A Brief Introduction to the p-Adics

Maria Qi
University of California Santa Barbara

Introduction

Suppose you wanted to prove the equation $3x^2+2y^2+z^2=0$ has no nontrivial solutions in the rationals. You could say that because every square of a real number is nonnegative, the only solution is (0, 0, 0). But what about $3x^2+2y^2-z^2=0$? For this you could instead show that the equation has no nontrivial solutions in $\mathbb{Q}_3$, which implies no nontrivial solutions in the rationals. What does this mean??

Given a prime number p, the p-adic numbers form a unique numbering system that represents numbers as a possibly infinite base-p expansion of powers of p. The p-adics in general are completions of the rational numbers with an unusual distance metric, which gives rise to unique geometric properties. Together with tools like Hensel's lemma, the local-global principle, and quadratic reciprocity, the p-adics are incredibly useful in solving many problems in number theory, particularly in finding integer and rational solutions to Diophantine equations.

Defining the *p*-Adics

The p-Adic Valuation on $\mathbb{Z}$ and $\mathbb{Q}$

Given a prime p, we can write any nonzero integer $n \in \mathbb{Z}$ as $n = p^k n'$, where $p \nmid n'$. Then we define the p-adic valuation as the function

$$v_p : \mathbb{Z} - \{0\} \to \mathbb{R}, \qquad v_p(n) = k.$$

We extend v_p to the field of rationals by defining

$$v_p(\frac{a}{b}) = v_p(a) - v_p(b), \quad \frac{a}{b} \in \mathbb{Q}, \qquad \text{with } v_p(0) = \infty.$$

Now we define the p-adic absolute value function $|\quad|_p : \mathbb{Q} \to \mathbb{R}$ as

$$|x|_p = p^{-v_p(x)}, \qquad |0|_p = 0.$$

In other words, the *p*-adic absolute value is the largest (by magnitude) integer power of p which divides a number x.

The Distance Function using $|\quad|_p$

Define a distance function $d(a,b) = |a-b|_p$. We can confirm that all four conditions of a distance function hold when using the p-adic absolute value:

1. $d(a,b) \geq 0$
2. $d(a,b) = d(b,a)$
3. $d(a,a) = 0$
4. $d(a,c) \leq d(a,b) + d(b,c)$

More interesting is the fact that p-adic absolute value satisfies the ultrametric inequality (this will come in handy later!):

5. $d(a,c) \leq max\{d(a,b), d(b,c)\}$

The Field of p-Adic Numbers

Let $\mathbb{Q}_p$ be the completion of $\mathbb{Q}$ using the p-adic metric defined above. Next, let $\mathbb{Z}_p = \{x \in \mathbb{Q}_p : |x|_p \leq 1\}$, such that $\mathbb{Z}_p$ is the completion of $\mathbb{Z}$ using the p-adic metric. Then we have $\mathbb{Q} \subset \mathbb{Q}_p$ and $\mathbb{Z} \subset \mathbb{Z}_p$.

The p-Adic Units

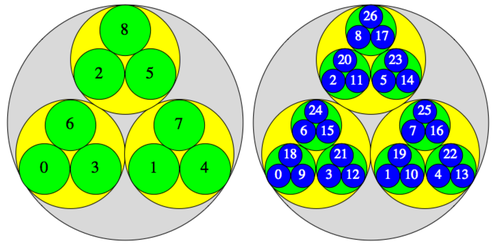
For a given p, the p-adic units are the invertible elements of $\mathbb{Z}_p$, defined as

$$\mathbb{Z}_p^\times = \{x \in \mathbb{Q}_p : |x|_p = 1\}.$$

Thus for any prime p, its p-adic units are the set of rationals $\{\frac{a}{b} \in \mathbb{Q} : p \nmid ab\}$.

Geometry of $\mathbb{Q}_p$

A fun result of the distance function on the p-adics is that "closeness" between two numbers is defined by similar divisibility by p^k. Larger numbers may be "closer" than consecutive integers, as shown below:



Visualization of the geometry of $\mathbb{Q}_3$ up to 3^1 and 3^2. Images by Heiko Knopse.

Interesting Results in $\mathbb{Q}_p$

Every Triangle in $\mathbb{Q}_p$ is Isosceles

Proof. First we will prove that $v_p(a) \neq v_p(b) \to v_p(a-b) = min\{v_p(a), v_p(b)\}$. Since $|a-b|_p = |b-a|_p$ implies $v_p(a-b) = v_p(b-a)$, we will only consider the case where $v_p(a) > v_p(b)$.

$$\begin{aligned} v_p(a-b) &= v_p(p^m a' - p^n b'), && p \nmid a', p \nmid b' \\ &= v_p(p^n(p^{m-n}a' - b')), && p \nmid b' \\ &= n = v_p(b). \end{aligned}$$

Similarly, $v_p(a-b) = v_p(a)$ when $v_p(a) < v_p(b)$. Thus $v_p(a-b) = min\{v_p(a), v_p(b)\}$.

Now we will prove that every triangle in $\mathbb{Q}_p$ is isosceles. Let $\triangle abc$ be a triangle in $\mathbb{Q}_p$. Then it has sides of length $d(a-c) = |a-c|_p, d(a-b) = |a-b|_p, d(b-c) = |b-c|_p$. Assume $|a-b|_p \neq |b-c|_p$, that is, $v_p(a-b) \neq v_p(b-c)$. Then, by the result above, we have

$$v_p(a-c) = v_p(a-b+b-c) = min\{v_p(a-b), v_p(b-c)\}.$$

This implies at least two sides of $\triangle abc$ must be of equal length. Thus every triangle in $\mathbb{Q}_p$ is isosceles. □

Every Point in an Open Ball is a Center of that Ball

Proof. For this proof we will rely on the ultrametric inequality in $\mathbb{Q}_p$. Consider an open ball $B(a,r) = \{b \in \mathbb{Q}_p : |b-a|_p < r\}$. Then for $b, x \in B(a,r)$, we have

$$|x-b|_p \leq max\{|x-a|_p, |b-a|_p\} < r,$$

which implies $x \in B(b,r)$. Thus $B(a,r) \subset B(b,r)$. Conversely, since $b \in B(a,r)$ implies $a \in B(b,r)$, we have

$$x \in B(b,r) \quad \Rightarrow \quad |x-a|_p \leq max\{|x-b|_p, |a-b|_p\} < r,$$

which implies $B(b,r) \subset B(a,r)$. Thus for $b \in B(a,r)$, we have $B(b,r) = B(a,r)$. □

Any Two Open Balls are Either Disjoint or Contained in One Another

Proof. This follows from the above result. Suppose $r_1 \leq r_2$. If $B(a,r_1) \cap B(b,r_2) \neq \varnothing$, then there exists a point c such that $c \in B(a,r_1) \cap B(b,r_2)$. Then we have $B(c,r_1) = B(a,r_1)$ and $B(c,r_2) = B(b,r_2)$, which implies $B(a,r_1) \subset B(b,r_2)$. □

(A similar method can be used to prove this and the above result with closed balls. The proof is left as an exercise to the reader.)

