle α with horizontal ($\mathbf{e}$) and passes through the origin. We denote this linear map by $F_\alpha : \mathbb{E}^2 \to \mathbb{E}^2$. Then we have $F_\alpha(\mathbf{e}) = \cos(2\alpha)\mathbf{e} + \sin(2\alpha)\mathbf{f}$, and $F_\alpha(\mathbf{f}) = \sin(2\alpha)\mathbf{e} - \cos(2\alpha)\mathbf{f}$. Therefore the matrix for F_α in the basis $\mathbf{e}, \mathbf{f}$, is

$$[F_\alpha]_{\mathbf{e},\mathbf{f}} = \begin{bmatrix} \cos(2\alpha) & \sin(2\alpha) \\ \sin(2\alpha) & -\cos(2\alpha) \end{bmatrix}.$$

Example 2.4. If in the previous example, we had chosen the basis $\mathbf{u}$ and $\mathbf{v}$, where $\mathbf{u}$ is a vector along the line of reflection, and $\mathbf{v}$ is perpendicular to it, then the matrix for F_α would be:

$$[F_\alpha]_{\mathbf{u},\mathbf{v}} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Definition 2.5. Given a square matrix A, define $\mathrm{tr}(A)$, the **trace** of A to be the sum of the diagonal entries of A.

Remark 2.6. The following observations are special cases of a general principle:

- $\mathrm{tr}([F_\alpha]_{\mathbf{e},\mathbf{f}}) = 0 = \mathrm{tr}([F_\alpha]_{\mathbf{u},\mathbf{v}})$.
- $\det([F_\alpha]_{\mathbf{e},\mathbf{f}}) = -1 = \det([F_\alpha]_{\mathbf{u},\mathbf{v}})$.

Definition 2.7. Let $\{\mathbf{e_1}, \ldots, \mathbf{e_n}\}$, and $\{\mathbf{f_1}, \ldots, \mathbf{f_n}\}$ be two bases for a vector space V. The basis change transformation $S : V \to V$ is the unique (invertible) linear map defined by $S : e_i \mapsto f_i$.

Example 2.8. Let $\mathbf{x}$ be a vector in V. Then $\mathbf{x} = \sum \alpha_i \mathbf{e}_i = \sum \beta_i \mathbf{f}_i$. Notice that

$$S\mathbf{x} = S \sum \alpha_i \mathbf{e}_i = \sum \alpha_i S \mathbf{e}_i = \sum \alpha_i \mathbf{f}_i.$$

Therefore if

$$[\mathbf{x}]_\mathbf{e} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}$$

$[S\mathbf{x}]_\mathbf{f} = [\mathbf{x}]_\mathbf{e}$. Also, $[S\mathbf{x}]_\mathbf{f} = [S]_\mathbf{f}[\mathbf{x}]_\mathbf{f}$ by the following exercise. Thus$[\mathbf{x}]_\mathbf{f} = [S^{-1}]_\mathbf{f}[\mathbf{x}]_\mathbf{e}$

Exercise 2.9. Let $\{\mathbf{e_1}, \ldots, \mathbf{e_n}\}$ be a basis for V, and let $\{\mathbf{f_1}, \ldots, \mathbf{f_n}\}$ be a basis for W. Let $\mathbf{v} \in V$, and let $A : V \to W$. Show that $[A]_{\mathbf{e},\mathbf{f}}[\mathbf{v}]_\mathbf{e} = [A\mathbf{v}]_\mathbf{f}$.

Exercise 2.10. Let V, W, Z be vector spaces over F. Let $\{\mathbf{e_1}, \ldots, \mathbf{e_n}\}$ be a basis for V, $\{\mathbf{f_1}, \ldots, \mathbf{f_n}\}$ be a basis for W, $\{\mathbf{g_1}, \ldots, \mathbf{g_n}\}$ be a basis for Z. Let $A : V \to W$, $B : W \to Z$. Show that $[BA]_{\mathbf{e},\mathbf{g}} = [B]_{\mathbf{f},\mathbf{g}}[A]_{\mathbf{e},\mathbf{f}}$.

