MATHEMATICS FROM EXAMPLES, SPRING 2023

Contents

1. Overview of the course 1
2. Measure theory and ergodic theory 4
2.1. An outlook 5
2.2. σ-algebras, measures, probability spaces 7
2.3. Measure-preserving functions 9
2.4. Recurrence 12
2.5. Lebesgue integral 13
2.6. Ergodicity 14
2.7. Ergodic theorems 18
2.8. Back to continued fractions 20
3. Topology 25
3.1. The Borsuk–Ulam theorem 25
3.2. Fundamental groups 28
3.3. Fundamental group of a circle and applications 31
3.4. The rectangular peg problem 34
4. Algebra 35
4.1. Rings 35
4.2. Ring of Gaussian integers 39
4.3. Applications 41
5. Modular forms 44
5.1. More applications of modular forms 45
5.2. Crash course on complex analysis 48
Bibliography 51

1. Overview of the course

Lecture 1

Example 1.1. Let $x \in (0,1)\backslash\mathbb{Q}$ be an irrational number. It can be written uniquely as a continued fraction

$$x = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{\ddots}}}}$$

where $a_1, a_2, \ldots$ are positive integers. How often does a positive integer k appear in this expression?

It turns out that for any given k, the frequency of k appearing in the continued fraction expression of x is the same for *almost every* $x \in (0,1)\backslash\mathbb{Q}$. In fact, for almost every $x \in (0,1)\backslash\mathbb{Q}$, we have

$$\lim_{n\to\infty} \frac{\#\{i \mid a_i = k, 1 \le i \le n\}}{n} = \frac{1}{\log 2} \log\left(\frac{(k+1)^2}{k(k+2)}\right).$$

To prove this, we will introduce some basic ideas of *measure theory* and *ergodic theory*.

Example 1.2. Consider the following *necklace-splitting problem*. Two thieves have stolen a precious necklace (opened, with two ends), on which there are d kinds of stones (diamonds, sapphires, rubies, etc.), an even number of each kind. The thieves do not know the values of stones of various kinds, so they want to divide the stones of each kind evenly. They would like to achieve this by as few cuts as possible. The question is, what is the minimum amount of cuts to divide the stones of each kind evenly?

It is not hard to show that at least d cuts may be necessary: Place the stones of the first kind first, then the stones of the second kind, and so on. The *necklace theorem* shows that this is the worst, what can happen. In other words, d cuts is always sufficient. Surprisingly, all known proofs of this theorem are *topological*.

Example 1.3. Let $C \subseteq \mathbb{R}^2$ be a simple closed curve. One considers the following *Rectangular Peg Problems*.

- Does there always exist four points on C such that they form the vertices of a rectangle?
- Even harder question: Fix a rectangle R. Does there always exist four points on C such that they form the vertices of a rectangle which is similar to R?

The first question was answered positively by Vaughan in 1981, which uses some basic *topology*. The second question was also answered positively quite recently by Greene and Lobb; their proof involves more advanced tools from *symplectic geometry*, which is beyond the scope of this course.

Example 1.4. Which positive integers n can be written as the sum of two squares $n = x^2 + y^2$?

To answer this question, it is natural to introduce the *ring* of *Gaussian integers* $\mathbb{Z}[i]$, since one has the factorization $x^2 + y^2 = (x + iy)(x - iy)$. The question then reduced to studying the properties of the ring $\mathbb{Z}[i]$.

Example 1.5. How many ways can a positive integer n be written as the sum of two (or more) squares?

The problem is closely related to the *Jacobi theta function*, which is a function defined for two complex variables $z \in \mathbb{C}$ and $\tau \in \mathbb{H}$:

$$\theta(z;\tau) = \sum_{n=-\infty}^{\infty} \exp(\pi i n^2 \tau + 2\pi i n z) = \sum_{n=-\infty}^{\infty} q^{n^2} u^n$$

where $q = \exp(\pi i \tau)$ and $u = \exp(2\pi i z)$. By taking $z = 0$ we have

$$\theta(0;\tau) = \sum_{n=-\infty}^{\infty} q^{n^2}.$$

Let us define $r_2(n)$ to be the number of ways that n can be written as the sum of two squares; to be more precise,

$$r_2(n) = \#\{(x, y) \in \mathbb{Z}^2 \mid x^2 + y^2 = n\}.$$

It is not hard to see that

$$\theta(0;\tau)^2 = \sum_{n=0}^{\infty} r_2(n) q^n.$$

The problem then reduces to understand $\theta(0;\tau)^2$. It turns out that $\theta(0;\tau)^2$ is a *modular form of weight* 1 *for the congruence subgroup* $\Gamma_1(4) \subseteq \mathrm{SL}(2,\mathbb{Z})$, and we can use the theory of modular forms to obtain an explicit formula of $r_2(n)$.

Example 1.6. Consider the following (generalized) theta functions

$$\theta[a, z_0](z;\tau) = \sum_{n=-\infty}^{\infty} \exp\left(\pi i (n+a)^2 \tau + 2\pi i (n+a)(z+z_0)\right)$$

where $a \in \mathbb{R}$, $z, z_0 \in \mathbb{C}$, and $\tau \in \mathbb{H}$. Note that $a = z_0 = 0$ recovers the Jacobi theta function. The generalized theta functions satisfy the following addition formula

$$\theta[0,0](z;\tau)^2 = \theta[0,0](0;2\tau)\cdot\theta[0,0](2z;2\tau) + \theta[1/2,0](0;2\tau)\cdot\theta[1/2,0](2z;2\tau).$$

We will discuss a geometric interpretation of this formula, in terms of *mirror symmetry* between *complex* and *symplectic geometry* of *elliptic curves*.

2. Measure theory and ergodic theory

Recall our motivating question: Let $x \in (0,1)\backslash\mathbb{Q}$ be an irrational number. It can be written uniquely as a continued fraction

$$x = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \frac{1}{\ddots}}}}$$

where $a_1, a_2, \ldots$ are positive integers. How often does a positive integer k appear in this expression? Below is the sketch of ideas toward answering this question.

- Define the *continued fraction map* $T\colon [0,1] \to [0,1]$ by $T(0) = 0$ and

$$T(x) = \frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor \text{ for } x \neq 0,$$

  where $\lfloor t \rfloor$ denotes the greatest integer less than or equal to t. In other words, $T(x)$ is the fractional part $\{\frac{1}{x}\}$ of $\frac{1}{x}$.
- Observe that $a_n = k$ if and only if $T^{n-1}(x) \in (\frac{1}{k+1}, \frac{1}{k}]$. Hence

$$\frac{\#\{i \mid a_i = k, 1 \leq i \leq n\}}{n} = \frac{1}{n}\sum_{i=0}^{n-1} \chi_{(\frac{1}{k+1},\frac{1}{k}]}(T^i(x))$$

  where χ is the characteristic function.
- Define the *Gauss measure* μ on $[0,1]$ to be

$$\mu(A) = \frac{1}{\log 2}\int_A \frac{1}{1+x}\,\mathrm{d}x \text{ for any measurable set } A \subseteq [0,1].$$

- Prove that the Gauss measure μ is *T-invariant* and *ergodic* .

- By *Birkhoff's pointwise ergodic theorem*, for almost every $x \in [0,1]\backslash\mathbb{Q}$ we have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \chi_{(\frac{1}{k+1},\frac{1}{k}]}(T^i(x)) = \int \chi_{(\frac{1}{k+1},\frac{1}{k}]} \,\mathrm{d}\mu = \mu\left(\left(\frac{1}{k+1},\frac{1}{k}\right]\right).$$

- The conclusion then follows from a simple calculation

$$\frac{1}{\log 2} \int_{\frac{1}{k+1}}^{\frac{1}{k}} \frac{1}{1+x} \,\mathrm{d}x = \frac{1}{\log 2} \log\left(\frac{(k+1)^2}{k(k+2)}\right).$$

In order to understand this approach and appreciate the powerful tools provided by ergodic theory (in our case, the pointwise ergodic theorem), we will discuss the following topics in this section:

- basic measure theory;
- basic ergodic theory;
- ergodic theorems and applications.

Some references that might be helpful include [4] and [9].

2.1. **An outlook.** Consider a map $T\colon X \to X$. In ergodic theory, one studies how *typical* orbits $\{x, T(x), T^2(x), \ldots\}$ are distributed. We would be interested in properties like *frequencies of visits*, *equidistribution*, *mixing*, etc.

Here is a basic example. Let $A \subseteq X$ be a subset, and x be an element of X. The number of visits of orbit of x to the subset A up to time n is given by

$$\#\{0 \le k \le n-1 \mid T^k(x) \in A\}.$$

A convenient way to write this quantity is as follows. Let $\chi_A\colon X \to \mathbb{R}$ be the characteristic function of the subset A: $\chi_A(x) = 1$ if $x \in A$, and $\chi_A(x) = 0$ if $x \notin A$. Then we have

$$\sum_{k=0}^{n-1} \chi_A(T^k(x)) = \#\{0 \le k \le n-1 \mid T^k(x) \in A\}.$$

The *frequency* of visits up to time n is defined to be the average

$$\frac{1}{n} \sum_{k=0}^{n-1} \chi_A(T^k(x)) \in [0,1].$$

Question 2.1. *We are interested in the following questions.*

(a) *Does the frequency of visits converge to a limit as n tends to infinity? (for all points of $x \in X$? or only for a typical point?)*
(b) *If the limit exists, what does the frequency converge to?*

Another type of question concerns the equidistributioness. Let us consider specifically in the setting of the unit interval $[0, 1]$. We say a sequence of points $\{x_n\}$ in $[0, 1]$ is *equidistributed* if for all intervals $I \subseteq [0, 1]$ we have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \chi_I(x_k) = \text{length}(I).$$

An equivalent definition is for all continuous functions $f\colon [0, 1] \to \mathbb{R}$ we have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(x_k) = \int_0^1 f(x)\,\mathrm{d}x.$$

So if we have a dynamical system $T\colon [0, 1] \to [0, 1]$ (or $T\colon S^1 \to S^1$, where $S^1 \cong \mathbb{R}/\mathbb{Z} \cong [0, 1]/0 \sim 1$), we can ask whether orbits $\{x.T(x), T^2(x), \ldots\}$ are equidistributed or not.

Example 2.2. Consider the *rotation* map

$$R_\alpha\colon S^1 \to S^1; \qquad x \mapsto x + \alpha \pmod 1.$$

If $\alpha \in \mathbb{Q}$ is a rational number, then every orbit of R_α is periodic, therefore cannot be equidistributed. If $\alpha \notin \mathbb{Q}$ is irrational, then one can show that every orbit of R_α is equidistributed (this is often thought of as the first ergodic theorem to have been proved).

Example 2.3. Consider the *doubling map*

$$T_2\colon S^1 \to S^1; \qquad x \mapsto 2x \pmod 1.$$

It is not hard to see that there is a dense subset of X for which the orbit of T_2 is periodic, therefore not equidistributed. However, it turns out that for *almost all* $x \in X$ the orbit of T_2 is equidistributed.

We may also have maps (e.g. the continued fraction map) where the orbits are not equidistributed for almost all $x \in X$. To make these notions precise, we need to introduce some measure theory, which have the advantage of introducing a theory of integration that is suitable for our purposes.

2.2. **σ-algebras, measures, probability spaces.** Intuitively, a *measure* μ on a space X is a function on a collection of subsets of X, called *measurable sets*, which assigns to each measurable set A its *measure* $\mu(A) \geq 0$. You already know at least two natural examples of measures.

Example 2.4. Let $X = \mathbb{R}$. The Lebesgue measure λ on $\mathbb{R}$ assigns to each interval $[a, b]$ its length

$$\lambda([a, b]) = b - a = \int_a^b \mathrm{d}x.$$

Let $X = \mathbb{R}^2$. The Lebesgue measure λ on $\mathbb{R}^2$ assigns to each measurable set $A \subseteq \mathbb{R}^2$ its area

$$\lambda(A) = \int_A \mathrm{d}x\,\mathrm{d}y.$$

One might hope to assign a measure to all subsets of X. Unfortunately, if we want the measure to have reasonable and useful properties, this would lead to a contradiction in certain cases (we will see an example later). So we are forced to assign a measure only to a sub-collection of all subsets of X.

Let X be a set. Denote by $\mathbb{P}(X)$ the collection of all subsets of X.

Definition 2.5. A subset $\mathcal{B} \subseteq \mathbb{P}(X)$ is called a *σ-algebra* on X if

(a) the empty set $\emptyset \in \mathcal{B}$,
(b) $\mathcal{B}$ is closed under complementation: $A \in \mathcal{B}$ implies $X \backslash A \in \mathcal{B}$,
(c) $\mathcal{B}$ is closed under countable union: $A_1, A_2, \ldots \in \mathcal{B}$ implies $\cup_{n=1}^{\infty} A_i \in \mathcal{B}$.

Elements of the σ-algebra are called *measurable sets*.

Remark 2.6. Let $F \subseteq \mathbb{P}(X)$ be an arbitrary subset (may or may not be a σ-algebra). Then there exists a unique smallest σ-algebra which contains every set in F. It is called the *σ-algebra generated by F*.

An important example is the *Borel algebra* over any *topological space*: it is the σ-algebra generated by the *open sets*. For instance, the Borel algebra over $[0, 1]$ is the σ-algebra generated by the collection of open sub-intervals of $[0, 1]$.

Definition 2.7. Let X be a set and $\mathcal{B}$ be a σ-algebra on X. A function $\mu \colon \mathcal{B} \to \mathbb{R} \cup \{\infty\}$ is called a *measure* if

(a) $\mu(\emptyset) = 0$,
(b) (non-negativity) $\mu(E) \geq 0$ for all $E \in \mathcal{B}$,

(c) (countable additivity) for all countable collections $\{E_k\}_{k=1}^{\infty}$ of pairwise disjoint sets in $\mathcal{B}$, we have

$$\mu\left(\bigcup_{k=1}^{\infty} E_k\right) = \sum_{k=1}^{\infty} \mu(E_k).$$

The triple $(X, \mathcal{B}, \mu)$ is called a *measurable space*, and it is called a *probability space* if $\mu(X) = 1$.

Example 2.8. Let $X = [0, 1]$ and let $\mathcal{B}$ be the Borel algebra on X, i.e. the σ-algebra generated by all open subintervals (a, b). There exists a measure (the *Lebesgue* measure) $\lambda\colon \mathcal{B} \to \mathbb{R}$ such that $\lambda((a, b)) = b - a$. The triple $(X, \mathcal{B}, \lambda)$ forms a probability space.