Legendre's Three Squares Theorem

We will now use these results to outline a basic proof of a well-known Diophantine problem, Legendre's three-square theorem:

n can be written as the sum of three squares *if and only if* $n \neq 4^a(8b+7)$.

Proof. The sufficient direction can be easily proven with modular arithmetic. First, consider $n = 8b+7$. Since any integer squared has the form 0, 1, or 4 (mod 8), no combination of three integer squares can add to 7 (mod 8). Then, for $n = 4^a(8b+7)$ with $a \geq 1$, taking the equation $n = x^2+y^2+z^2$ *modulo* 4 tells us x^2, y^2, z^2 are all even. Setting $x = 2x_1, y = 2y_1, z = 2z_1$, we get $4^{a-1}(8b+7) = x_1^2+y_1^2+z_1^2$. Repeat this step until the exponent is reduced to 0, and we once again have $8b+7 = x_a^2+y_a^2+z_a^2$, which as we know is impossible. Thus no integer of the form $4^a(8b+7)$ can be written as the sum of three squares.

The necessary direction is more difficult and requires using advanced strategies in p-adic analysis. To start, we will prove that every square in $\mathbb{Q}_2$ is of the form $4^a(8b+1)$. We first note that every $d \in \mathbb{Q}_2$ is of the form $2^a(d')$, where $a \geq 0$ and $d' \in \mathbb{Z}_2^\times$ is a 2-adic unit (in other words, $d' \equiv 1 \bmod 2$). As this implies $d'^2 \equiv 1 \bmod 8$, it follows that every square in $\mathbb{Q}_2$ is of the form $4^a(d'^2)$ which is congruent to $4^a(1 \bmod 8)$.

Now suppose n is not of the form $4^a(8b+7)$. Then, $-n$ is not a square in $\mathbb{Q}_2$. One fundamental property commonly used to solve Diophantine equations in the p-adics is the Hasse-Minkowski theorem, which is one of many local-global principles: a quadratic form (any polynomial where every term is of degree 2) has solutions in $\mathbb{Q}$ *if and only if* it has solutions in all $\mathbb{Q}_p$, including $\mathbb{Q}_\infty = \mathbb{R}$. Using results by Kurt Hensel and David Hilbert, we find that for any $n > 0$, $n = x^2+y^2+z^2$ has solutions in $\mathbb{Q}_p$ for all odd primes $p \geq 3$ as well as the real numbers, but this does not hold for all $n \in \mathbb{Q}_2$. Particularly, the equation holds in $\mathbb{Q}_2$ *if and only if* $-n$ is not a square in $\mathbb{Q}_2$. From our earlier results we see that when $-n$ is a square, we have $n \equiv -4^a(1 \bmod 8) \equiv 4^a(-1 \bmod 8) \equiv 4^a(7 \bmod 8) = 4^a(8b+7)$. Thus:

$$n \neq x^2+y^2+z^2 \Rightarrow -n \text{ is a square in } \mathbb{Q}_2 \Rightarrow n = 4^a(8b+7).$$

Now we have proven the necessary condition. Thus any positive integer n can be expressed as the sum of three squares *if and only if* n is not of the form $4^a(8b+7)$.

□

Acknowledgements

I would like to thank my mentor Marcos Reyes for his guidance and the UCSB Directed Reading Program for the opportunity to work on this project.

References

[1] David M. Burton. *Elementary Number Theory*. Allyn & Bacon, 1980.

[2] Catherine Crompton. "Some Geometry of the p-adic Rationals". In: *Rose-Hulman Undergraduate Mathematics Journal* (2007).

[3] Fernando Q. Gouvêa. *p-adic Numbers: An Introduction*. Springer Cham, 2020.

[4] Xingyu Wang. "p-Adic Numbers, Hasse-Minkowski Theorem, and its Applications". In: *UChicago REU Papers* (2019).

Elementary, My Dear Topos: Logic Beyond Sets

Alex Cao, Yifan Dai, Elian Gal-on, Dmitri Anh-Minh Tran

University of California, Santa Barbara

Subobjects

A *subobject* of A is an (equivalence class of) monomorphisms into A. Two subobjects $X \xrightarrow{m} A$ and $Y \xrightarrow{m'} A$ are *equivalent* if there is an isomorphism $X \xrightarrow{\phi} Y$ such that the below commutes.

$$\begin{array}{ccc} X & \xrightarrow{\phi} & Y \\ & {\scriptstyle m}\searrow & \downarrow{\scriptstyle m'} \\ & & A \end{array}$$

We will conflate a subobject, its source, and its equivalence class. In **Set**, a subobject of A can be identified with its image, i.e. a subset of A.
The *characteristic function* of a subset $X \subseteq A$ is the function $\chi_A : A \to \{0,1\}$ defined by $\chi_X(a) = 0$ if and only if $a \in X$. The association $X \leftrightarrow \chi_X$ is a bijection between $\mathrm{Sub}(A)$ and $\mathrm{Hom}(A, \{0,1\})$.
In any category, a morphism $1 \xrightarrow{\mathrm{true}} \Omega$ is a *subobject classifier* if, for any subobject $X \xrightarrow{m} A$, there is exactly one $A \xrightarrow{f} \Omega$ such that the following is a pullback

$$\begin{array}{ccc} X & \longrightarrow & 1 \\ {\scriptstyle m}\downarrow & & \downarrow{\scriptstyle \mathrm{true}} \\ A & \xrightarrow{f} & \Omega \end{array}$$

In **Set**, if we set $\Omega = \{0,1\}$ and let $\mathrm{true} : \{0\} \to \{0,1\}$ send 0 to 0, then Ω is a subobject classifier and f is the characteristic function of X.
Thus, true classifies subsets in two senses: 1. $\mathrm{Sub}(A) \cong \mathrm{Hom}(A, \Omega)$, naturally in A; 2. Ω is a subobject classifier in **Set**. These conditions turn out to be equivalent.