Exercise 2.11. If B, C are $n \times k$ matrices that $B\mathbf{x} = C\mathbf{x}$ for all $\mathbf{x} \in F^k$ then $B = C$.

Remark 2.12. Let $\mathbf{e}$, $\mathbf{e}'$ be two bases for V with change of basis map $S : \mathbf{e}_i \mapsto \mathbf{e}_i'$. Let $\mathbf{f}$, $\mathbf{f}'$ be two bases for W with change of basis map $T : \mathbf{f}_i \mapsto \mathbf{f}_i'$. Let $A : V \to W$. We want to compare $[A]_{\mathbf{ef}}$ with $[A]_{\mathbf{e'f'}}$. Let $\mathbf{x} \in V$. We have the following identities:

- $[A\mathbf{x}]_\mathbf{f} = [A]_{\mathbf{ef}}[\mathbf{x}]_\mathbf{e}$
- $[A\mathbf{x}]_{\mathbf{f'}} = [A]_{\mathbf{e'f'}}[\mathbf{x}]_{\mathbf{e'}} = [A]_{\mathbf{e'f'}}[S^{-1}]_{\mathbf{e'e'}}[\mathbf{x}]_\mathbf{e}$
- $[A\mathbf{x}]_{\mathbf{f'}} = [T^{-1}]_{\mathbf{f'f'}}[A\mathbf{x}]_\mathbf{f}$.

Therefore for all $\mathbf{x} \in V$,

$$([T^{-1}]_{\mathbf{f'f'}}[A]_{\mathbf{ef}})[\mathbf{x}]_\mathbf{e} = ([A]_{\mathbf{e'f'}}[S^{-1}]_{\mathbf{e'e'}})[\mathbf{x}]_\mathbf{e}.$$

However, by the previous exercise, this is only possible if

$$[T^{-1}]_{\mathbf{f'f'}}[A]_{\mathbf{ef}} = [A]_{\mathbf{e'f'}}[S^{-1}]_{\mathbf{e'e'}}.$$

Remark 2.13. From the previous remark, we deduce that

$$[A]_{\mathbf{ef}} = [T]_{\mathbf{f'f'}}[A]_{\mathbf{e'f'}}[S^{-1}]_{\mathbf{e'e'}} = [TAS]_{\mathbf{e'f'}}.$$

Corollary 2.14. *Let $S = A = T$. Then*

$$[S]_{\mathbf{ee}} = [TAS^{-1}]_{\mathbf{ff}} = [S]_{\mathbf{ff}}.$$

Corollary 2.15. *If $V = W$, $e = f$, then $A_{new} = S^{-1}A_{old}.S$*

Definition 2.16. Two $n \times n$ matrices A, B are **similar** if there exists some invertible S such that $B = S^{-1}AS$. We write $A \sim B$.

Definition 2.17. The **characteristic polynomial** f_A of a matrix A is the determinant $f_A(x) = \det(xI - A)$, where I is the $n \times n$ identity matrix.

Exercise 2.18. If A and B are $n \times n$ matrices, then $\det(AB) = \det(A)\det(B)$.

Exercise 2.19. If $A \sim B$ then $\det(A) = \det(B)$.

Exercise 2.20. If $A \sim B$ then $f_A = f_B$.

Definition 2.21. If A is a matrix, the roots of f_A are called the **eigenvalues** of A.

Definition 2.22. If T is a linear transformation, then the **characteristic polynomial** of T is the characteristic polynomial of a matrix for T in some basis. The last exercise shows that this is well-defined.

Definition 2.23. If T is a linear transformation, then the **eigenvalues** of T are the roots of the characteristic polynomial.

Remark 2.24. If A is an upper triangular matrix, then the eigenvalues of A are the diagonal entries of A, taken with multiplicity.

Exercise 2.25. Find two 2×2 matrices, A and B such that $f_A = f_B$ but A and B are not similar.

Exercise 2.26. If A is an $n \times n$ matrix over a field F, and f_A has n distinct roots in F, then A is "diagonalizable," i.e., A is similar to a diagonal matrix. (What are the entries of this diagonal matrix?)