Remark 2.9. Given a probability space $(X, \mathcal{B}, \mu)$, one can regard X as the space of all possible *events*, and $\mu(A)$ gives the probability of an event occurs in a measurable subset $A \subseteq X$.

Example 2.10. Let us consider a discrete example. Let $X = \{1, \ldots, n\}$, and let $\mathcal{B} = \mathbb{P}(X)$ be the σ-algebra consists of all subsets of X. Choose any $0 \leq p_1, \ldots, p_n \leq 1$ such that $\sum p_i = 1$. Then one can define a measure $\mu\colon \mathcal{B} \to \mathbb{R}$ by

$$\mu(\{i_1, \ldots, i_k\}) = p_{i_1} + \cdots + p_{i_k}.$$

Remark 2.11. In this remark, we show that in general it is necessary to restrict the definition of measure on a subset $\mathcal{B} \subseteq \mathbb{P}(X)$, as opposed to defining it on the *whole* collection of subsets of X. Consider the Lebesgue measure $\lambda\colon \mathcal{B} \to \mathbb{R}_{\geq 0}$ on $X = \mathbb{R}$. It satisfies the following properties:

- λ has the countable additivity property in the definition of measure,
- if two subsets of A and B are related by a translation, then $\lambda(A) = \lambda(B)$,
- $\lambda([0, 1]) = 1$.

We show that unfortunately it is not possible to extend the definition of λ to *all* subsets of $\mathbb{R}$ that still satisfy these three properties.

Let us consider the example constructed by Vitali in 1905. A *Vitali set* is a subset $V \subseteq [0, 1]$ of real numbers such that, for each real number r, there is exactly one number $v \in V$ such that $v - r \in \mathbb{Q}$. Equivalently, V is constructed

by choosing a representative in $[0,1]$ of each element of the quotient group $\mathbb{R}/\mathbb{Q}$.

Let $q_1, q_2, \ldots$ be an enumeration of the rational numbers in $[-1,1]$ (recall that $\mathbb{Q}$ is *countable*). Consider the translated sets $V_k = V + q_k$ for $k = 1, 2, \ldots$. It is not hard to show the following:

- V_k's are pairwise disjoint,
- $[0,1] \subseteq \cup_{k=1}^{\infty} V_k \subseteq [-1,2]$.

Assume the contrary that it is possible to extend the definition of Lebesgue measure to *all* subsets of $\mathbb{R}$ which satisfies the properties above. Then we have

$$1 \leq \sum_{k=1}^{\infty} \lambda(V_k) \leq 3.$$

Since Lebesgue measure is translation invariant, we have $\lambda(V_k) = \lambda(V)$, hence

$$1 \leq \sum_{k=1}^{\infty} \lambda(V) \leq 3.$$

But this is impossible: If $\lambda(V) = 0$ then $\sum_{k=1}^{\infty} \lambda(V) = 0$; if $\lambda(V) > 0$ then $\sum_{k=1}^{\infty} \lambda(V) = \infty$. Contradiction.

2.3. **Measure-preserving functions.**

Definition 2.12. Let $(X, \mathcal{B}, \mu)$ and $(Y, \mathcal{C}, \nu)$ be two probability spaces.

- A map $T\colon X \to Y$ is called *measurable* if $T^{-1}(A) \in \mathcal{B}$ for any $A \in \mathcal{C}$.
- Furthermore, a measurable function T is called *measure-preserving* if $\mu(T^{-1}(A)) = \nu(A)$ for any $A \in \mathcal{C}$.
- If $T\colon X \to X$ is measure-preserving, then we say $(X, \mathcal{B}, \mu, T)$ is a *measure-preserving system*.

Exercise. Let X be a topological space and $\mathcal{B}$ be the Borel σ-algebra on X (which is generated by open sets of X). Show that any *continuous* map $T\colon X \to X$ is measurable.

Exercise. To show a measurable map $T\colon X \to Y$ is measure-preserving, it is enough to check $\mu(T^{-1}(A)) = \nu(A)$ holds for a generating set of $\mathcal{C}$.

Example 2.13 (Rotation on S^1). Consider the circle $S^1 \cong \mathbb{R}/\mathbb{Z}$, which can be obtained by identifying the two endpoints of $[0,1]$. One equips S^1 with the

Lebesgue measure. It is easy to show that the rotation

$$R_\alpha \colon S^1 \to S^1; \qquad x \mapsto x + \alpha \pmod 1$$

is measure-preserving for any α.

Example 2.14 (Doubling map on S^1). Define the *doubling map*

$$T_2 \colon S^1 \to S^1; \qquad x \mapsto 2x \pmod 1.$$

Let us show that it is measure-preserving. It is enough to check this on intervals: we have $\mu(T_2^{-1}(a,b)) = \mu(a,b)$ since

$$T_2^{-1}(a,b) = \left(\frac{a}{2}, \frac{b}{2}\right) \cup \left(\frac{a+1}{2}, \frac{b+1}{2}\right).$$

Note that the measure-preserving property cannot be seen by studying "forward iterates": $\mu(T_2(a,b)) \neq \mu(a,b)$ in general.

Example 2.15. Define the $(\frac{1}{2}, \frac{1}{2})$-measure $\mu_{(1/2,1/2)}$ on the finite set $\{1,2\}$ by

$$\mu_{(1/2,1/2)}(\{1\}) = \mu_{(1/2,1/2)}(\{2\}) = \frac{1}{2}.$$

Consider the space of infinite product $X = \{1,2\}^{\mathbb{N}}$, which models the set of possible outcomes of the infinitely repeated toss of a coin. Given a finite subset $I \subseteq \mathbb{N}$ and a map $a \colon I \to \{1,2\}$, we define the *cylinder set* associated to I and a to be

$$I(a) = \{x \in X \mid x_j = a(j) \text{ for all } j \in I\},$$

i.e. one specifies the outcome of the j-th thows for all $j \in I$. We define $\mathcal{B}$ to be the σ-algebra generated by all cylinder sets, and define a measure $\mu \colon \mathcal{B} \to \mathbb{R}$ via

$$\mu(I(a)) = \left(\frac{1}{2}\right)^{\#|I|}.$$

Consider the *left shift map* $\sigma \colon X \to X$ defined by

$$\sigma(x_1, x_2, \ldots) = (x_2, x_3, \ldots).$$

It is easy to see that $(X, \mathcal{B}, \mu, \sigma)$ is a measure-preserving system.

In fact, this system is *measurably isomorphic* to the doubling map T_2 on S^1, which roughly means that they are identical except on a measure zero set.

Indeed, consider the map $\phi\colon X \to S^1 \cong [0,1]/0 \sim 1$ where

$$\phi(x_1, x_2, \ldots) = \sum_{n=1}^{\infty} \frac{x_n}{2^n}.$$

Then we have $\phi \circ \sigma = T_2 \circ \phi$. Below is the precise definition of the notion of measurably isomorphic.

Definition 2.16. We say two measure-preserving systems $(X, \mathcal{B}, \mu, T)$ and $(Y, \mathcal{C}, \nu, S)$ are *measurably isomorphic* if there exists $X' \in \mathcal{B}$ and $Y' \in \mathcal{C}$ such that:

- $\mu(X') = \nu(Y') = 1$,
- $T(X') \subseteq X'$, $S(Y') \subseteq Y'$,
- there exists a bijective map $\phi\colon X' \to Y'$ such that both ϕ and ϕ^{-1} are measurable and measure-preserving, and
- $\phi \circ T(x) = S \circ \phi(x)$ for any $x \in X'$.

Example 2.17 (Bernoulli shift). Consider the two-sided infinite set

$$\begin{aligned} X &= \{1, \ldots, n\}^{\mathbb{Z}} \\ &= \{x = (\ldots, x_{-1}, x_0, x_1, \ldots) \mid x_i \in \{1, \ldots, n\} \text{ for all } i\}. \end{aligned}$$

which gives the sample space of the outcome of throwing an n-sided die (each appears with probabilities $p_1, \ldots, p_n$) infinitely many times. Let us define a σ-algebra and a measure on X. Given a finite subset $I \subseteq \mathbb{Z}$ and a map $a\colon I \to \{1, \ldots, n\}$, we define the *cylinder set* associated to I and a to be

$$I(a) = \{x \in X \mid x_j = a(j) \text{ for all } j \in I\},$$

i.e. one specifies the outcome of the j-th thows for all $j \in I$. We define $\mathcal{B}$ to be the σ-algebra generated by all cylinder sets, and define a measure $\mu\colon \mathcal{B} \to \mathbb{R}$ via

$$\mu(I(a)) = \prod_{j \in I} p_{a(j)}.$$

Now, consider the left shift map $\sigma\colon X \to X$ defined by $\sigma(x)_i = x_{i+1}$. It clearly preserves the measure of all cylinder sets, hence $(X, \mathcal{B}, \mu, \sigma)$ is a measure-preserving system. The map σ is called the *Bernoulli shift*.

2.4. **Recurrence.** One of the central themes in ergodic theory is *recurrence*, which concerns how points in measurable dynamical systems return close to themselves under iterations.

Theorem 2.18 (Poincaré recurrence)**.** *Let $T\colon X \to X$ be a measure-preserving transformation on a probability space $(X, \mathcal{B}, \mu)$, and let $E \in \mathcal{B}$ be a measurable set with $\mu(E) > 0$. Then almost every point $x \in E$ returns to E infinitely many often under iterations of T. More precisely, there exists a measurable set $F \subseteq E$ such that $\mu(F) = \mu(E)$, and for every point $x \in F$ the sequence of points $\{T^n(x)\}_{n=1}^{\infty}$ returns to E infinitely many times.*

Proof. Let

$$B = \{x \in E \mid T^n(x) \notin E \text{ for all } n \geq 1\}.$$

It is an easy exercise to show that B is measurable. Using the definition of B, one can show that the sets $B, T^{-1}B, T^{-2}B, \ldots$ are pairwise disjoint. Hence

$$\sum_{k=0}^{\infty} \mu(T^{-k}B) = \mu\left(\bigcup_{k=0}^{\infty} T^{-k}B\right) \leq \mu(X) = 1.$$

Therefore we have $\mu(B) = 0$, since T is measure-preserving.

Observe that the points of the union

$$\bigcup_{k=0}^{\infty} (T^{-k}B \cap E)$$

are precisely those points of E which do not return to E infinitely many often. Therefore, it suffices to show that the measure of the above union is zero.

$$\mu\left(\bigcup_{k=0}^{\infty} (T^{-k}B \cap E)\right) \leq \mu\left(\bigcup_{k=0}^{\infty} T^{-k}B\right) = \sum_{k=0}^{\infty} \mu(T^{-k}B) = 0$$

since $\mu(B) = 0$ and T is measure-preserving. □

Remark 2.19. The key step of the proof is to show that $\mu(B) = 0$, which is essentially the pigeon-hole principle: the sets $B, T^{-1}B, T^{-2}B, \ldots$ are disjoint and with the same measure, so they can not fit into a space of finite measure ($\mu(X) = 1$) unless $\mu(B) = 0$. The recurrence property does not hold for spaces of infinite measure (can you give an example?).

Remark 2.20. If one further assumes that the map $T\colon X \to X$ is *ergodic*, then one can show that the *frequency* of return to the set E is precisely $\mu(E) > 0$.

2.5. **Lebesgue integral.**

Definition 2.21. Let $(X, \mathcal{B}, \mu)$ be a probability space. A function $f\colon X \to \mathbb{R}$ is called *measurable* if $f^{-1}(A) \in \mathcal{B}$ for any (Borel) measurable set $A \subseteq \mathbb{R}$.

We would like to define the *(Lebesgue) integral* $\int f \,\mathrm{d}\mu$ of measurable functions f. First, a function $g\colon X \to \mathbb{R}$ is called *simple* if

$$g(x) = \sum_{j=1}^{m} c_j \chi_{A_j}(x)$$

for some constants $c_j \in \mathbb{R}$ and *disjoint* measurable sets $A_j \in \mathcal{B}$. In this case, the integral of g is defined to be

$$\int g \,\mathrm{d}\mu = \sum_{j=1}^{m} c_j \mu(A_j).$$

Second, one can show that for any *non-negative* measurable function $f\colon X \to \mathbb{R}_{\geq 0}$, there exists a pointwise increasing sequence of simple functions $(g_n)_{n\geq 1}$ which converges to g_n pointwisely converges to f. This allows us to define

$$\int f \,\mathrm{d}\mu = \lim_{n\to\infty} \int g_n \,\mathrm{d}\mu.$$

A non-negative measurable function $f\colon X \to \mathbb{R}_{\geq 0}$ is called *integrable* if $\int f \,\mathrm{d}\mu < \infty$.

Finally, for a general measurable function $f\colon X \to \mathbb{R}$, one can decompose it into $f = f^+ - f^-$ where $f^+(x) = \max\{f(x), 0\}$. Both f^+, f^- are non-negative measurable functions. The function f is called *integrable* if both f^+, f^- are integrable, and its integral is defined to be

$$\int f \,\mathrm{d}\mu = \int f^+ \,\mathrm{d}\mu - \int f^- \,\mathrm{d}\mu.$$

Notation. Let $(X, \mathcal{B}, \mu)$ be a measurable space. Define

$$L^1_\mu = \left\{ f\colon X \to \mathbb{R} : f \text{ is measurable and } \|f\|_1 \coloneqq \int |f| \,\mathrm{d}\mu < \infty \right\}.$$

Similarly, define

$$L^2_\mu = \left\{ f\colon X \to \mathbb{R} : f \text{ is measurable and } \|f\|_2 \coloneqq \left(\int |f|^2 \,\mathrm{d}\mu \right)^{1/2} < \infty \right\}.$$

The following theorem provides an important characterization of measure-preserving maps.