Elementary Topoi

A topos $\mathcal{E}$ is a category that has all finite limits and is equipped with an object Ω and a functor $P : \mathcal{E}^{\mathrm{op}} \to \mathcal{E}$ (the "powerset functor") , such that

$$\mathrm{Sub}_{\mathcal{E}} A \cong \mathrm{Hom}_{\mathcal{E}}(A, \Omega) \tag{1}$$

$$\mathrm{Hom}_{\mathcal{E}}(B \times A, \Omega) \cong \mathrm{Hom}_{\mathcal{E}}(A, PB) \tag{2}$$

naturally in A. Recall Equation 1 asserts that $\mathcal{E}$ has a subobject classifier. An example is the category of sets, where $\Omega = 2$ (corresponding to true/false), and P is simply the powerset function.
A topos gives a way to define logical predicates on A as a subobject $A' \to A$, its truth value being a characteristic morphism $\chi : A \to \Omega$. For instance in $A = \mathbb{N} \in$ **Set**, the statement "n is even" corresponds to a morphism that sends all even numbers to 0.
In a more general topos, Ω can have more than two elements which gives rise to more truth values than True/False.
Other logical connectives can be modeled as well: the conjunction $P \wedge Q$ is represented by the pull back of χ_P and χ_Q, and the implication $\implies$ is constructed with Ω^{Ω}.

Exponentials

Given A, B in a category, B^A is the *exponential object* of B and A if $\mathrm{Hom}(X \times A, B) \cong \mathrm{Hom}(X, B^A)$ natuarlly in X. We will show that exponential of any object exists in a topos. The construction is similar to the construction of B^A in sets: it is a subset of $A \times B$. We consider $P(A \times B)$. Using **Set** as an example, a subset S of $A \times B$ is a function only when for all $a \in A$, $\{b : (a,b) \in S\}$ is a singleton for all a. The predicate of "being a singleton" can be formulated as a morphism to Ω. Thus we let $\phi : P(A \times B) \to \Omega$ be the morphism that asserts that a subobject represents a function. Then B^A is the pull back of ϕ along $\mathrm{true} : 1 \to \Omega$.

Heyting Algebras

A lattice is a poset with all binary products (greatest lower bounds) and binary coproducts (least upper bounds). A Heyting algebra is a Cartesian closed poset with all finite products and coproducts, i.e., a lattice with 0,1, and all exponentials $y^x = (x \Rightarrow y)$ for elements x, y of the lattice. These exponentials satisfy the identities $(x \Rightarrow x) = 1$, $x \wedge (x \Rightarrow y) = x \wedge y$, $y \wedge (x \Rightarrow y) = y$, and $x \Rightarrow (y \wedge z) = (x \Rightarrow y) \wedge (x \Rightarrow z)$; negation may be defined in a Heyting algebra as $x \Rightarrow 0$. For any small category **C** and any contravariant functor P into **C**, the poset $\mathrm{Sub}_{\widehat{\mathbf{C}}}(P)$ of subfunctors of P is a Heyting algebra, where $\widehat{\mathbf{C}}$ denotes $\mathbf{Sets}^{\mathbf{C}^{\mathbf{op}}}$. In any category **C** with finite limits, an internal Heyting algebra is an object L equipped with morphisms $\wedge, \vee, \Rightarrow : L \times L \to L$ satisfying the identities which define a Heyting algebra, i.e., an object whose subobjects form a Heyting algebra. These equations may be diagrammatized in order to define the operations as morphisms of **C**. For example, letting $\Rightarrow_{L^3}$ denote $(\Rightarrow \times \Rightarrow) \circ (1_L \times \tau \times 1_L) \circ (\pi_1 \times 1_{L \times L \times L})$,

$$\begin{array}{ccc} L \times L \times L & \xrightarrow{1_L \times \wedge} & L \times L \\ {\scriptstyle \Rightarrow_{L^3}}\downarrow & & \downarrow{\scriptstyle \Rightarrow} \\ L \times L & \xrightarrow[\wedge]{} & L \end{array}$$

commutes when $x \Rightarrow (y \wedge z) = (x \Rightarrow y) \wedge (x \Rightarrow z)$.

Direct Image

We can extend the notion of the *direct image* of a subset under a morphism of sets to general topoi. In particular, we consider an arbitrary monomorphism $k : B' \to B$ in a topos $\mathcal{E}$. We can define a morphism $\exists_k : PB' \to PB$ between the power objects of B' and B. The construction of $\exists_k$ is contained in the diagram below:

$$\begin{array}{ccccc} U & \longrightarrow & 1 & = & 1 \\ {\scriptstyle u_{B'}}\downarrow & & \downarrow{\scriptstyle \mathrm{true}} & & \\ B' \times PB' & \xrightarrow{\in_{B'}} & \Omega & & \downarrow{\scriptstyle \mathrm{true}} \\ {\scriptstyle k \times 1}\downarrow & & & & \\ B \times PB' & & \xrightarrow{e_k} & & \Omega \end{array}$$

$$PB' \xrightarrow{\exists_k} PB$$

The *Beck-Chevalley Condition* for $\exists$: if m is the pullback of a monomorphism k along an arbitrary arrow g in a topos, as in the left-hand square, then the right-hand square will commute:

$$\begin{array}{ccc} C' & \xrightarrow{g'} & B' \\ {\scriptstyle m}\downarrow & & \downarrow{\scriptstyle k} \\ C & \xrightarrow{g} & B \end{array} \qquad \begin{array}{ccc} PB' & \xrightarrow{Pg'} & PC' \\ {\scriptstyle \exists_k}\downarrow & & \downarrow{\scriptstyle \exists_m} \\ PB & \xrightarrow{Pg} & PC \end{array}$$

For all monomorphisms k, the composite

$$PB' \xrightarrow{\exists_k} PB \xrightarrow{Pk} PB'$$

is the identity. This is because the commutative diagram

$$\begin{array}{ccc} B' & \xrightarrow{1} & B' \\ {\scriptstyle 1}\downarrow & & \downarrow{\scriptstyle k} \\ B' & \xrightarrow{k} & B \end{array}$$

is a pullback, $\exists_1 = 1$, and $P1 = 1$. Applying the Beck-Chevalley condition completes the proof that this composition of morphisms is the identity.
Furthermore, there exists another "direct image" $k_! : \mathrm{Sub}_{\mathcal{E}} B' \to \mathrm{Sub}_{\mathcal{E}} B$ which is to $\mathrm{Sub}_{\mathcal{E}} B'$ what $\exists_k$ is to PB'. For any monomorphism $k' : B'' \to B'$ and $k : B' \to B$, $k \circ k'$, treated as a subobject of B, is the "direct image" under k of k' treated as a subobject.

Factoring morphisms

We can factor any function as a surjection followed by an injection

$$\begin{array}{ccccc} & & f & & \\ & \overset{\frown}{} & & \overset{\frown}{} & \\ A & \xrightarrow{e} & \mathrm{im} f & \xrightarrow{m} & B \end{array}$$

where m is the inclusion map. In any category with finite limits and colimits, we can factorize $f = me$, where m is a monomorphism and e is an epimorphism. (Many authors define topos to have finite colimits. This is not strictly necessary; see [1], p. 176.) We will give a categorical construction in **Set** and merely assert that it works in general. (For details, see [1], p. 184.)
Let C be the pushout of f and f. Explicitly, C contains two copies of B, where the elements in $\mathrm{im} f$ are identified with one another. x sends B to the first copy, and y sends B to the second. Let $M \xrightarrow{m} B$ be the equalizer of x and y. Explicitly, $M \cong \mathrm{im} f$. From the pushout diagram defining C, we see that f also equalizes x and y. By the universal property of equalizer, there is a unique $A \xrightarrow{e} M$ such that the following commutes

$$\begin{array}{ccccc} M & \xrightarrow{m} & B & \underset{y}{\overset{x}{\rightrightarrows}} & C \\ {\scriptstyle e}\uparrow & \nearrow_{f} & & & \\ A & & & & \end{array}$$

$f = me$ is the desired factorization.

Heyting Algebra Structure of Topoi

For any object A of a topos $\mathcal{E}$, the poset $\mathrm{Sub}_{\mathcal{E}}(A)$ is a Heyting algebra, and for any morphism $k : A \to B$, the induced map $k^{-1} : \mathrm{Sub}_{\mathcal{E}} B \to \mathrm{Sub}_{\mathcal{E}} A$, defined pointwise by the pullback along k of any subobject inclusion map of A, is a homomorphism of Heyting algebras.
To show that $\mathrm{Sub}_{\mathcal{E}}(A)$ is a poset, it suffices to consider that for any monomorphism $U, V :\rightarrowtail 1$, U^V is also monomorphism, so that the lattice $\mathrm{Sub}_{\mathcal{E}}(1)$ has exponentials, and so is a Heyting algebra. The result follows from the fact that the slice category of a topos over any of its objects A is also a topos, and $\mathrm{Sub}_{\mathcal{E}}(A) \cong \mathrm{Sub}_{\mathcal{E}/A}(1)$ in any topos. Furthermore, the pullback functor along k preserves the subobject classifier and exponentials, and commutes with the morphisms ι_A, ι_B identifying subobjects of A with their inclusion maps to A, so the following diagram demonstrates that k^{-1} preserves Heyting algebra structure:

$$\begin{array}{ccc} \mathrm{Sub}_{\mathcal{E}}(B) & \xrightarrow{k^{-1}} & \mathrm{Sub}_{\mathcal{E}}(A) \\ {\scriptstyle \iota_B}\downarrow & & \downarrow{\scriptstyle \iota_B} \\ \mathcal{E}/B & \xrightarrow{k^*} & \mathcal{E}/A \end{array}$$

Likewise, for any object A of a topos $\mathcal{E}$, the exponential object PA is an internal Heyting algebra, with the property that for any morphism $k : A \to B$, the induced map $Pk : PA \to PB$ is a homomorphism of Heyting algebras.

Acknowledgements

Thank you to Choomno Moos for mentoring us.

References

[1] Saunders Mac Lane and Ieke Moerdijk. *Sheaves in Geometry and Logic*. Springer-Verlag, 1994.

Data–Driven Patterns in Elliptic Curves

Dulce Rodriguez and Daniel Ramos
University of California Santa Barbara

What Are Elliptic Curves?

An elliptic curve over $\mathbb{Q}$ is a smooth cubic projective curve E defined over $\mathbb{Q}$, with at least one rational point $\mathcal{O} \in E(\mathbb{Q})$ that we call the origin. For simplicity, we concentrate on trying to find all rational points on a curve

$$E(\mathbb{Q}) : y^2 = x^3 + ax + b, \quad a, b \in \mathbb{Q}$$

with discriminant $\Delta = -16(4a^3 + 27b^2) \neq 0$ to ensure non-singularity. The set of rational points $E(\mathbb{Q})$, defined as

$$\{(x, y) \in E \mid x, y \in \mathbb{Q}\} \cup \{\mathcal{O}\}$$

where $\mathcal{O}$ = [0,1,0] is the point at infinity. This set forms a finite generated abelian group:

$$E(\mathbb{Q}) \cong \mathbb{Z}^r \oplus E(\mathbb{Q})_{\text{tors}}$$

where r is the rank (infinite-order points) and $E(\mathbb{Q})_{\text{tors}}$ is the finite torsion subgroup.
Elliptic curves play a central role in number theory, cryptography, and algebraic geometry.

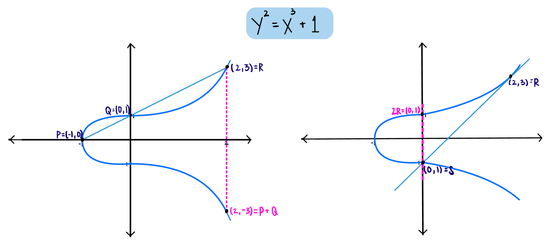
How Does Point Addition Work?

Point addition and doubling gives elliptic curves their group structure:
Point Addition: Given two distinct points P and Q on an elliptic curve, their sum $P + Q$ is defined as follows:

- Draw the straight line that intersects P and Q.
- This line will generally intersect the curve at a third point, say R.
- Reflect R across the x-axis to get $P + Q$.

Point Doubling ($2P$): If $P = Q$, we take the tangent line at P and repeat the same process:

- Compute where the tangent at P intersects the curve again.
- Reflect that intersection point across the x-axis to obtain $2P$.