Theorem 2.22. *Let $(X, \mathcal{B}, \mu)$ be a probability space. A map $T\colon X \to X$ is measure-preserving if and only if*

$$\int f \,\mathrm{d}\mu = \int f \circ T \,\mathrm{d}\mu \qquad \text{for all } f \in L^1_\mu.$$

Proof. First, we prove the "if" part. Take $f = \chi_B$ for any $B \in \mathcal{B}$, one gets

$$\mu(T^{-1}B) = \int \chi_{T^{-1}B} \,\mathrm{d}\mu = \int \chi_B \circ T \,\mathrm{d}\mu = \int \chi_B d\mu = \mu(B).$$

Conversely, if T is measure-preserving, then the integral equality holds for any simple functions. For any $f \in L^1_\mu$, one can take an increasing sequence (f_n) of simple functions such that $\lim f_n = f$ pointwise. Hence we also have $\lim f_n \circ T = f \circ T$. By dominated convergence theorem,

$$\int f \,\mathrm{d}\mu = \lim_{n\to\infty} \int f_n \,\mathrm{d}\mu = \lim_{n\to\infty} \int f_n \circ T \,\mathrm{d}\mu \int f \circ T \,\mathrm{d}\mu$$

□

Remark 2.23. The Lebesgue integral is more general than the *Riemann integral*: The Lebesgue integral allows a countable infinity of discontinuities, while Riemann integral allows only a finite number of discontinuities. As an example, consider the set $A = \mathbb{Q} \cap [0, 1]$ of rational numbers in $[0, 1]$. It is an easy exercise of Riemann integral to show that the characteristic function $\chi_A\colon [0, 1] \to \mathbb{R}$ is not integrable. On the other hand, the set A is measurable and its Lebesgue measure is $\lambda(A) = 0$. Therefore, χ_A is Lebesgue measurable and

$$\int \chi_A \,\mathrm{d}\lambda = \lambda(A) = 0.$$

Lecture 2

2.6. **Ergodicity.**

Definition 2.24. Let $(X, \mathcal{B}, \mu)$ be a probability space. A measure-preserving transformation $T\colon X \to X$ is said to be *ergodic* if for any $B \in \mathcal{B}$,

$$T^{-1}B = B \implies \mu(B) = 0 \text{ or } \mu(B) = 1.$$

In words, it is impossible to split X into T-invariant subsets of positive measures.

Non-example. Consider the rotation map $R_\alpha(x) = x + \alpha \pmod 1$ on the circle S^1. It is not hard to show that if α is rational then R_α is not ergodic. For instance, when $\alpha = \frac{1}{2}$, the set $B = (0, \frac{1}{4}) \cup (\frac{1}{2}, \frac{3}{4})$ satisfies $R_\alpha^{-1} B = B$ but $\mu(B) = \frac{1}{2}$. We will see later that if α is irrational then R_α is ergodic.

Example 2.25. Let us show that the *Bernoulli shifts* σ are ergodic. First, we claim that the Bernoulli shifts are *mixing*, i.e.

$$\lim_{n\to\infty} \mu(B \cap \sigma^{-n} B') = \mu(B)\mu(B') \qquad \text{for all } B, B' \in \mathcal{B}.$$

It is easy to see that the statement is true if B and B' are both finite unions of cylinder sets. By Kolmogorov extension theorem (which we will not discuss here), for any measurable set B and any $\epsilon > 0$, there exists a finite union of cylinder sets A such that $\mu(A \Delta B) < \epsilon$. (Here $A \Delta B \coloneqq (A \backslash B) \cup (B \backslash A)$.) It is then an easy exercise to show the mixing property.

Second, we claim that mixing implies ergodic. Let $B = \sigma^{-1} B$ be a measurable σ-invariant set. By the mixing property, we have

$$\mu(B) = \lim_{n\to\infty} \mu(B \cap \sigma^{-n} B) = \mu(B)^2.$$

Hence $\mu(B) \in \{0, 1\}$.

Remark 2.26. As the proof above suggests, the concept of ergodicity is closely related to the idea of *mixing*, meaning, given a measurable set $A \subseteq X$, how the set $T^{-n} A$ is spread around the whole space X under large iterations n?

It can be proved that a measure-preserving system $(X, \mathcal{B}, \mu, T)$ is ergodic if and only if it is *weak-mixing* (a weaker condition than *mixing*), i.e.

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mu(A \cap T^{-n} B) = \mu(A)\mu(B) \text{ for all } A, B \in \mathcal{B}.$$

A proof of this fact can be found in [4, Section 2.7].

The following theorem is very useful for proving a system is ergodic (or non-ergodic).

Theorem 2.27. *For a measure-preserving system $(X, \mathcal{B}, \mu, T)$, the following are equivalent.*

(a) T is ergodic.

(b) *For any $f\colon X \to \mathbb{R}$ measurable, if $f \circ T = f$ almost everywhere, then f is constant almost everywhere.*

Proof. It is easy to see that (b) implies (a): Suppose $T^{-1}B = B$. Take $f = \chi_B$. Then we have χ_B is constant almost everywhere, thus $\mu(B) \in \{0, 1\}$.

Conversely, suppose T is ergodic, and $f\colon X \to \mathbb{R}$ be a measurable function such that $f \circ T = f$ almost everywhere. By the following exercise, one can redefine f on a set of measure zero such that the redefined function, which we still call f, is T-invariant *everywhere*. We would like to show that f is constant almost everywhere. Consider the level sets

$$A_t = \{x \in X \mid f(x) > t\}, \qquad t \in \mathbb{R}.$$

Then

$$\begin{aligned} T^{-1}A_t &= \{x \in X \mid T(x) \in A_t\} \\ &= \{x \in X \mid f(T(x)) > t\} \\ &= \{x \in X \mid f(x) > t\} \\ &= A_t. \end{aligned}$$

Since T is ergodic, for each t we have $\mu(A_t) \in \{0, 1\}$. It is then easy to show that f has to be constant almost everywhere. □

Exercise. Let $(X, \mathcal{B}, \mu, T)$ be a measure-preserving system. Let $f\colon X \to \mathbb{R}$ be a measurable function that is T-invariant almost everywhere, i.e. $f \circ T(x) = f(x)$ holds for μ-almost every $x \in X$.

(a) Consider the set

$$E = \bigcup_{k \geq 0} T^{-k}(A), \qquad \text{where } A = \{x \in X \mid f(T(x)) \neq f(x)\}.$$

Prove that $\mu(E) = 0$ and that $T^{-1}(E) \subseteq E$.

(b) Define a new function $\tilde{f}\colon X \to \mathbb{R}$ by

$$\tilde{f}(x) = \begin{cases} f(x) & \text{if } x \notin E, \\ 0 & \text{if } x \in E. \end{cases}$$

Prove that $f = \tilde{f}$ almost everywhere, and that $\tilde{f} \circ T = \tilde{f}$ *everywhere*.

Remark 2.28. One can show that in the characterization theorem above, instead of considering all measurable functions, it is enough to consider only the integrable functions $f \in L^1_\mu$ or the square-integrable functions $f \in L^2_\mu$. More precisely, for a measure-preserving system $(X, \mathcal{B}, \mu, T)$, the following statements are all equivalent:

(a) T is ergodic.
(b) For any $f\colon X \to \mathbb{R}$ measurable, if $f \circ T = f$ almost everywhere, then f is constant almost everywhere.
(c) For any $f \in L^1_\mu$, if $f \circ T = f$ almost everywhere, then f is constant almost everywhere.
(d) For any $f \in L^2_\mu$, if $f \circ T = f$ almost everywhere, then f is constant almost everywhere.

Using this remark and some basic knowledge of *Fourier series*, one can easily show that the rotation maps and the doubling map of S^1 are ergodic. Let $f\colon S^1 \cong \mathbb{R}/\mathbb{Z} \to \mathbb{R}$ be a square-integrable function, i.e. $f \in L^2(S^1)$. Results of Fourier series imply that there exists a *unique* collection of complex numbers $\ldots, c_{-2}, c_{-1}, c_0, c_1, c_2, \ldots$, called the *Fourier coefficients* of f, such that

$$f(x) = \sum_{n\in\mathbb{Z}} c_n e^{2\pi i n x} \qquad \text{for a.e. } x \in \mathbb{R}/\mathbb{Z}.$$

Moreover, we have $\|f\|_2 = \sum_{n\in\mathbb{Z}} |c_n|^2 < \infty$.

Example 2.29. Consider the rotation map $R_\alpha(x) = x + \alpha \pmod 1$ on the circle S^1 where α is irrational. By Remark 2.28, it suffices to show that for any $f \in L^2(S^1)$, if $f \circ R_\alpha = f$ almost everywhere, then f is constant almost everywhere. Let the Fourier series of f be $\sum_{n\in\mathbb{Z}} c_n e^{2\pi i n x}$. Then

$$\left(\sum_{n\in\mathbb{Z}} c_n e^{2\pi i n x}\right) \circ R_\alpha = \sum_{n\in\mathbb{Z}} c_n e^{2\pi i n (x+\alpha)} = \sum_{n\in\mathbb{Z}} c_n e^{2\pi i n \alpha} e^{2\pi i n x}.$$

By the uniqueness of the Fourier coefficients, we have

$$c_n \left(1 - e^{2\pi i n \alpha}\right) = 0 \qquad \text{for all } n \in \mathbb{Z}.$$

Suppose α is irrational, then $1 - e^{2\pi i n\alpha} \neq 0$ for all $n \in \mathbb{Z}\backslash\{0\}$, thus we have $c_n = 0$ for all $n \in \mathbb{Z}\backslash\{0\}$. Hence $f(x) = \sum_{n\in\mathbb{Z}} c_n e^{2\pi i n x} = c_0$ is constant almost everywhere. (Can you identify at where this argument fails for α rational?)

Example 2.30. We show that the doubling map $T_2\colon S^1 \to S^1$ is ergodic. Let $f \in L^2(S^1)$ with $f \circ T = f$ almost everywhere. Let $\sum_{n\in\mathbb{Z}} c_n e^{2\pi i n x}$ be the Fourier series of f, where $\|f\|_2^2 = \sum_{i\in\mathbb{Z}} |a_i|^2 < \infty$. Then

$$\left(\sum_{n\in\mathbb{Z}} c_n e^{2\pi i n x}\right) \circ T_2 = \sum_{n\in\mathbb{Z}} c_n e^{2\pi i n (2x)} = \sum_{n\in\mathbb{Z}} c_n e^{2\pi i (2n) x}$$

By the uniqueness of the Fourier coefficients, we have $c_n = c_{2n}$ for all $n \in \mathbb{Z}$. This implies that $c_n = 0$ for all $n \neq 0$ since $\|f\|_2^2 < \infty$. Hence f is a constant function almost everywhere.

2.7. **Ergodic theorems.** Let X be the *phase space* of a physical system (e.g. the points of X can represent configurations of positions and velocities of particles in a box). A measurable function $f\colon X \to \mathbb{R}$ represents an *observable* of the system, i.e. a quantity that can be measured (e.g. velocity, temperature, position, etc.). The value $f(x)$ is the measurement of the observable f that one gets when the system is in the state x. *Time evolution* of the system, if measured by discrete time units, can be given by a transformation $T\colon X \to X$, so that if $x \in X$ is the initial state of the system, then $T(x)$ is the state of the system after one time unit. The map T is measure-preserving if the system is in equilibrium.

In order to measure a physical quantity, one usually measures repeatedly in time and consider their average. The average of the first n measurements is given by

$$\frac{1}{n}\sum_{j=0}^{n-1} f(T^j x).$$

This quantity is called the *time average*. On the other hand, the *space average* of the observable f is simply

$$\int f \,\mathrm{d}\mu.$$

In physics, one would like to know the space average of the observable; but since experimentally it is easier to compute the time average, it is natural to ask whether the time average gives a good approximation of the space average as $n \to \infty$.

Boltzmann's Hypothesis was that for almost every initial state $x \in X$ the time averages of any observable f converge to the space average as time tends

to infinity. Unfortunately, this is not true for general measure-preserving map T. On the other hand, *under the assumption that T is ergodic*, the conclusion of Boltzmann's Hypothesis is true, and this is exactly the content of Birkhoff's ergodic theorem. Finding the right condition under which Boltzmann's Hypothesis holds motivated the definition of ergodicity, and gave birth to the study of ergodic theory.

Theorem 2.31. *Let $(X, \mathcal{B}, \mu, T)$ be a measure-preserving system on a probability space, and let $f\colon X \to \mathbb{R}$ be an integrable function.*

(a) The limit

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} f(T^j x) = f^*(x)$$

converges almost everywhere to a T-invariant integrable function f^*, where

$$\int f^* \, \mathrm{d}\mu = \int f \, \mathrm{d}\mu.$$

(b) Moreover, if T is ergodic, then

$$f^*(x) = \int f \, \mathrm{d}\mu$$

almost everywhere.

A proof of the theorem can be found in [4, Section 2.6]. Note that the second part of the statement is an easy corollary of the first part using Theorem 2.27.

Remark 2.32. Note that for an ergodic system $(X, \mathcal{B}, \mu, T)$ and a measurable function $f\colon X \to \mathbb{R}$, the ergodic theorem only guarantees the limit

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} f(T^j x) = \int f \, \mathrm{d}\mu$$

almost everywhere; the equality may not be satisfied by *every* points of X. For instance, consider the doubling map $T_2\colon S^1 \to S^1$ which is ergodic. Choose any measurable function $f\colon S^1 \to \mathbb{R}$ such that $\int f \, \mathrm{d}\mu \neq f(0)$. Then the above equality is not satisfied at the point $x = 0 \in S^1$.

Example 2.33 (Frequency of visits). Let $(X, \mathcal{B}, \mu, T)$ be a measure-preserving ergodic system, and let $A \subseteq X$ be measurable set with $\mu(A) > 0$. We would like to understand the frequency of visits:

$$\frac{\#\{0 \le k \le n-1 \mid T^k(x) \in A\}}{n} = \frac{1}{n} \sum_{k=0}^{n-1} \chi_A(T^k(x)).$$

Applying Birkhoff's pointwise ergodic theorem to $f = \chi_A$, one gets

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \chi_A(T^k(x)) = \int \chi_A \, \mathrm{d}\mu = \mu(A).$$

2.8. **Back to continued fractions.**

Definition 2.34. A *continued fraction* is an expression of the form

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cdots}}},$$

denotes alternatively by $[a_0; a_1, a_2, a_3, \ldots]$, where $a_0 \in \mathbb{Z}_{\ge 0}$ and $a_n \in \mathbb{Z}_{>0}$ for all $n \ge 1$. This expression can be finite (when the represented number is rational) or infinite (when the represented number is irrational).