Why Rank and Torsion Matter

When the **Rank** is zero, there are only a few rational points. On the other hand, if the rank is at least one, there are infinitely many, and these can be added together to create new points. The rank is at the center of the famous **Birch and Swinnerton-Dyer Conjecture**, which predicts how the rank connects to deeper properties of the curve.
Torsion subgroup is the part of the curve that contains points that eventually "loop back" to the identity when added to themselves a finite number of times (finite-order points). Over the rational numbers, the torsion subgroup is always one of just 15 possible types, thanks to a result called **Mazur's Theorem**. Both rank and torsion give us insight into the curve's structure. We explored patterns in torsion and discriminant values and their relation to rank.

Our Data Science Approach

Goal: Explore statistical and predictive relationships among rank, torsion, and discriminant.
Data Source: Sample of 1 million elliptic curves from the 3.8 million available in the LMFDB (L-Functions and Modular Forms Database), queried using the lmfdb-lite Python library.
Key Fields:

Rank (r), Discriminant (Δ), Torsion structure, Conductor (N), Maximal Rank

Workflow:

1. Connect to LMFDB mirror via 'psycopg2'
2. Extract and clean data using 'pandas'
3. Apply log-scaling to $|\Delta|$, one-hot encode torsion types
4. Build regression and classification models (Random Forest, Logistic Regression)

Statistical Modeling

Figure 1

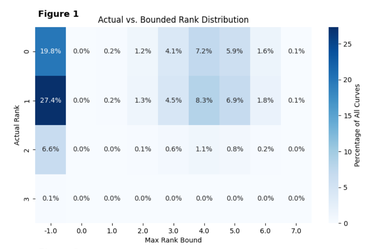


Figure 2

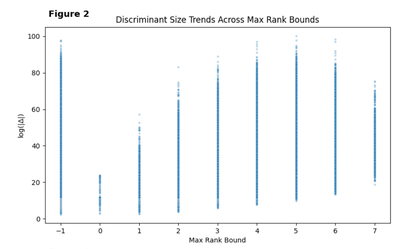


Figure 3

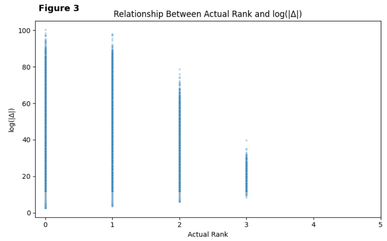
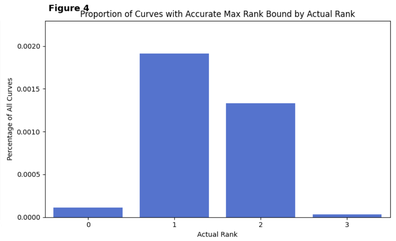


Figure 4



Graph Note: The "1" column appears only in the top two graphs and represents curves that don't satisfy the max rank condition. It's included for reference.

Model classification performance on elliptic curve rank prediction

Logistic Regression

Class	Precision	Recall	Support
0 (rank 0)	0.45	0.57	119,633
1 (rank > 0)	0.65	0.53	180,367
Accuracy	0.56		300,000
Weighted avg	0.57	0.55	300,000

Decision Tree

Class	Precision	Recall	Support
0 (rank 0)	0.44	0.64	119,633
1 (rank > 0)	0.66	0.46	180,367
Accuracy	0.53		300,000
Weighted avg	0.57	0.53	300,000

Max Rank

The **rank** is one of the most important properties describing the structure of an elliptic curve. There isn't a method or formula that computes the rank for elliptic curves due to its complexity. However, we can compute the **max rank** under certain conditions to create an interval given by:

$$0 \leq \operatorname{rank}(E) \leq \max \operatorname{rank}(E).$$

This allows us to narrow down our understanding of the elliptic curves rank. The bound is found by using the following statement. Let $E/\mathbb{Q}$ be any elliptic curve with a non-trivial point of 2-torsion, and let a (resp. m) be the number of primes of additive (resp. multiplicative) bad reduction of $E/\mathbb{Q}$. Then:

$$\operatorname{rank}_{\mathbb{Z}}(E(\mathbb{Q})) \leq m + 2a - 1.$$

The max rank used in the graphs was calculated using an algorithm we developed that involved the curves' numerical properties found in the database.

Findings and Discussion

Graph Discussion

Figure 1: We can see the majority of the curves in our sample fall in the "-1" column, further demonstrating how difficult obtaining knowledge about the rank can be. Interestingly, rank 0 and 1 curves had the highest frequency of bounded ranks in our sample, specifically a max range of 3, 4 or 5.
Figure 2: We can see the size of the "-1" column's Δ stretches the entire y-axis, exhibiting no apparent trend. However, the interval where the discriminant lies increases in size and value-wise as the max rank bound increases.
Figure 3: Here we see the graph exhibits a downwards trend. Large Δ tend to trigger more primes of bad reduction, the variables used in computing the max rank bound. This appears to overestimate their rank significantly.
Figure 4: This table illustrates the tremendous unlikelihood of an elliptic curves rank being equal to their bound. This refers to the **percentage** of curves in the sample whose actual rank matches their bound.

Table Insights

We applied several models that classified an elliptic curve as having rank 0 or greater than 0 based on the discriminant, max rank, and the torsion order. We chose to display the 2 most efficient models, a logistic regression and a binary decision tree. Advanced methods such as Support Vector Machines were not very effective as they require heavy computing power and smaller sample sizes were ineffective. The imbalance of curves with rank 0/rank above 0, limited us to methods that were efficient with large samples and accounted for imbalanced data. To facilitate **table interpretation**, we've included some definitions. **Precision:** out of the predictive positives, how many were actually correct. **Recall:** out of the actual positives, how many did the model guess correctly.
Limitations: Rank is difficult to predict precisely; the BSD conjecture remains unsolved.

Acknowledgments

Special thanks to our DRP mentor, Marcos Reyes, the UCSB Direct Reading Program, and the developers of LMFDB and 'lmfdb-lite'.

References

He, Y.-H., Lee, K.-H., Oliver, T. (2022). Machine-learning arithmetic curves. arXiv.https://arxiv.org/abs/2203.13705
Lozano-Robledo, Á. (2011). Elliptic curves, modular forms, and their L-functions (Vol. 58). American Mathematical Society.

KEEPING UP WITH THE JONESES

How to Compute the Jones Polynomial

Amy Wang and Valerie Yu

Formal Intro to Knots and Links

A **link diagram** is a 4-valent plane graph with extra structure associated to its vertices. We call a vertex in a link diagram a **crossing**, and the additional structure associated to it indicates the manner in which we represent the crossings in the diagram as an "over" or "under" crossing.

Formally, a link diagram is a triple (D, V, σ) where D is a 4-valent graph embedded in the plane, V is the set of vertices (crossings) of D, and σ is a function on V specifying the crossing information (i.e. which strands are the over and under strand).

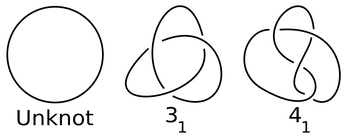
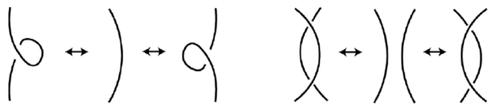


Figure 1: The three most basic knots

Reidemeister moves are local transformations of a link diagram. There are three types of moves:



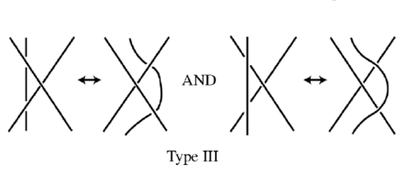


Figure 2: Reidemeister moves

A **link** is defined as an *equivalence class* of link diagrams under the Reidemeister moves. Each disjoint circle represented in the link diagram is called a **component**. A link with one component is called a **knot**.

The simplest link is the class of the empty graph $0 := (\emptyset, \emptyset, \sigma)$, which is known as the *unknot* and traditionally drawn as a plane circle, as shown in Figure 1.

A link **orientation** is a choice of direction in which the link is traversed on each component. An *oriented link diagram* is a link diagram D together with a consistent direction assigned to each edge of the graph, forming closed directed loops. Orientation allows us to define operations such as the connected sum, where the joining of arcs must respect the orientation.

The **connected sum** $K_1 \# K_2$ of two oriented knots is formed by removing a small open arc from each knot, and connecting the resulting endpoints with two new arcs that join them smoothly and preserve orientation. This operation results in a new knot diagram that locally looks like the two knots joined end-to-end.

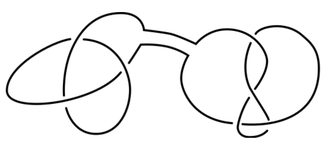


Figure 3: Connected sum of two trefoils

The operation of connect sum makes the set of knots into a monoid. In fact, it is a free monoid isomorphic to $(\mathbb{N}, \times)$. Hence, it makes sense to define prime and composite knots: a knot is called **composite** if there exist nontrivial knots K_1 and K_2 such that $K = K_1 \# K_2$. Otherwise, K is called **prime**.

The Jones Polynomial

A **link invariant** is a quantity that remains unchanged under the Reidemeister moves. Invariants are essential for distinguishing non-equivalent knots and links and organizing tabulated data.

Laurent polynomials are polynomials that allow for both positive and negative powers:

$$f(q) = a_n q^n + \cdots + a_0 + \cdots + a_{-m} q^{-m}.$$

The **Kauffman bracket polynomial** associates a Laurent polynomial to a link diagram D, and is defined recursively by the following rules:

1. $\langle \bigcirc \rangle = 1$.
2. $\langle \times \rangle = A \langle) (\rangle + A^{-1} \langle) (\rangle$.
3. $\langle D \cup \bigcirc \rangle = (-A^2 - A^{-2}) \langle D \rangle$.

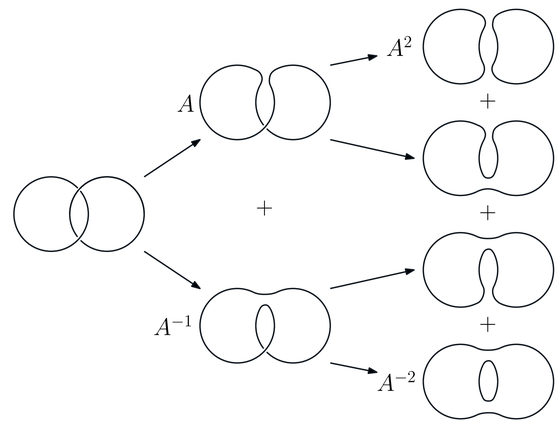


Figure 4: Kauffman bracket polynomial computation for the Hopf link

The Kauffman bracket is not invariant under the first Redeimeister move. To fix this, we define the **writhe** $w(D)$ of an oriented link diagram D is the sum of the signs of its crossings:

$$w(D) = \sum_{c \in V} \operatorname{sign}(c).$$

Each crossing is assigned $+1$ if it is a positive (right-handed) crossing, or -1 if it is a negative (left-handed) crossing.

The **Jones polynomial** $V_L(q)$ is a Laurent polynomial in $\mathbb{Z}[q^{\pm 1/2}]$ assigned to an oriented link L. It is a powerful invariant that can distinguish many (but not all) links.

We define $V_L(q)$ from an oriented diagram D of the link L, using its Kauffman bracket and writhe:

$$V_L(q) = (-A^3)^{-w(D)} \langle D \rangle \quad \text{where } q = A^{-4}.$$

This normalization ensures invariance under the first Reidemeister moves, making $V_L(q)$ a true link invariant.

Based on Figure 4, the Kauffman bracket polynomial of the Hopf link is $\langle D \rangle = -A^4 - A^{-4}$, and the writhe is $w(D) = 2$. So, the Jones polynomial of the Hopf link is:

$$V_L(q) = (-A^3)^{-w(D)} \langle D \rangle = (-A^3)^{-2}(-A^4 - A^{-4}) = A^{-2} - A^{-10} = -q^{\frac{1}{2}} - q^{\frac{5}{2}}.$$

The Jones polynomial is multiplicative under connected sum:

$$V_{K_1 \# K_2}(q) = V_{K_1}(q) \cdot V_{K_2}(q).$$

The Jones polynomial respects the structure of knots under connected sum.

Planar Diagram Notation

The **planar diagram (PD) notation** represents a knot as a list of vertices along with a numbering for each of the four incident edges. Each vertex (denoted X) can be positive or negative, depending on whether the crossing is right-handed or left-handed. The edges are written counterclockwise.