Exercise. Fix a sequence $(a_n)_{n\ge 0}$ where $a_0 \in \mathbb{Z}_{\ge 0}$ and $a_n \in \mathbb{Z}_{>0}$ for all $n \ge 1$. Denote the partial expressions as

$$\frac{p_n}{q_n} = [a_0; a_1, \ldots, a_n]$$

where p_n, q_n are coprime positive integers. Then they satisfy the recursive relation

$$\begin{pmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix}.$$

Therefore, we have

$$p_{n+1} = a_{n+1} p_n + p_{n-1}, \qquad q_{n+1} = a_{n+1} q_n + q_{n-1}.$$

Also, by taking the determinants of the matrix equation, we get

$$p_n q_{n-1} - p_{n-1} q_n = (-1)^{n+1}.$$

Hence

$$\begin{aligned}\frac{p_n}{q_n} &= \frac{p_{n-1}}{q_{n-1}} + (-1)^{n+1}\frac{1}{q_{n-1}q_n} \\ &= a_0 + \frac{1}{q_0q_1} - \frac{1}{q_1q_2} + \cdots + (-1)^{n+1}\frac{1}{q_{n-1}q_n}\end{aligned}$$

by induction, and show that

$$x = \lim_{n\to\infty}[a_0; a_1, \ldots, a_n] = \lim_{n\to\infty}\frac{p_n}{q_n} = a_0 + \sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{q_{n-1}q_n}.$$

Moreover, we have

$$\frac{p_0}{q_0} < \frac{p_2}{q_2} < \cdots < \frac{p_{2n}}{q_{2n}} < \cdots < x < \cdots < \frac{p_{2n+1}}{q_{2n+1}} < \cdots < \frac{p_3}{q_3} < \frac{p_1}{q_1}.$$

The rational numbers $\frac{p_n}{q_n}$ are called the *convergents* of the continued fraction for x, and they provide very rapid rational approximation to x. We have

$$\left|x - \frac{p_n}{q_n}\right| < \frac{1}{q_nq_{n+1}}.$$

The numbers q_n and p_n grow exponentially as $n \to \infty$: using the recursive relation, one can show that both p_n and q_n are greater than $2^{(n-2)/2}$.

In fact, the continued fraction convergents provide the *optimal* rational approximants of an irrational number in the following sense.

Proposition 2.35. *Let $x > 0$ be an irrational number, $[a_0; a_1, \ldots]$ be its associated continued fraction, and $\frac{p_n}{q_n}$ be its convergents defined above. For any $1 \le q < q_n$ and any $p_n > 0$, we have*

$$\left|x - \frac{p_n}{q_n}\right| < \left|x - \frac{p}{q}\right|.$$

Definition 2.36. Define the *continued fraction map* $T\colon [0,1] \to [0,1]$ by $T(0) = 0$ and

$$T(x) = \frac{1}{x} - \left\lfloor\frac{1}{x}\right\rfloor \text{ for } x \neq 0,$$

where $\lfloor t\rfloor$ denotes the greatest integer less than or equal to t. In other words, $T(x)$ is the fractional part $\{\frac{1}{x}\}$ of $\frac{1}{x}$.

For our purpose, we would like to find a measure on $[0,1]$ such that the continued fraction map T is measure-preserving. Unfortunately, the usual Lebesgue measure on $[0,1]$ does not work. For instance,

$$T^{-1}\left(0,\frac{1}{2}\right)=\left(\frac{2}{3},1\right)\cup\left(\frac{2}{5},\frac{1}{2}\right)\cup\left(\frac{2}{7},\frac{1}{3}\right)\cup\cdots,$$

which has measure strictly greater than $1/2$ with respect to the standard Lebesgue measure.

Definition 2.37. Define the *Gauss measure* μ on $[0,1]$ to be

$$\mu(A)=\frac{1}{\log 2}\int_A\frac{1}{1+x}dx \text{ for any measurable set } A\subseteq[0,1].$$

Exercise. The Gauss measure is "comparable" with the standard Lebesgue measure λ on $[0,1]$: Show that

$$\frac{\lambda(B)}{2\log 2}\leq\mu(B)\leq\frac{\lambda(B)}{\log 2}\qquad\text{for any measurable set } B\subseteq[0,1].$$

Proposition 2.38. *The continued fraction map T preserves the Gauss measure μ.*

Proof. It suffices to show it for $A=[0,b]$ for all $b>0$. Observe that

$$T^{-1}[0,b]=\bigcup_{n=1}^{\infty}\left[\frac{1}{b+n},\frac{1}{n}\right].$$

It is an easy exercise to show that

$$\begin{aligned}\mu(T^{-1}[0,b])&=\frac{1}{\log 2}\sum_{n=1}^{\infty}\int_{\frac{1}{b+n}}^{\frac{1}{n}}\frac{1}{1+x}dx\\&=\frac{1}{\log 2}\int_0^b\frac{1}{1+x}dx\\&=\mu([0,b]).\end{aligned}$$

□

We now move on to prove the *ergodicity* of the continued fraction map T with respect to the Gauss measure. Notice that in terms of the continued fraction expansion, T behaves similar to the shift map in that

$$T([a_1,a_2,\ldots])=[a_2,a_3,\ldots].$$

We therefore would like to pursue a method of proof similar to the proof of the ergodicity of Bernoulli shifts: we want to control the size of the *cylinder sets* and their *intersections*.

Exercise. Given an n-tuple $a = (a_1, \ldots, a_n) \in \mathbb{Z}_{>0}^n$ of positive integers, define the cylinder set

$$I(a) = \{[x_1, x_2, \ldots] \mid x_i = a_i \text{ for } 1 \leq i \leq n\} \subseteq [0, 1].$$

- $I(a)$ is a subinterval of $[0, 1]$ with length $\frac{1}{q_n(q_n+q_{n-1})}$, where $\frac{p_n}{q_n}$ is the convergent of $[a_1, \ldots, a_n]$.
- Since $q_n \geq 2^{(n-2)/2}$, the length of $I(a) = I([a_1, \ldots, a_n])$ shrinks to zero as $n \to \infty$. Use this to show that the cylinder sets $I(a)$ for all possible strings of positive integers generate the Borel σ-algebra on $[0, 1]$.

Proposition 2.39. *The continued fraction map T on $[0, 1]$ is ergodic with respect to the Gauss measure μ.*

Proof. The key step of the proof is to show that

$$\mu(T^{-n}A \cap I(a)) \asymp \mu(A)\mu(I(a)) \qquad \text{for any measurable set } A, \tag{2.1}$$

i.e. there exist constants $C_1, C_2 > 0$ which are independent of the choice of A (but may depend on $I(a)$), such that

$$C_1\mu(T^{-n}A \cap I(a)) \leq \mu(A)\mu(I(a)) \leq C_2\mu(T^{-n}A \cap I(a)).$$

We first prove that T is ergodic assuming (2.1). Let $B \subseteq [0, 1]$ be a measurable set with $T^{-1}B = B$. By (2.1) we have

$$\mu(B \cap I(a)) \asymp \mu(B)\mu(I(a)).$$

Since the cylinder sets generate the Borel σ-algebra of A, we have

$$\mu(B \cap A) \asymp \mu(B)\mu(A) \qquad \text{for any measurable set } A.$$

By applying this to $A = X \backslash B$, we obtain $\mu(B)\mu(X \backslash B) = 0$, which concludes the proof.

We now proceed to prove (2.1). Recall that the Gauss measure μ is comparable with the Lebesgue measure λ, thus it suffices to show

$$\lambda(T^{-n}A \cap I(a)) \asymp \lambda(A)\lambda(I(a)) \qquad \text{for any measurable set } A$$

As usual, it suffices to show it for any interval $A = [d, e]$. It is an exercise to show that $T^{-n}A \cap I(a)$ is an interval with endpoints given by

$$\frac{p_n + p_{n-1}d}{q_n + q_{n-1}d} \qquad \text{and} \qquad \frac{p_n + p_{n-1}e}{q_n + q_{n-1}e}.$$

Therefore

$$\begin{aligned}
\lambda(T^{-n}A \cap I(a)) &= \frac{e-d}{(q_n + q_{n-1}d)(q_n + q_{n-1}e)} \\
&= \lambda(A)\lambda(I(a))\frac{q_n(q_n + q_{n-1})}{(q_n + q_{n-1}d)(q_n + q_{n-1}e)} \\
&\asymp \lambda(A)\lambda(I(a)).
\end{aligned}$$

□

Example 2.40. This answers our motivating question: By applying Birkhoff's pointwise ergodic theorem, we have

$$\begin{aligned}
\lim_{n\to\infty} \frac{\#\{i \mid a_i = k, 1 \le i \le n\}}{n} &= \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \chi_{(\frac{1}{k+1}, \frac{1}{k}]}(T^i(x)) \\
&= \int \chi_{(\frac{1}{k+1}, \frac{1}{k}]} \, \mathrm{d}\mu \\
&= \mu\left(\left(\frac{1}{k+1}, \frac{1}{k}\right]\right) \\
&= \frac{1}{\log 2} \int_{\frac{1}{k+1}}^{\frac{1}{k}} \frac{1}{1+x} \, \mathrm{d}x = \frac{1}{\log 2} \log\left(\frac{(k+1)^2}{k(k+2)}\right)
\end{aligned}$$

for almost every $x \in (0, 1)$.

Example 2.41. The following result also is an application of the pointwise ergodic theorem: for almost every $x \in (0, 1)$, the rate of approximation of the continued fractions is given by

$$\lim_{n\to\infty} \frac{1}{n} \log \left| x - \frac{p_n(x)}{q_n(x)} \right| = \frac{-\pi^2}{6 \log 2}.$$

3. Topology

3.1. **The Borsuk–Ulam theorem.** Let us consider the following *continuous* version of the necklace splitting problem. We say a probability measure μ on $[0,1]$ is *continuous* if $\int_0^x \mathrm{d}\mu$ is continuous in x.

Question 3.1. *Let $\mu_1, \dots, \mu_n$ be continuous probability measures on $[0,1]$. Does there exist a partition of $[0,1]$ into $n+1$ intervals $I_0, \dots, I_n$ and signs $\epsilon_0, \dots, \epsilon_n \in \{\pm 1\}$ such that*

$$\sum_{j=0}^{n} \epsilon_j \cdot \mu_i(I_j) = 0 \quad \text{for all } 1 \le i \le n \text{ ?}$$

Remark 3.2. In the original necklace splitting problem, the n measures μ_i corresponds to the n kinds of precious stones, the interval $[0,1]$ is separated into $n+1$ subintervals by n cuts, and the signs ± 1 determine the corresponding portion of the necklace belongs to which one of the two thieves.

An affirmative answer to the above continuous version would imply an affirmative answer to the original necklace splitting problem. For more details, cf. [6].

There is a clever way to encode the divisions of the necklace by points of the *n-dimensional sphere* S^n. With every point of the sphere

$$S^n = \left\{(x_0, \dots, x_n) \in \mathbb{R}^{n+1} \mid x_0^2 + \cdots + x_n^2 = 1\right\}$$

we associate a division of the interval $[0,1]$ into $n+1$ parts, of lengths $x_0^2, \dots, x_n^2$; i.e. we cut the interval at the points $0 = z_0 \le z_1 \le \cdots z_n \le z_{n+1} \le 1$. The sign ϵ_j for the j-th interval $[z_{j-1}, z_j]$ is chosen as $\operatorname{sign}(x_j)$. This defines a continuous map $g\colon S^n \to \mathbb{R}^n$, where its i-th component is given by

$$g_i(x) = \sum_{j=0}^{n} \operatorname{sign}(x_j) \cdot \mu_i\left([z_{j-1} - z_j]\right).$$

The function g clearly satisfies $g(-x) = -g(x)$ for all $x \in S^n$. We would like to show that $g(x) = 0$ for some $x \in S^n$. It follows directly from the *Borsuk–Ulam theorem.*

Theorem 3.3 (Borsuk–Ulam)**.** *Let $f\colon S^n \to \mathbb{R}^n$ be a continuous map. Then there exists an $x \in S^n$ such that $f(-x) = f(x)$.*

Lecture 3

For instance, the case $n = 2$ can be illustrated by saying that at any moment, there is always a pair of antipodal points on the Earth's surface with equal temperatures and equal pressures.

Exercise. For any $n \geq 1$, the following statements are equivalent:

- For every continuous map $f\colon S^n \to \mathbb{R}^n$ there exists a point $x \in S^n$ such that $f(-x) = f(x)$.
- For every *antipodal* continuous map $f\colon S^n \to \mathbb{R}^n$ (antipodal means $f(-x) = -f(x)$ for all $x \in S^n$), there exists $x \in S^n$ such that $f(x) = 0$.
- There is no antipodal map $f\colon S^n \to S^{n-1}$.
- There is no continuous map $f\colon B^n \to S^{n-1}$ that is antipodal on the boundary, i.e. satisfies $f(-x) = -f(x)$ for all $x \in S^{n-1} = \partial B^n$.

Remark 3.4. As a direct corollary, there is no continuous map $f\colon B^n \to S^{n-1}$ that is the *identity* on the boundary $\partial B^n = S^{n-1}$, which implies the *Brouwer fixed point theorem.*

As an another corollary of the Borsuk–Ulam theorem, one can show the following *ham sandwich theorem.* The informal statement that gave the ham sandwich theorem its name is this: "For every sandwich made of ham, cheese, and bread, there is a planar cut that simultaneously halves the ham, the cheese, and the bread."

Theorem 3.5 (Ham sandwich theorem)**.** *For any compact sets $A_1, \ldots, A_n \subseteq \mathbb{R}^n$, there exists a hyperplane dividing each of them into two subsets of equal measure.*

One can prove a more general version of ham sandwich theorem in terms of measures. We say a measure on $\mathbb{R}^n$ is a *finite Borel measure* if all open subsets of $\mathbb{R}^n$ are measurable and $0 < \mu(\mathbb{R}^n) < \infty$. For instance, for any compact set $A \subseteq \mathbb{R}^n$, one can define a finite Borel measure μ_A by $\mu_A(X) \coloneqq \lambda(X \cap A)$.

Theorem 3.6 (Ham sandwich theorem for measures)**.** *For any finite Borel measures $\mu_1, \ldots, \mu_n$ on $\mathbb{R}^n$, there exists a hyperplane h such that*

$$\mu_i(h^+) = \frac{1}{2}\mu_i(\mathbb{R}^n) \quad \text{for } 1 \leq i \leq n$$

where h^+ denotes one of the half-spaces defined by h.

Proof. Let $u = (u_0, \dots, u_n)$ be a point of the sphere $u \in S^n$. If at least one of the components $u_1, \dots, u_n$ is nonzero, we assign u the half-space

$$h^+(u) = \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid u_1 x_1 + \cdots + u_n x_n \le u_0\}.$$

It is clear that antipodal points of S^n correspond to opposite half-spaces. For $u = (\pm 1, 0, \dots, 0) \in S^n$, we have by the same formula

$$h^+((+1, 0, \dots, 0)) = \mathbb{R}^n,$$
$$h^+((-1, 0, \dots, 0)) = \emptyset.$$

Define a continuous function $f\colon S^n \to \mathbb{R}^n$ where the i-th component is

$$f_i(u) \coloneqq \mu_i(h^+(u)).$$

By the Borsuk–Ulam theorem, there exists $x \in S^n$ such that $f(-x) = f(x)$. Then the boundary of the half space $h^+(x)$ is the desired hyperplane. □

Let us discuss the proof of the Borsuk–Ulam theorem. For $n = 1$, the theorem follows easily from the intermediate value theorem. One can prove the $n = 2$ case using some basic knowledge of *fundamental groups* of topological spaces. We will be discussing this in more details in later subsections.

For $n \ge 3$, the proofs usually are more involved; let us sketch a proof here.

- Assume the contrary that there exists an antipodal map $f\colon S^n \to S^{n-1}$. This descends to a continuous map $g\colon \mathbb{RP}^n \to \mathbb{RP}^{n-1}$. Here $\mathbb{RP}^n \cong S^n/\mathbb{Z}_2$ is the n-dimensional *real projective space*.
- One can show that such g induces an isomorphism $g_*\colon \pi_1(\mathbb{RP}^n) \to \pi_1(\mathbb{RP}^{n-1})$ between the *fundamental groups*.
- By the *Poincaré–Hurewicz theorem*, we have an isomorphism $g_*\colon H_1(\mathbb{RP}^n, \mathbb{Z}) \to H_1(\mathbb{RP}^{n-1}, \mathbb{Z})$ between the *homology groups*.
- By the *universal coefficient theorem*, we have an induced *ring homomorphism* between the *cohomology rings*

  $$\mathbb{F}_2[b]/b^n \cong H^*(\mathbb{RP}^{n-1}, \mathbb{F}_2) \xrightarrow{g^*} H^*(\mathbb{RP}^n, \mathbb{F}_2) \cong \mathbb{F}_2[a]/a^{n+1}$$

  which sends $b \mapsto a$. But then we get that $b^n = 0$ is sent to $a^n \ne 0$, a contradiction.

Remark 3.7. The real projective space $\mathbb{RP}^n$ is the topological space that parametrizes the 1-dimensional subspaces of $\mathbb{R}^{n+1}$. It can be defined by quotienting the scaling action:

$$\mathbb{RP}^n = (\mathbb{R}^{n+1}\backslash\{0\})/\mathbb{R}^*.$$

Thus $\mathbb{RP}^n$ can also be formed by identifying antipodal points of S^n. It is a smooth compact manifold, and is a special case of *Grassmannians* $\mathrm{Gr}(k, n+1)$ which parametrizes the k-dimensional subspaces of $\mathbb{R}^{n+1}$.

In the following, we will introduce the notion of *fundamental groups* of topological spaces, and prove the Borsuk–Ulam theorem for $n = 2$. A nice reference in which you can find all these notions mentioned above is a book by Hatcher [5].

3.2. **Fundamental groups.** Let us start with recalling the definition of *topological spaces* and *continuous maps* between them.

Definition 3.8. A *topology* on a set X is a collection τ of subsets of X satisfying the following axioms:

- The empty set and X itself belong to τ.
- Any arbitrary (finite or infinite) union of members of τ belongs to τ.
- The intersection of any finite number of members of τ belongs to τ.

Members of τ are called *open subsets* of X (with respect to this topology).

Definition 3.9. A map $f\colon X \to Y$ between topological spaces is called *continuous* if

$$U \subseteq Y \text{ is an open subset } \implies f^{-1}(U) \subseteq X \text{ is an open subset.}$$

The map f is called a *homeomorphism* if it is bijective, and both f and f^{-1} are continuous. In this case, X and Y are said to be *homeomorphic*.

The *fundamental groups* of topological spaces will be defined in terms of *loops* and their deformations.

Definition 3.10. Let X be a topological space.

- A *path* in X is a continuous map $\gamma\colon I \to X$ where $I = [0, 1]$.
- Its *inverse path* $\gamma^{-1}\colon I \to X$ is defined by $\gamma^{-1}(t) = \gamma(1 - t)$.
- A path is called a *loop* if $\gamma(0) = \gamma(1)$. It can be considered as a map $\gamma\colon S^1 \to X$, with *basepoint* $x_0 = \gamma(0) = \gamma(1)$.

- If $\gamma(t) = x_0 \in X$ for all $t \in [0,1]$, then such γ is called a *constant path*, and denoted by i_{x_0}.
- If γ_1 and γ_2 are two loops satisfying $\gamma_1(1) = \gamma_2(0)$, we define their *composition* or *product path* to be

$$(\gamma_1 \cdot \gamma_2)(s) = \begin{cases} \gamma_1(2s), & 0 \le s \le 1/2 \\ \gamma_2(2s-1), & 1/2 \le s \le 1 \end{cases}$$

Definition 3.11. Two paths γ_0, γ_1 with the same endpoints x_0, x_1 are called *homotopic* if there exists a continuous map $F\colon I \times I \to X$ such that

- $F(s,0) = \gamma_0(s)$ and $F(s,1) = \gamma_1(s)$ for all $s \in [0,1]$.
- $F(0,t) = x_0$ and $F(1,t) = x_1$ for all $t \in [0,1]$.

In this case, we will denote $\gamma_0 \simeq \gamma_1$.

Example 3.12. Any two paths γ_0, γ_1 in $\mathbb{R}^n$ having the same endpoints x_0, x_1 are homotopic via the linear homotopy $F(s,t) = (1-t)\gamma_0(s) + t\gamma_1(s)$.

Exercise. The relation of homotopy on paths with fixed endpoints is an *equivalence relation*, i.e.

- $\gamma \simeq \gamma$.
- If $\gamma_1 \simeq \gamma_2$, then $\gamma_2 \simeq \gamma_1$.
- If $\gamma_1 \simeq \gamma_2$ and $\gamma_2 \simeq \gamma_3$, then $\gamma_1 \simeq \gamma_3$.

We denote the homotopy class of γ as $[\gamma]$.

Exercise. Let $\gamma_1, \gamma_2, \beta_1, \beta_2$ be paths in X. Suppose $\gamma_1 \simeq \gamma_2$, $\beta_1 \simeq \beta_2$, and $\gamma_1(1) = \gamma_2(1) = \beta_1(0) = \beta_2(0)$. Prove that $\gamma_1 \cdot \beta_1 \simeq \gamma_2 \cdot \beta_2$.

This shows that the *composition* (or *product*) can be defined on homotopy classes:

$$[\gamma] \cdot [\beta] \coloneqq [\gamma \cdot \beta].$$

Exercise. This exercise shows that the product on homotopy classes has *associativity*. Let $\gamma_1, \gamma_2, \gamma_3$ be paths in X satisfying $\gamma_1(1) = \gamma_2(0)$ and $\gamma_2(1) = \gamma_3(0)$. Prove that

$$([\gamma_1] \cdot [\gamma_2]) \cdot [\gamma_3] = [\gamma_1] \cdot ([\gamma_2] \cdot [\gamma_3]).$$

Note that the equality is not true without considering their homotopy classes: $(\gamma_1 \cdot \gamma_2) \cdot \gamma_3 \neq \gamma_1 \cdot (\gamma_2 \cdot \gamma_3)$ in general.

Exercise. Let γ be a path from x_0 to x_1 in X. Prove that

$$[\gamma] \cdot [\gamma^{-1}] = [i_{x_0}], \quad [\gamma^{-1}] \cdot [\gamma] = [i_{x_1}], \quad [\gamma] \cdot [i_{x_1}] = [\gamma] = [i_{x_0}] \cdot [\gamma].$$

We are now ready to define the fundamental group.

Definition 3.13. The *fundamental group* of X at the basepoint x_0, denoted by $\pi_1(X, x_0)$, is defined to be the set of all homotopy classes $[\gamma]$ of loops $\gamma\colon I \to X$ with basepoint x_0, where

- the group structure given by the product $[\gamma_1] \cdot [\gamma_2] = [\gamma_1 \cdot \gamma_2]$,
- the identity element is $[i_{x_0}]$,
- the inverse of an element $[\gamma]$ is given by $[\gamma^{-1}]$.

Example 3.14. Hold a mug in your hand. Now, without letting go of the mug and without spilling the coffee, see if you can rotate the mug *two full turns* and return your hand, arm, and cup to their original positions. If you can do that, can you do the same trick with only *one* full turn? (*No!*)

Continuously rotating a mug is equivalent to following a path in SO(3), the space of rotations in $\mathbb{R}^3$, and if you start and end the mug in the same orientation, you have traced a loop in SO(3). The reason this trick works for 2 twists but not 1 twist is because $\pi_1(\mathrm{SO}(3)) \cong \mathbb{Z}/2\mathbb{Z}$.

Proposition 3.15. *Suppose X is path-connected, i.e. for any two points $x_0, x_1 \in X$, there exists a path $\gamma\colon I \to X$ such that $\gamma(0) = x_0$ and $\gamma(1) = x_1$. Then the isomorphic class of the fundamental group $\pi_1(X, x_0)$ is independent of the choice of the basepoint x_0, i.e. for any two points $x_0, x_1 \in X$ we have $\pi_1(X, x_0) \cong \pi_1(X, x_1)$.*

Proof. Let γ be a path connecting x_0 and x_1. It is easy to check that

$$\pi_1(X, x_0) \to \pi_1(X, x_1); \qquad [\beta] \mapsto [\gamma^{-1}] \cdot [\beta] \cdot [\gamma]$$

and

$$\pi_1(X, x_1) \to \pi_1(X, x_0); \qquad [\beta] \mapsto [\gamma] \cdot [\beta] \cdot [\gamma^{-1}]$$

are group homomorphisms inverse with each other. Thus $\pi_1(X, x_0) \cong \pi_1(X, x_1)$. □

Proposition 3.16. *A continuous map $f\colon X \to Y$ induces a group homomorphism*

$$f_*\colon \pi_1(X, x_0) \to \pi_1(Y, f(x_0)); \qquad [\gamma] \mapsto [f \circ \gamma].$$

Proof. One can verify that the map preserves homotopy equivalences and compositions. The proposition then follows easily. □

3.3. **Fundamental group of a circle and applications.** Consider the circle

$$S^1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\} = \{(\cos(2\pi s), \sin(2\pi s)) \in \mathbb{R}^2 \mid s \in \mathbb{R}\}$$

and choose a basepoint $x_0 = (1, 0) \in S^1$. This subsection is devoted to prove the following theorem.

Theorem 3.17. *The fundamental group $\pi_1(S^1, x_0) \cong \mathbb{Z}$ is an infinite cyclic group generated by the homotopy class of the loop $\omega(s) = (\cos(2\pi s), \sin(2\pi s))$.*

Note that $[\omega]^n = [\omega_n]$ where $\omega_n(s) = (\cos(2\pi ns), \sin(2\pi ns))$ for all $n \in \mathbb{Z}$. The theorem is therefore equivalent to the statement that every loop in S^1 based at $(1, 0)$ is homotopic to ω_n for a unique $n \in \mathbb{Z}$.

The main idea is to compare paths in S^1 with paths in $\mathbb{R}$ via the map

$$p\colon \mathbb{R} \to S^1; \qquad s \mapsto (\cos(2\pi s), \sin(2\pi s)).$$

Consider the path $\widetilde{\omega_n}(s) = ns$ in $\mathbb{R}$, which starts at 0 and ends at ns. The relation $\omega_n = p\widetilde{\omega_n}$ is expressed by saying that $\widetilde{\omega_n}$ is a *lift* of ω_n.

Definition 3.18. Let X be a topological space. A *covering space* of X consists of a space $\widetilde{X}$ and a map $p\colon \widetilde{X} \to X$ such that: for each point $x \in X$ there is an open neighborhood U of x such that $p^{-1}(U)$ is a union of disjoint open sets each of which is mapped homeomorphically onto U by p.

Example 3.19. Here are some basic examples of covering spaces of S^1.

- The map $p\colon \mathbb{R} \to S^1$ where $s \mapsto (\cos(2\pi s), \sin(2\pi s))$ is a covering map.
- The map $S^1 \to S^1$ where $(\cos(2\pi s), \sin(2\pi s)) \mapsto (\cos(2\pi ns), \sin(2\pi ns))$ is a covering map for any nonzero integer n. In terms of complex numbers, the map can be expressed as $z \mapsto z^n$.

Exercise. Below are two basic facts about covering spaces $p\colon \widetilde{X} \to X$.

(a) For each path $f\colon I \to X$ starting at a point $x_0 \in X$ and each $\widetilde{x_0} \in p^{-1}(x_0)$, there is a unique lift $\widetilde{f}\colon I \to \widetilde{X}$ of f starting at $\widetilde{x_0}$.

(b) For each homotopy $F\colon I \times I \to X$ starting at a point $x_0 \in X$ and each $\widetilde{x_0} \in p^{-1}(x_0)$, there is a unique lifted homotopy $\widetilde{F}\colon I \times I \to \widetilde{X}$ of f starting at $\widetilde{x_0}$.

Proof of Theorem 3.17. Let $f\colon I \to S^1$ be a loop at the basepoint $x_0 = (1, 0)$. We would like to show that it is homotopic to ω_n for a unique $n \in \mathbb{Z}$. By (a) there is a unique lift $\widetilde{f}$ of the loop f starting at 0. Note that the path $\widetilde{f}$ ends at some integer n since $p\widetilde{f}(1) = f(1) = x_0$. Recall that $\widetilde{f}$ and $\widetilde{\omega_n}$ are homotopic since they can be linearly homotopic with each other in $\mathbb{R}$. Thus $[f] = [\omega_n]$.

To show that n is uniquely determined by $[f]$, suppose there is $\omega_m \simeq \omega_n$ for some $m, n \in \mathbb{Z}$. Let F be a homotopy from ω_m to ω_n. By (b) it lifts to a homotopy $\widetilde{F}$ starting at 0, therefore the endpoints of $\widetilde{\omega_m}$ and $\widetilde{\omega_n}$ coincide. Hence $m = n$. □

Remark 3.20. For a covering space $p\colon \widetilde{X} \to X$, a homeomorphism $d\colon \widetilde{X} \to \widetilde{X}$ is called a *deck transformation* if $p \circ d = p$. Together with the composition of maps, the set of deck transformation forms a group $\mathrm{Deck}(p)$. For instance, for the n-sheeted covering space $S^1 \to S^1$ given by $z \mapsto z^n$, the deck transformations are the rotations of S^1 through angles that are multiples of $2\pi/n$, so the deck transformation group is $\mathbb{Z}/n\mathbb{Z}$. Similarly, the deck transformation group of the covering space $\mathbb{R} \to S^1$ is isomorphic to $\mathbb{Z} \cong \pi_1(S^1)$.