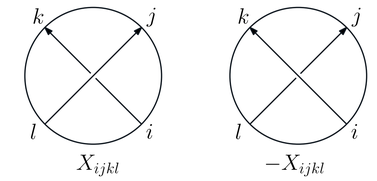


Figure 5: Right-Handedness and Left-Handedness for PD Notation

The right-handed trefoil knot can be represented in PD notation as follows:

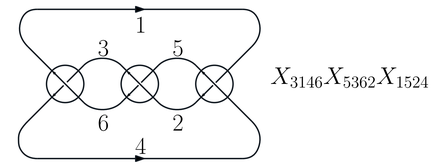


Figure 6: PD Code for a right-handed trefoil

To reconstruct a knot from PD notation, draw all the positive and negative vertices and label its four edges counterclockwise. Then connect like numbers, respecting the orientation of the edges (i.e. outbound edges connect to inbound edges).

PD notation allows us to more easily calculate the Kauffman bracket polynomial (and therefore the Jones polynomial).

Knot Tabulation

Based on [1], we implemented a Python script to generate a nearly complete tabulation of prime knots up to 10 crossings. For alternating knots, we adapted code from [2]. For non-alternating knots, our program computes the Jones polynomial for each knot diagram, distinguishing knots based on their PD notation. *See our code here: https://github.com/awaang/knottery.*

We generated a total of 248 distinct prime knots: 165 alternating and 83 non-alternating. This is two fewer than the 250 listed in Khesin's tabulation because two pairs of non-alternating knots share the same Jones polynomial. This confirms that the Jones polynomial is not a complete knot invariant.

Acknowledgements

We thank our mentor Choomno Moos for their invaluable guidance as well as the UCSB Directed Reading Program for the opportunity to pursue this project.

References

[1] Andrey Boris Khesin. *The 250 Knots with up to 10 Crossings*. 2017. arXiv: 1705.10319 [math.GT]. URL: https://arxiv.org/abs/1705.10319.

[2] Alex Gaither and Mihir Mantri. *The First 197 Alternating Knots*. UCSB DRP. 2024.

[3] Colin Adams. *The Knot Book*. Providence, RI: W.H. Freeman and Company, 2000.

UC **SANTA BARBARA**

Higher Algebraic K-Theory

Pico Gilman[1]

[1]University of California Santa Barbara
Mentored by Jitendra Rathore

Introduction

Algebraic K-theory studies a family of functors that associate rings and algebraic varieties (or more generally exact categories) to abelian groups, known as algebraic K-groups and denoted K_n. The classical theory originated in the 1950s and 1960s through the foundational work of Grothendieck on K_0, followed by Bass and Whitehead on K_1, and Milnor on K_2, establishing the lower K-groups.

A major advancement came in the 1970s when Quillen defined higher K-groups K_n for all natural n, unifying and vastly extending the scope of algebraic K-theory. These groups capture deep arithmetic and geometric information about rings and, more generally, algebraic varieties.

Over time, the theory has revealed deep connections across diverse mathematical disciplines, particularly serving as conceptual bridges amongst algebraic number theory, algebraic geometry, arithmetic geometry, and algebraic topology. This poster outlines the definition of K-groups, their properties, some explicit computations, and their significant connections to number theory and arithmetic geometry.

In this poster, R is a commutative ring with unity and all modules are finitely generated, although the former is not usually necessary.

K_0 (Grothendieck group)

A **projective module** is a summand of a free-module. Projective modules form a monoid $\mathbf{P}$ via $P + P' \coloneqq P \oplus P'$. Over PIDs or local rings, projective modules are free (i.e. have a basis). The group K_0 is defined to be formal group closure of $\mathbf{P}$.

Example. Consider $R = \mathbb{Z}$. Because $\mathbb{Z}$ is a PID, finitely generated projective modules over R are free. Thus, $\mathbf{P} \cong \{\mathbb{Z}^n | n \in \mathbb{N}\} \cong \mathbb{N}$, where the isomorphisms are as monoids. The group closure is thus $\mathbb{Z}$, and thus $K_0(R) \cong \mathbb{Z}$.

Likewise, the K_0 of any PID or local ring is $\mathbb{Z}$. Additionally, the tensor product $\otimes$ distributes over $\oplus$, and thus gives K_0 the structure of an commutative ring with unity $1 \coloneqq R$. By examining the above definition, K_0 can be seen as a functor from commutative rings with unity to abelian groups.

Proposition. The functor K_0 respects binary products. That is, $K_0(R \times R') \cong K_0(R) \times K_0(R')$.

Example. We have $K_0(\mathbb{Z}[\zeta_p]) \cong K_0(\mathbb{Z}[\mu_p])$, where the right side denotes the group ring, and μ_p is the pth roots of unity.

K_0 of a Dedekind domain

A **Dedekind domain** R is a domain such that for any $I \subset J$ ideals there is some ideal $J' \subseteq R$ with $I = JJ'$. In a Dedekind domain, any non-zero fractional ideal (R modules $I \subset \mathrm{Frac}(R) = K$ with $rI \subseteq R$ for some nonzero $r \in R$) is invertible; that is, there is some J with $IJ = R$. Let $\mathcal{I}_K$ be the set of non-zero fractional ideals. It can be given an abelian group structure by $I +_G J \coloneqq IJ$ and contains as a subgroup the principal ideals $\mathcal{P}_K$. The quotient $\mathsf{Cl}(K) \coloneqq \mathcal{I}_K/\mathcal{P}_K$ is the **class group**, which measures lack of unique factorization. The size of the class group is called the **class number**.

Theorem. R is a PID if and only if R is a UFD if and only if $\mathsf{Cl}(K) \cong 1$.

Example. If K is a number field (i.e. $[K : \mathbb{Q}] < \infty$) then $\mathcal{O}_K$ (the elements which are roots of monic polynomials over $\mathbb{Z}$) is a Dedekind domain. We get an exact sequence

$$0 \to \mathcal{O}_K^\times \to K^\times \to \mathcal{I}_K \to \mathsf{Cl}(K) \to 0.$$

Lemma. A finitely generated R-module M is projective if and only if it is isomorphic to a direct sum of ideals.

Theorem. (Steinitz) The map $I_1 \oplus \cdots \oplus I_k \to (k, I_1 \cdots I_k) \in \mathbb{N} \times \mathsf{Cl}(K)$ induces an isomorphism $K_0(R) \cong \mathbb{Z} \oplus \mathsf{Cl}(K)$.