The covering space $p\colon \mathbb{R} \to S^1$ where $s \mapsto (\cos(2\pi s), \sin(2\pi s))$ is the *universal cover* of S^1: any covering space of S^1 can be covered by the universal cover. For instance, the covering space $S^1 \xrightarrow{z^n} S^1$ can be covered by $p_n\colon \mathbb{R} \to S^1$ where $s \mapsto (\cos(2\pi s/n), \sin(2\pi s/n))$; we have $z^n \circ p_n = p$. The deck transformation group of p_n is given by $n\mathbb{Z}$. In general, there is a one-to-one correspondence:

$$\{\text{covering space of } X\} \leftrightarrow \{\text{subgroups of } \pi_1(X)\}$$

where a covering space $p\colon \widetilde{X} \to X$ corresponds to the subgroup $p_*(\pi_1(\widetilde{X}))$ of $\pi_1(X)$. Moreover, the deck transformation group of p is isomorphic to $N(p_*(\pi_1(\widetilde{X})))/p_*(\pi_1(\widetilde{X}))$, where $N(p_*(\pi_1(\widetilde{X})))$ is the normalizer subgroup of $p_*(\pi_1(\widetilde{X}))$ in $\pi_1(X)$.

Theorem 3.21 (Borsuk–Ulam in dimension 2)**.** *There is no antipodal map* $f\colon S^2 \to S^1$.

Proof. Assume the contrary that such map f exists. Define a loop η circling the equator

$$\eta\colon I \to S^2; \qquad s \mapsto (\cos(2\pi s), \sin(2\pi s), 0),$$

and consider the loop $g = f \circ \eta \colon I \to S^1$.

On the one hand, the loop η in S^2 is homotopic to a constant map, thus so is the loop g in S^1. In other words, $[g] = 0$ in $\pi_1(S^1) \cong \mathbb{Z}$.

On the other hand, since $f(-x) = -f(x)$, we have

$$g\left(s + \frac{1}{2}\right) = -g(s) \qquad \text{for all } s \in \left[0, \frac{1}{2}\right].$$

Let $\widetilde{g} \colon I \to \mathbb{R}$ be a lift of g. Then for each $s \in \left[0, \frac{1}{2}\right]$ we have

$$\widetilde{g}\left(s + \frac{1}{2}\right) = \widetilde{g}(s) + \frac{q}{2} \qquad \text{for some odd integer } q.$$

Note that q depends continuously on $s \in \left[0, \frac{1}{2}\right]$, so it must be a constant for all $s \in \left[0, \frac{1}{2}\right]$ since it is of integer value. In particular, we have

$$\widetilde{g}(1) = \widetilde{g}(0) + q.$$

Thus $[g] \neq 0$ in $\pi_1(S^1) \cong \mathbb{Z}$ since q is odd. Contradiction. □

Theorem 3.22 (Fundamental theorem of algebra)**.** *Every non-constant polynomial with complex coefficients has a root in* $\mathbb{C}$*.*

Proof. Consider a complex polynomial $p(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_0$. Assume the contrary that $p(z)$ has no roots in $\mathbb{C}$, then for each $r \geq 0$

$$f_r(s) = \frac{p(re^{2\pi is})/p(r)}{|p(re^{2\pi is})/p(r)|}$$

defines a loop in S^1 based at 1. As r varies, f_r is a homotopy of loops in S^1 based at 1. Since f_0 is the trivial loop, we have $[f_r] = 0$ in $\pi_1(S^1)$ for all $r \geq 0$.

On the other hand, for r sufficiently large, on the circle $|z| = r$ we have

$$|z^n| > (|a_0| + \cdots + |a_{n-1}|)|z^{n-1}| \geq |a_{n-1}z^{n-1} + \cdots + a_0|.$$

Thus the polynomial $p_t(z) = z^n + t(a_{n-1}z^{n-1} + \cdots + a_0)$ has no zero on the circle $|z| = r$ when $0 \leq t \leq 1$. Replacing p by p_t in the formula above and letting t go from 1 to 0, one obtains a homotopy from the loop f_r to the loop $\omega_n(s) = e^{2\pi ins}$, thus $[f_r] = [\omega_n]$ in $\pi_1(S^1)$. We then conclude that $n = 0$. □

3.4. **The rectangular peg problem.** Let $C \subseteq \mathbb{R}^2$ be a continuous simple closed curve. Does there always exist four points on C such that they form the vertices of a rectangle? Below is the sketch of ideas toward answering this question (affirmatively).

- Denote M the *moduli space* of unordered pairs of points in C: each (unordered) pair of points c_1, c_2 in C corresponds to a unique point in M.
- Observe that M is naturally topologically equivalent to a Möbius strip, where its boundary can be identified with the curve C.
- Define a continuous function $f_C \colon M \to \mathbb{R}^3$ which sends a pair of points $c_1 = (x_1, y_1), c_2 = (x_2, y_2)$ on the curve C to the point

$$\left(\frac{x_1 + x_2}{2}, \frac{y_1 + y_2}{2}, \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \right) \in \mathbb{R}^3$$

  where the first two coordinates give the midpoint of c_1, c_2, and the third coordinate is the distance between c_1 and c_2.
- Observe that the rectangular peg problem has an affirmative answer for a curve C if and only if f_C is not injective.
- Observe that one gets the *real projective plane* $\mathbb{RP}^2$ by gluing the Möbius strip with a disk along their boundaries.
- Assume the contrary that there exists a curve C such that f_C is injective. Then one gets an embedding of the real projective plane $\mathbb{RP}^2$ into $\mathbb{R}^3$.
- Use topological tools to show that there is no embedding of $\mathbb{RP}^2$ into $\mathbb{R}^3$. This concludes the proof.

One way to show the last statement, namely there is no embedding of $\mathbb{RP}^2$ into $\mathbb{R}^3$, is by consider the *orientability* of the real projective plane $\mathbb{RP}^2$. It is known that $\mathbb{RP}^2$ is *non-orientable*: this can be rigorously proved by computing the homology groups of $\mathbb{RP}^2$. On the other hand, assume the contrary that there exists an embedding of $\mathbb{RP}^2$ into $\mathbb{R}^3$, then the image would bound a compact region in $\mathbb{R}^3$ (by the *generalized Jordan curve theorem*). The outward-pointing normal vector field would then give an orientation of $\mathbb{RP}^2$. Contradiction.

4. Algebra

Which positive integers n can be written as the sum of two squares? To answer this question, it is convenient to consider the factorization in the *ring* of *Gaussian integers* $\mathbb{Z}[i]$:

$$n = x^2 + y^2 = (x + iy)(x - iy).$$

One would also like to study other number rings; for instance, to understand the Diophantine equation $n = x^2 - 5y^2$, one would like to do factorizations in the ring $\mathbb{Z}[\sqrt{5}]$.

It is important to be aware that not all number rings have the same properties. For instance, the ring of Gaussian integers $\mathbb{Z}[i]$ is a *Unique Factorization Domain (UFD)*, but the ring $\mathbb{Z}[\sqrt{5}]$ is not: there are factorizations

$$(3 + \sqrt{5})(3 - \sqrt{5}) = 4 = 2 \cdot 2$$

where $3 \pm \sqrt{5}$ and 2 are all *irreducible* elements of $\mathbb{Z}[\sqrt{5}]$, so there are two truly different factorizations of 4 in $\mathbb{Z}[\sqrt{5}]$.

We will begin our discussions with the general notion of *rings*, then gradually specialized to commutative rings, integral domains, unique factorization domains, principal ideal domains, Euclidean domains. It turns out that the ring of Gaussian integers $\mathbb{Z}[i]$ is an *Euclidean domain* (a condition stronger than UFD), which will allow us to completely classify the integers that can be written as the sum of two squares. A nice reference for this part (and abstract algebra in general) is a book of Artin [1].

4.1. **Rings.**

Definition 4.1. A *ring* is a set R equipped with two binary operations $+$ (addition) and $\cdot$ (multiplication) satisfying:

(1) R is an abelian group under addition, namely:
   - $(a + b) + c = a + (b + c)$ for all $a, b, c \in R$.
   - $a + b = b + a$ for all $a, b \in R$.
   - There is an element $0 \in R$ such that $a + 0 = a$ for all $a \in R$.
   - For each $a \in R$ there exists $-a \in R$ such that $a + (-a) = 0$.

(2) R is a monoid under multiplication, namely:
   - $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.
   - There is an element $1 \in R$ such that $a \cdot 1 = a = 1 \cdot a$ for all $a \in R$.

(3) Multiplication is distributive with respect to addition, namely:
- $a \cdot (b + c) = a \cdot b + a \cdot c$ for all $a, b, c \in R$.
- $(b + c) \cdot a = b \cdot a + c \cdot a$ for all $a, b, c \in R$.

Note that the multiplication symbol $\cdot$ is often omitted: for instance, ab means $a \cdot b$.

Definition 4.2. A ring R is said to be *commutative* if $ab = ba$ for all $a, b \in R$.

Non-example. The set of 2×2 real matrices forms a ring under the standard matrix additions and multiplications. It is not commutative.

Remark 4.3. Whether a ring is commutative has profound implications on its behavior. *Commutative algebra*, the theory of commutative rings, is a major branch of ring theory. Its development has been greatly influenced by problems and ideas of *algebraic number theory* and *algebraic geometry*. If you are interested, a standard textbook on commutative algebra is [2].

Commutative rings resemble familiar number systems, and various definitions for commutative rings are designed to formalize properties of the integers.

Definition 4.4. A nonzero commutative ring R is called an *integral domain* if the product of any two nonzero elements is nonzero.

Non-example. The quotient ring $\mathbb{Z}/6\mathbb{Z}$ is a commutative ring, but is not an integral domain.

Non-example. The quotient ring $\mathbb{Z}[x]/(x^2 - 1)$ is a commutative ring, but is not an integral domain.

In order to introduce the definition of unique factorization domain, we need to define the notion of *units*.

Definition 4.5. An element $u \in R$ is called a *unit* if there exists $v \in R$ such that $uv = vu = 1$. In other words, a unit is an invertible element for the multiplication of the ring.

Example 4.6. Here are some basic examples:
- The units of $\mathbb{Z}$ are 1 and -1.
- The units of $\mathbb{Z}[i]$ are 1, -1, i, and $-i$.
- The units of $M_2(\mathbb{R})$ are all invertible matrices.
- The ring $\mathbb{Z}[\sqrt{3}]$ has infinitely many units: for instance, $(2 + \sqrt{3})$ and its powers are units of the ring. In general, the ring of integers in a number field can be determined by the *Dirichlet's unit theorem*.

Definition 4.7. An element of an integral domain R is called *irreducible* if it is not a unit, and is not the product of two non-unit elements.

Remark 4.8. An element of an integral domain R is called *prime* if, whenever $a \mid bc$ (i.e. $bc = ax$ for some $x \in R$), then $a \mid b$ or $a \mid c$. In an integral domain, every prime element is irreducible, but the converse is not true in general. For instance, in the ring $\mathbb{Z}[\sqrt{-5}]$, it can be shown that 3 is irreducible. However, it is not a prime element since

$$3 \mid (2+\sqrt{-5})(2-\sqrt{-5}) = 9$$

but 3 does not divide either of the two factors.

Definition 4.9. An integral domain R is said to be a *unique factorization domain* (or UFD for short) if every nonzero element $x \in R$ can be written as a product

$$x = up_1 \cdots p_n$$

where u is a unit and p_i's are irreducible, and this representation is unique in the following sense: If we also have

$$x = vq_1 \cdots q_m$$

where v is a unit and q_i's are irreducible, then $m = n$, and there exists a bijective map $\sigma\colon \{1, \ldots, n\} \to \{1, \ldots, n\}$ such that $p_i = w_i q_{\sigma(i)}$ for some units w_i.

Non-example. The quadratic ring $\mathbb{Z}[\sqrt{-5}]$ is an integral domain, but is not a UFD:

$$2 \cdot 3 = 6 = (1+\sqrt{-5})(1-\sqrt{5}).$$

One can show that $2, 3, 1+\sqrt{-5}, 1-\sqrt{-5}$ are all irreducible, and the only units of $\mathbb{Z}[\sqrt{-5}]$ is ± 1, therefore these truly are two different factorizations.

One important class of examples of UFDs are given by *principal ideal domains* (PID).

Definition 4.10. An ideal I of a commutative ring R is an additive subgroup of R which is closed under multiplications: more precisely,

- $(I, +)$ is a subgroup of $(R, +)$.
- For every $r \in R$ and $x \in I$, the product rx is in I.

An ideal is called *principal* if it can be generated by a single element, i.e. it is of the form $xR = \{xr \mid r \in R\}$.

Definition 4.11. An integral domain R is called a *principal ideal domain* (PID) if every ideal of R is principal.

Non-example. $\mathbb{Z}[x]$ is a UFD, but is not a PID: for instance, the ideal $\langle 2, x\rangle$ can not be generated by a single polynomial.

Theorem 4.12. *Every PID is a UFD.*

Proof. Let R be a PID. First, we show that R satisfies the *ascending chain condition* (ACC) on ideals; namely, whenever there are ideals

$$I_1 \subseteq I_2 \subseteq \cdots \subseteq I_n \subseteq \cdots$$

then there is some $N > 0$ such that $I_n = I_N$ for all $n \geq N$. Consider the union

$$I = \bigcup_{n \geq 1} I_n$$

which is also an ideal of R. Thus $I = (a)$ for some $a \in I$, and there exists $N > 0$ such that $a \in I_N$. This shows that R satisfies ACC.

Second, we show that every irreducible elements of R is prime. Let $a \in R$ be an irreducible element. Suppose $a \mid bc$ for some $b, c \in R$. We would like to show that $a \mid b$ or $a \mid c$ holds. Let us consider the ideal (a, b). Since R is PID, there exists $x \in R$ such that $(x) = (a, b)$. In particular, $a = xy$ for some $y \in R$. Since a is irreducible, x or y has to be a unit.

- If y is a unit, then $(a) = (x) = (a, b)$, thus $a \mid b$ as desired.
- If x is a unit, then $(1) = (x) = (a, b)$, so there exists $c, d \in R$ such that $ac + bd = 1$. Multiplying both sides with c, one gets $ac^2 + bcd = c$. Note that the left hand side is a multiple of a since $a \mid bc$, thus we obtain $a \mid c$.

Now we are ready to show that R is a UFD. First, we show that any nonzero nonunit element of R can be written as a product of irreducible elements. Assume the contrary that there exists nonzero nonunit element of R that cannot be written as a product of irreducibles. Denote the collection of such elements by S. Since R satisfies ACC, there exists $r \in S$ such that $(r) \nsubseteq (s)$ for any $s \in S \backslash \{r\}$. In particular, r is not irreducible, so it can be written as $r = xy$ for some nonunit elements $x, y \in R$. Since $(r) \subseteq (x)$ and $(r) \subseteq (y)$, we have $x, y \notin S$, therefore x and y both can be written as a product of irreducibles. But then we get $r = xy$ can also be written as a product of irreducibles. Contradiction.

Finally, we show that the factorization is unique. Suppose

$$a = up_1 \cdots p_n = vq_1 \cdots q_m$$

where u, v are units and p_i, q_i's are irreducibles (therefore are primes by what we proved earlier). Then $p_1 \mid vq_1 \cdots q_m$, thus it must divide some q_j. Since p_1 and q_j are both primes, they are the same up to a unit. We may continue this process and match each prime factor on both sides. □

Definition 4.13. An integral domain R is said to be a *Euclidean domain* if there exists a function $N\colon R\backslash\{0\} \to \mathbb{Z}_{\geq 0}$ (called a *norm function*) such that:

- For all nonzero elements $a, b \in R$, there exists $q, r \in R$ such that $a = qb + r$ and either $r = 0$ or $N(r) < N(b)$.
- For all nonzero elements $a, b \in R$ we have $N(a) \leq N(ab)$.

Non-example. The ring $\mathbb{Z}\left[\frac{1+\sqrt{-19}}{2}\right]$ is a PID, but is not a Euclidean domain.

Example 4.14. Here are some basic examples of Euclidean domains.

- The ring of integers $\mathbb{Z}$, with $N(a) = |a|$.
- The ring of Gaussian integers $\mathbb{Z}[i]$, with $N(a + ib) = a^2 + b^2$ (we will discuss more details later).
- The ring of polynomials $\mathbb{R}[x]$ over $\mathbb{R}$ (can be replaced by any *field*), with $N(P) = \deg(P)$.

Theorem 4.15. *Every Euclidean domain is a PID.*

Proof. Let R be a Euclidean domain. Let $I \subseteq R$ be a nonzero ideal. Then there exists a nonzero element $a \in I$ such that $N(a)$ is minimal among all elements of the ideal. We claim that $I = (a)$. For any $b \in I$, there exists $q, r \in R$ such that $b = qa + r$ where $r = 0$ or $N(r) < N(a)$. Since $a, b \in I$, we have $r \in I$, thus $N(r) \geq N(a)$ by the minimality. Therefore we have $r = 0$ and $b \in (a)$. □

4.2. **Ring of Gaussian integers.**

Definition 4.16. The norm function on the ring of Gaussian integers $\mathbb{Z}[i]$ is defined to be

$$N(a + ib) = (a + ib)(a - ib) = a^2 + b^2.$$

Exercise. Here are some basic properties of the norm function.

- $N(\alpha) = 0$ if and only if $\alpha = 0$.
- $N(\alpha\beta) = N(\alpha)N(\beta)$ for all $\alpha, \beta \in \mathbb{Z}[i]$.
- $N(\alpha) = 1$ if and only if α is a unit of $\mathbb{Z}[i]$.
- $\{1, -1, i, -i\}$ are the only units of $\mathbb{Z}[i]$.

Theorem 4.17. *$\mathbb{Z}[i]$ is a Euclidean domain.*

Proof. Let a, b be nonzero elements of $\mathbb{Z}[i]$. Observe that the set $b\mathbb{Z}[i]$ forms a lattice of squares with side length $|b| = \sqrt{N(b)}$. Then the distance between a and the lattice point closest to it (say bq) is no bigger than $|b|/\sqrt{2}$. Let $r = a - bq \in \mathbb{Z}[i]$. Then

$$N(r) = |r|^2 \leq \frac{|b|^2}{2} = \frac{N(b)}{2} < N(b).$$

□

Lemma 4.18. *If $\pi \in \mathbb{Z}[i]$ is such that $N(\pi)$ is a prime number, then π is a prime in $\mathbb{Z}[i]$.*

Proof. If $\pi = \alpha\beta$ in $\mathbb{Z}[i]$, then $N(\pi) = N(\alpha)N(\beta)$. So either $N(\alpha)$ or $N(\beta)$ is 1, which means that either α or β is a unit. □

Lemma 4.19. *Let q be a prime number with $q = 3$ (mod 4). Then q is a prime in $\mathbb{Z}[i]$.*

Proof. If $q = \alpha\beta$ in $\mathbb{Z}[i]$, then $q^2 = N(\alpha)N(\beta)$. Note that $q = N(\alpha) = a^2 + b^2$ is impossible since $q = 3 \pmod 4$. Thus either $N(\alpha)$ or $N(\beta)$ is 1. □

Lemma 4.20. *Let p be a prime number with $p = 1$ (mod 4). Then there exists a Gaussian prime π such that $p = \pi\overline{\pi}$.*

Proof. First, we claim that there exists an integer $c \in \mathbb{Z}$ such that $c^2 = -1 \pmod p$. This can be easily proved by assuming the fact that the multiplicative group $\mathbb{Z}_p^*$ of the finite field $\mathbb{Z}_p$ is cyclic. Let a be a generator of the multiplicative group $\mathbb{Z}_p^*$ (which has $p - 1$ elements), i.e.

$$\mathbb{Z}_p^* = \{1, a, a^2, \ldots, a^{p-2}\}.$$

Observe that -1 is the unique order two element of $\mathbb{Z}_p^*$, thus $a^{\frac{p-1}{2}} = -1 \pmod p$. The claim then follows from the assumption that $p = 1 \pmod 4$.

By the claim, we have $p \mid (c+i)(c-i)$ in $\mathbb{Z}[i]$. It is easy to show that p does not divide $c+i$ or $c-i$. Therefore p is *not* a Gaussian prime. Hence there exists nonunit elements $\alpha, \beta \in \mathbb{Z}[i]$ such that $p = \alpha\beta$. By comparing the norms on both sides, we obtain $N(\alpha) = N(\beta) = p$. Therefore both α and β are Gaussian primes. It is then easy to check that they are complex conjugate with each other. □

Proposition 4.21. *Up to multiplying by units, all the Gaussian primes are the following:*

- $1+i$ *(which is of norm* 2*),*
- π *and* $\overline{\pi}$*, where* $p = \pi\overline{\pi}$ *is a prime number with* $p = 1 \pmod 4$ *(the norms of* π *and* $\overline{\pi}$ *are both* p*),*
- q*, where* q *is a prime number with* $q = 3 \pmod 4$ *(which is of norm* q^2*).*

Proof. Let α be a Gaussian prime. Then we can find a Gaussian prime π in the above list so that $\pi \mid N(\alpha) = \alpha\overline{\alpha}$. So either π or $\overline{\pi}$ divides α. Thus α is also in the above list. □

4.3. **Applications.** Let us apply the arithmetic of $\mathbb{Z}[i]$ to solve a classic problem: finding all *Pythagorean triples*. A Pythagorean triples is $(x, y, z) \in \mathbb{Z}^3_{>0}$ where $x^2 + y^2 = z^2$. It suffices to only look for *primitive* Pythagorean triples, i.e. $\gcd(x, y, z) = 1$. Also, observe that x and y cannot both be odd, so may assume that x is odd and y is even.

Theorem 4.22. *Let* $(x, y, z) \in \mathbb{Z}^3_{>0}$ *be a primitive Pythagorean triples with* x *odd and* y *even. Then there exists coprime integers* a, b *with* $a > b > 0$ *and* $a \neq b \pmod 2$ *such that*

$$x = a^2 - b^2, \quad y = 2ab, \quad z = a^2 + b^2.$$

Proof. Let $\alpha = x + iy \in \mathbb{Z}[i]$, so $N(\alpha) = x^2 + y^2 = z^2$. The idea is to show that α is a *square* in $\mathbb{Z}[i]$; writing $\alpha = (a+ib)^2$ gives the desired result. We have

$$z^2 = N(\alpha) = (x+iy)(x-iy).$$

We claim that $x+iy$ and $x-iy$ are coprime in $\mathbb{Z}[i]$. Assume the contrary that there exists a Gaussian integer π that divides both $x+iy$ and $x-iy$. Then it also divides $2x$ and $2y$. Since x, y are coprime, π has to divide 2. Therefore

$\pi = 1+i$ (up to a unit). But $1+i$ does not divide $x+iy$ since $x \neq y \pmod 2$. Contradiction.

Hence $x+iy$ and $x-iy$ are coprime in $\mathbb{Z}[i]$. As their product is a square, unique factorization in $\mathbb{Z}[i]$ implies that each of them is a square (up to a unit). Using $-1 = i^2$, each of them must be a square or i times a square.

If $x+iy = i(a+ib)^2$, then $x = -2ab$ which contradicts with the assumption that x is odd. Therefore $x+iy$ is a square. □

Next, we solve the sum of two squares problem.

Theorem 4.23. *Let $n = a \cdot b^2$ be an integer with a square-free. Then n can be written as a sum of two squares if and only if no prime $q = 3 \pmod 4$ divides a.*

Proof. The "if" part: For each prime p dividing a, there is a Gaussian prime pi_p such that $p = \pi_p\overline{\pi}_p$. Let $x+iy = b \cdot \prod_{p|a} \pi_p$. Then $x^2+y^2 = n$.

The "only if" part: Suppose $n = x^2+y^2 = (x+iy)(x-iy)$. If a prime $q = 3 \pmod 4$ divides n, as it is a Gaussian prime, it divides $x+iy$ or $x-iy$, which implies that q divides both $x+iy$ and $x-iy$. Thus q^2 divides n. The statement can then be proved by induction on b. □

In the upcoming section, we will use the theory of *modular forms* to count the number

$$r_2(n) = \#\{(x_1, x_2) \in \mathbb{Z}^2 \mid x_1^2 + x_2^2 = n\}.$$

Here is a sketch of the main idea. One can show that

$$E_1^\chi(q) = \frac{1}{4} + \sum_{n=1}^{\infty} \left(\sum_{d|n} \chi(d) \right) q^n \in M_1(\Gamma_1(4)), \quad \text{where } \chi(d) = \begin{cases} 1 & \text{if } n = 1 \pmod 4 \\ -1 & \text{if } n = 3 \pmod 4 \\ 0 & \text{if } n \text{ is even} \end{cases}$$

and the space $M_1(\Gamma_1(4))$ of modular form of weight 1 for the group $\Gamma_1(4) \subseteq \mathrm{SL}(2,\mathbb{Z})$ is one-dimensional, therefore is generated by the function $E_1^\chi(q)$. On the other hand, one can also show that

$$\theta(q)^2 = \sum_{n=0}^{\infty} r_2(n) q^n \in M_1(\Gamma_1(4)).$$

Thus $\theta(q)^2$ is a scalar multiple of $E_1^\chi(q)$. The coefficient of the constant term of $\theta(q)^2$ is $r_2(0) = 1$, while the coefficient of the constant term of $E_1^\chi(q)$ is $1/4$.

Hence one obtains

$$\theta(q)^2 = 4E_1^{\chi}(q).$$

By comparing the coefficients on both sides, we get an explicit formula for $r_2(n)$:

$$r_2(n) = 4\sum_{d|n} \chi(d).$$

Let us give another proof of the formula using the properties of the ring of Gaussian integers. The number $r_2(n)$ can also be interpreted as the number of Gaussian integers with norm n. Thus

$$\sum_{n\geq 1} \frac{r_2(n)}{n^s} = \sum_{0\neq\alpha\in\mathbb{Z}[i]} \frac{1}{N(\alpha)^s}.$$

Denote the set of all Gaussian primes (up to units) by $\mathcal{P}$. Then we have

$$\begin{aligned}\sum_{0\neq\alpha\in\mathbb{Z}[i]} \frac{1}{N(\alpha)^s} &= 4\prod_{\pi\in\mathcal{P}} \frac{1}{1-N(\pi)^{-s}} \\ &= 4\cdot\frac{1}{1-2^{-s}}\cdot\prod_{p=1(\text{mod }4)} \frac{1}{(1-p^{-s})^2}\prod_{q=3(\text{mod }4)} \frac{1}{1-q^{-2s}} \\ &= \zeta(s)\cdot L(s,\chi).\end{aligned}$$

Here $\zeta(s)$ is the Riemann zeta function

$$\zeta(s) = \sum_{n\geq 1}\frac{1}{n^s} = \prod_{p\in\mathbb{Z}\text{ prime}} \frac{1}{1-p^{-s}}$$

and $L(s,\chi)$ is the Dirichlet L-series

$$L(s,\chi) = \sum_{n=1}^{\infty}\frac{\chi(n)}{n^s} = \prod_{p\in\mathbb{Z}\text{ prime}} \frac{1}{1-\chi(p)p^{-s}}.$$

So we have

$$\frac{1}{4}\sum_{n\geq 1}\frac{r_2(n)}{n^s} = \left(\sum_{m\geq 1}\frac{1}{m^s}\right)\left(\sum_{d=1}^{\infty}\frac{\chi(d)}{d^s}\right).$$

Thus

$$\frac{1}{4}r_2(n) = \sum_{md=n}\chi(d) = \sum_{d|n}\chi(d).$$

5. Modular forms

The German mathematician Martin Eichler once stated that there were five fundamental operations of mathematics: addition, subtraction, multiplication, division, and *modular forms*. We will discuss the basic concepts of modular forms and provide some applications, including the sums of squares problem. We would like to understand the counting

$$r_k(n) = \#\{(x_1, \dots, x_k) \in \mathbb{Z}^2 \mid x_1^2 + \cdots + x_k^2 = n\}.$$

Consider the *theta function*

$$\theta(\tau) = \sum_{n=-\infty}^{\infty} e^{2\pi i \tau n^2} = \sum_{n=-\infty}^{\infty} q^{n^2} = 1 + 2\sum_{n=1}^{\infty} q^{n^2} \quad (q = \exp(2\pi i \tau)).$$

It is not hard to see that

$$\theta(\tau)^k = \sum_{n=0}^{\infty} r_k(n) q^n.$$

The problem then reduces to understanding the coefficients of powers of the theta function. Below is the sketch of ideas toward answering this problem.

- The theta function is a *holomorphic function* $\theta\colon \mathbb{H} \to \mathbb{C}$ defines on the upper half plane $\mathbb{H}$.
- The theta function satisfies $\theta(\tau + 1) = \theta(\tau)$ and the transformation formula

$$\theta\left(\frac{\tau}{4\tau + 1}\right) = \sqrt{4\tau + 1}\theta(\tau) \quad \text{for all } \tau \in \mathbb{H}.$$

- Then, for k even, the function $\theta(\tau)^k$ is a *modular form of weight* $k/2$ *for the group* $\Gamma_1(4) \subseteq \mathrm{SL}(2, \mathbb{Z})$, i.e. $\theta^k \in M_{k/2}(\Gamma_1(4))$.
- It turns out that modular forms are rather rare: for instance, for $k = 4$, the space $M_2(\Gamma_1(4))$ is a complex vector space of dimension 2. Moreover, one can write down an explicit basis of $M_2(\Gamma_1(4))$:

$$M_2(\Gamma_1(4)) = \mathrm{Span}\{E_2(\tau) - 2E_2(2\tau), E_2(\tau) - 4E_2(4\tau)\},$$

  where E_2 is certain *Eisenstein series*

$$E_2(\tau) = -\frac{1}{24} + \sum_{n=1}^{\infty} \sigma_1(n) q^n, \quad \sigma_1(n) = \sum_{d|n} d.$$

- By computing the first few coefficients of $\theta(\tau)^4$ and the basis $E_2(\tau) - 2E_2(2\tau), E_2(\tau) - 4E_2(4\tau)$, one can then conclude that

$$\theta(\tau)^4 = 8(E_2(\tau) - 4E_2(4\tau)).$$

- Therefore, we have

$$r_4(n) = 8\left(\sum_{d|n} d - 4\sum_{d|\frac{n}{4}} d\right) = 8\sum_{\substack{d|n \\ 4\nmid d}} d.$$

- A closed formula of $r_k(n)$ can be obtained similarly for even k.

In order to understand this argument, we will discuss the following topics in this section.

- basic complex analysis, holomorphic functions;
- basics of modular forms.

Some references that might be helpful include [3], [7], and [8].

5.1. **More applications of modular forms.** Let us discuss the *j-invariant* first. Classically, the j-invariant was studied as a parameterization of *elliptic curves* over $\mathbb{C}$. Every elliptic curve E over $\mathbb{C}$ is a complex torus, and thus can be identified with a rank 2 lattice. This lattice can be rotated and scaled (which preserve the isomorphism class), so that it is generated by 1 and $\tau \in \mathbb{H}$. This lattice corresponds to the elliptic curve

$$y^2 = 4x^3 - g_2(\tau)x - g_3(\tau),$$

where

$$g_2(\tau) = \frac{4\pi^4}{3}E_4(\tau), \qquad g_3(\tau) = \frac{8\pi^6}{27}E_6(\tau),$$

and

$$E_4(\tau) = 1 + 240\sum_{r\geq 1}\sigma_3(r)q^r, \qquad E_6(\tau) = 1 - 504\sum_{r\geq 1}\sigma_5(r)q^r$$

are *Eisenstein series* (which are *modular forms* of weight 4 and 6, respectively), where $q = e^{2\pi i\tau}$ and $\sigma_k(r) = \sum_{d|r} d^k$. The isomorphic class of elliptic curves is uniquely determined by the j-invariant

$$j(\tau) = 1728\frac{g_2(\tau)^3}{g_2(\tau)^3 - 27g_3(\tau)^2}.$$

It is the *unique* (up to scalar multiplication) holomorphic function on $\mathbb{H}$ that is invariant under the $\mathrm{SL}(2,\mathbb{Z})$-action and has a simple pole at infinity. In fact, any meromorphic modular function (i.e. invariant under $\mathrm{SL}(2,\mathbb{Z})$-action) on $\mathbb{H}$ is a rational function of $j(\tau)$.

Consider

$$e^{\pi\sqrt{163}} = 262537412640768743.99999999999925...$$

which is very close to an integer. This remarkable phenomenon can be easily deduced using the fact that

$$j\left(\frac{1+\sqrt{-163}}{2}\right) \in \mathbb{Z}.$$

together with the q-expansion of the j-function

$$j(\tau) = \frac{1}{q} + 744 + 196884q + 21493760q^2 + O(q^3), \qquad \text{where } q = e^{2\pi i\tau}.$$

Consider primitive positive-definite quadratic forms $Q(x,y) = ax^2 + bxy + cy^2$, where $a, b, c \in \mathbb{Z}$, $\gcd(a,b,c) = 1$, $a > 0$, and $D = b^2 - 4ac < 0$. Two such quadratic forms are *equivalent* if and only if $D = D'$ and

$$j\left(\frac{b+\sqrt{-D}}{2a}\right) = j\left(\frac{b'+\sqrt{-D}}{2a'}\right).$$

For each possible discriminant D there are only finitely many equivalence classes, thus we get a finite set of j-values for each discriminant. The big theorem is that these values are the solutions of a monic algebraic equation with integer coefficients. In particular, when there is only one equivalence class for D, the j-invariant of the corresponding quadratic form must be an integer. The above phenomenon then follows from the fact that all positive-definite integer quadratic forms of discriminant $D = -163$ are equivalent. In fact, 163 is the largest number satisfying this property; other numbers are: $1, 2, 3, 7, 11, 19, 43, 67$; for instance, we also have

$$e^{\pi\sqrt{67}} \approx \mathbb{Z} + 0.0000013; \qquad e^{\pi\sqrt{43}} \approx \mathbb{Z} + 0.00022.$$

These results on the j-function are one of the starting points of the theory of *complex multiplications*.

Another surprising result is a connection between the j-function and the *monster group*.

Theorem 5.1. *Every finite simple group is isomorphic to one of the following groups:*

- *a member of one of three infinite classes of:*
  - *the cyclic groups of prime order,*
  - *the alternating groups A_n for $n \geq 5$,*
  - *the groups of Lie type*
- *one of the 27 sporadic groups.*

Among the 27 sporadic groups, the *monster group* M has the largest order of roughly 8×10^{53}. The minimal dimension of a faithful complex representation of the monster group is 196883, which happens to be very close to one of the coefficients in the q-expansion

$$j(\tau) = \frac{1}{q} + 744 + 196884q + 21493760q^2 + 864299970q^3 + 20245856256q^4 + \cdots .$$

In fact, the dimensions of the irreducible representations of M are: $r_1 = 1$, $r_2 = 196883$, $r_3 = 21296876$, $r_4 = 842609326$, $r_5 = 18538750076$, etc., and the coefficients of the q-expansion of j-function satisfies

$$\begin{aligned} 196884 &= r_1 + r_2 \\ 21493760 &= r_1 + r_2 + r_3 \\ 864299970 &= 2r_1 + 2r_2 + r_3 + r_4 \\ 20245856256 &= 3r_1 + 3r_2 + r_3 + 2r_4 + r_5 \\ &\cdots \end{aligned}$$

Very roughly, this can be explained by the fact that there exists a *vertex operator algebra* which admits an infinite-dimensional graded representation of the monster group, whose graded dimensions are the coefficients of the j-function. The precise content of this statement and their detailed properties (Conway–Norton conjecture) are proved by Borcherds, who won the Fields Medal in 1998 in part for his solution of the conjecture.

Let us consider a more elementary application of modular forms. Consider the functions

$$\sigma_3(r) = \sum_{d|r} d^3 \quad \text{and} \quad \sigma_7(r) = \sum_{d|r} d^7.$$

They satisfy a relation

$$\sigma_7(r) = \sigma_3(r) + 120 \sum_{p+q=r} \sigma_3(p)\sigma_3(q).$$

This is not an easy statement to prove. Using the fact that

$$E_4(\tau) = 1 + 240 \sum_{r\geq 1} \sigma_3(r) q^r \quad \text{and} \quad E_8(\tau) = 1 + 480 \sum_{r\geq 1} \sigma_7(r) q^r$$

are modular forms of weight 4 and 8, respectively; together with the fact the space of modular forms of weight 8 is one-dimensional, one deduces $E_4(\tau)^2 = E_8(\tau)$. The above relation then follows from comparing the coefficients of both sides of the equation.

5.2. **Crash course on complex analysis.** Let $U \subseteq \mathbb{C}$ be an open subset of the complex plane. A function $f\colon U \to \mathbb{C}$ is called *holomorphic* if for every $z_0 \in U$, the limit

$$\lim_{z\to z_0} \frac{f(z) - f(z_0)}{z - z_0} \text{ exists.}$$

In other words, it is holomorphic if the derivative in the "complex sense" exists. If the limit exists, it will be denoted by $f'(z_0) \in \mathbb{C}$. This is exactly the complex analogue of the *differentiable* functions over $\mathbb{R}$. However, holomorphic functions possess many nicer properties than differentiable functions.

Example 5.2. Holomorphic functions satisfy the "local determine global" principle. Namely, suppose there are two holomorphic functions f, g on a (connected) open set $U \subseteq \mathbb{C}$ such that their values agree on an open subset $V \subseteq U$, i.e. $f(z) = g(z)$ for all $z \in V$. Then, no matter how small the open subset V is, we would have $f(z) = g(z)$ for all $z \in U$.

This is not true for smooth functions over $\mathbb{R}$. For instance, the smooth function

$$f(x) = \begin{cases} e^{-1/x^2} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

is identical with the zero function on $\mathbb{R}_{<0}$, but they are obvious not identical on the whole real line.

Example 5.3. Another important result is that if $f\colon U \to \mathbb{C}$ is holomorphic, then its derivative $f'\colon U \to \mathbb{C}$ is also holomorphic. This implies that any

holomorphic is infinitely differentiable, i.e. $f, f', f'', f''', \dots$ exist. Moreover, for any $z_0 \in U$ the power series

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(z_0)}{n!}(z - z_0)^n$$

converges in a neighborhood of z_0, and the limit coincides with $f(z)$. These are again not true for differentiable functions over $\mathbb{R}$.

These results, together with other basic theorems in complex analysis, including Liouville's theorem, Morera's theorem, residue formula, argument principle, etc., essentially all are corollaries of a single theorem, the *Cauchy integral theorem*. To state the theorem, we need to define the notion of *path integrals*.

Definition 5.4. A *parametrized smooth curve* in $U \subseteq \mathbb{C}$ is a map

$$\gamma\colon [a, b] \to U; \qquad \gamma(t) = x(t) + iy(t)$$

such that

- $x(t), y(t)$ are differentiable, and $x'(t), y'(t)$ are continuous,
- $\gamma'(t) = (x'(t), y'(t)) \neq (0, 0)$ for all $t \in (a, b)$.

Example 5.5. $\gamma\colon [0, \pi] \to \mathbb{C}$ where $\gamma(t) = e^{it} = \cos(t) + i\sin(t)$ parametrizes the upper half of the unit circle (going counterclockwise). Note that there are infinitely many ways to represent a curve. For instance, $\gamma'\colon [0, 2\pi] \to \mathbb{C}$ where $\gamma'(s) = e^{is/2}$ also parametrizes the upper half of the unit circle with the same orientation.

Definition 5.6. Two parametrized smooth curves $\gamma\colon [a, b] \to \mathbb{C}$ and $\gamma'\colon [c, d] \to \mathbb{C}$ are said to be *equivalent* if there exists a smooth bijective map $\varphi\colon [c, d] \to [a, b]$ so that $\gamma(\varphi(s)) = \gamma'(s)$ and $\varphi'(s) > 0$ for all $s \in [c, d]$.

Note that the condition $\varphi'(s) > 0$ guarantees that the two curves have the same orientations.

Definition 5.7. A *piecewise parametrized smooth curve* in $U \subseteq \mathbb{C}$ is a continuous map

$$\gamma\colon [a, b] \to U$$

such that there exists $a < p_1 < \cdots < p_n < b$ so that

$$\gamma|_{[a,p_1]}, \dots, \gamma_{[p_n,b]}$$

are parametrized smooth curves.

Definition 5.8. Let $\gamma\colon [a,b] \to \mathbb{C}$ be a piecewise parametrized smooth curve on an open set $U \subseteq \mathbb{C}$, and let $f\colon U \to \mathbb{C}$ be a continuous function. The *integral of f along γ* is defined to be

$$\int_\gamma f(z)\,\mathrm{d}z := \int_a^b f(\gamma(t)) \cdot \gamma'(t)\,\mathrm{d}t.$$

Exercise. Show that if γ and γ' are equivalent, then

$$\int_\gamma f(z)\,\mathrm{d}z = \int_{\gamma'} f(z)\,\mathrm{d}z \quad \text{for any } f.$$

In other words, the integral depends only on the underlying curve (and its orientation).

Exercise. Show that if γ and γ' parametrizes the same curve but with opposite orientations, then

$$\int_\gamma f(z)\,\mathrm{d}z = -\int_{\gamma'} f(z)\,\mathrm{d}z \quad \text{for any } f.$$

The following is perhaps the most important example of path integrals.

Example 5.9. Consider the unit circle parametrizes counterclockwisely $\gamma\colon [0,2\pi] \to \mathbb{C}$ where $\gamma(t) = e^{it}$. The function $f(z) = \frac{1}{z}$ is continuous (in fact, holomorphic) on $\mathbb{C}\backslash\{0\}$, so it makes sense to compute the path integral of f along the unit circle.

$$\int_\gamma f(z)\,\mathrm{d}z = \int_0^{2\pi} \frac{1}{e^{it}} \cdot ie^{it}\,\mathrm{d}t = 2\pi i.$$

We are now ready to state the Cauchy integral theorem.

Theorem 5.10 (Cauchy integral theorem)**.** *Let γ be a simple closed curve in $\mathbb{C}$. Suppose f is holomorphic on an open set containing γ* and its interior*, then*

$$\int_\gamma f(z)\,\mathrm{d}z = 0.$$

(need zeros, poles, residue formula)

(what I need for elliptic functions and modular forms discussion: argument principle; Laurent expansion)

Bibliography

[1] M. Artin. *Algebra* (Second Edition). Pearson Education, 2011.

[2] M. Atiyah and I. G. Macdonald. *Introduction to commutative algebra*. Westview Press, Boulder, CO, 2016.

[3] F. Diamond and J. Shurman. *A first course in modular forms*. Graduate Texts in Mathematics, 228. Springer-Verlag, New York, 2005.

[4] M. Einsiedler and T. Ward. *Ergodic theory with a view towards number theory*. Graduate Texts in Mathematics, 259. Springer-Verlag London, Ltd., London, 2011.

[5] A. Hatcher. *Algebraic topology*. Cambridge University Press, Cambridge, 2002.

[6] J. Matousek. *Using the Borsuk-–Ulam Theorem*. Lectures on topological methods in combinatorics and geometry. Universitext. Springer-Verlag, Berlin, 2003.

[7] J.-P. Serre. *A course in arithmetic*. Graduate Texts in Mathematics, No. 7. Springer-Verlag, New York-Heidelberg, 1973.

[8] E. M. Stein and R. Shakarchi. *Complex analysis*. Princeton Lectures in Analysis, 2. Princeton University Press, Princeton, NJ, 2003.

[9] P. Walter. *An introduction to ergodic theory*. Graduate Texts in Mathematics, 79. Springer-Verlag, New York-Berlin, 1982.