Example. We have $K_0(\mathbb{Z}[(1+\sqrt{-5})/2]) \cong \mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z}$, $K_0(\mathbb{Z}[\zeta_{37}]) \cong \mathbb{Z} \oplus \mathbb{Z}/37\mathbb{Z}$, and $K_0(\mathbb{Z}[\zeta_{29}]) \cong \mathbb{Z} \oplus (\mathbb{Z}/2\mathbb{Z})^{\otimes 3}$.

K_1 (Whitehead Group)

Consider the map $\mathsf{GL}_n(R) \hookrightarrow \mathsf{GL}_{n+1}(R)$ via $A \mapsto \begin{bmatrix} A & 0 \\ 0 & 1 \end{bmatrix}$. This is a directed system with limit $\mathsf{GL}(R)$, which inherits a group structure. Informally, this is the set of infinite matrices with only finitely many differences from the identity. Let $E_n(R)$ denote the subgroup of $\mathsf{GL}_n(R)$ generated by elementary matrices $e_{ij}(r)$ (i.e. identity with the ij entry changed to be r for $i \neq j$). We likewise have $E_n(R) \hookrightarrow E_{n+1}(R)$ and thus some limit $E(R) \subseteq \mathsf{GL}(R)$.

Definition. The K_1 of a ring R is defined to be $\mathsf{GL}(R)/E(R)$.

Proposition. We have $E(R) = [\mathsf{GL}(R), \mathsf{GL}(R)]$, the commutator subgroup of $\mathsf{GL}(R)$.

Corollary. We have $K_1(R) \cong H_1(\mathsf{GL}(R), \mathbb{Z})$, where H_i denotes group homology.

Observation. The determinant of every element of $E_n(R)$ is 1, and thus the unit group $R^\times$ of R will be a direct summand of K_1. Consequently, we have $K_1(R) \cong R^\times \oplus (\mathsf{SL}(R)/E(R))$, where the second summand is denoted SK_1.

Theorem. Let R be a local ring or a Euclidean domain. Then $K_1(R) \cong R^\times$.

Example. Let $R = \mathbb{Z}/15\mathbb{Z} \cong \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/5\mathbb{Z}$. Then $K_1(R) \cong (\mathbb{Z}/3\mathbb{Z})^\times \times (\mathbb{Z}/5\mathbb{Z})^\times \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}$. We also have $K_1(\mathbb{Z}) \cong \{\pm 1\}$ and $K_1(\mathbb{R}) \cong \mathbb{R}^\times \cong \mathbb{R} \times \mathbb{Z}/2\mathbb{Z}$.

Connection to Algebraic Topology. (Mayer-Vietoris) For $I \subseteq R$ an ideal mapped isomorphically via $f : R \to S$ such that quotienting and mapping to S commute, we have the following exact sequence

$$K_1(R) \to K_1(S) \oplus K_1(R/I) \xrightarrow{\pm} K_1(S/I) \xrightarrow{\partial} K_0(R) \to K_0(S) \oplus K_0(R/I) \xrightarrow{\pm} K_0(S/I).$$

K_2 (Milnor Group)

For $n \geq 3$, consider the group $\mathsf{St}_n(R)$ defined by generators $x_{ij}(r)$ for $r \in R$ with the relations

$$x_{ij}(r)x_{ij}(s) = x_{ij}(r+s), \quad [x_{ij}(r), x_{kl}(s)] = \begin{cases} 1 & i \neq l, j \neq k \\ x_{kj}(-rs) & i = l, j \neq k \\ x_{il}(rs) & i \neq l, j = k. \end{cases}$$

There is a natural homomorphism ϕ_n via $x_{ij}(r) \mapsto e_{ij}(r) \subseteq E_n(R)$. Now put $\mathsf{St}(R)$ as the direct limit of $\mathsf{St}_n(R)$. The ϕ define a surjective map $\mathsf{St}(R) \to E(R)$.

Definition. $K_2(R) \coloneqq \mathsf{Ker}(\phi)$. Consequently, the following is exact

$$0 \to K_2(R) \to \mathsf{St}(R) \xrightarrow{\phi} \mathsf{GL}(R) \to K_1(R) \to 0.$$

Theorem. (Steinberg) The center of $\mathsf{St}(R)$ is precisely $K_2(R)$.

Corollary. We have $K_2(R) \cong H_2(E(R), \mathbb{Z})$.

For $R = \mathbb{Z}$, we can compute $K_2(\mathbb{Z}) \cong \mathbb{Z}/2\mathbb{Z}$, but this is a unwieldy definition in general. However, for $R = F$, a field, it reduces by the following theorem.

Theorem. (Matsumoto) The group $K_2(F)$ is isomorphic to the group generated by $\{x, y\}$ with $x, y \in F^\times$ that is linear on each coordinate and satisfies $\{x, 1-x\} = 1$ for $x \neq 0, 1$.

This can be naturally seen via $\{r, s\} = [x_{12}(r), x_{13}(s)]$ and checking the identities.

Computation. We have $-r = (1-r)/(1-r^{-1})$, so

$$\{r, -r\} = \{r, 1-r\}\{r, 1-r^{-1}\}^{-1} = \{r^{-1}, 1-r^{-1}\} = 1.$$
$$\{r, s\}\{s, r\} = \{r, (-r)s\}\{s, (-s)r\} = \{rs, -rs\} = 1.$$

Corollary. We have $K_2(\mathbb{F}_q) = 1$, where $\mathbb{F}_q$ denotes the finite field on q elements.

Proof. Consider some generator x of $\mathbb{F}_q^\times$. Then $\{x, x\}$ generates $K_2(\mathbb{F}_q)$. For q even, we have $\{x, x\} = \{x, -x\} = 1$. For q odd, find some u non-square such that $1 - u$ is not a square by the pigeon hole principle. Then for some odd numbers n, m, we have

$$1 = \{u, 1-u\} = \{x^n, x^m\} = \{x, x\}^{nm} = \{x, x\},$$

where the last equality is since $\{x, x\}^2 = 1$. □

Higher K-groups

For any group G, we write BG to denote the classifying space of G.

Quillen + construction. Given a topological space X and perfect $H \trianglelefteq G = \pi_1(X)$, we say that $f : X \to X^+$ is the $+$ construction if H is the kernel of $f_* : \pi_1(X) \to \pi_1(X^+)$ and induces isomorphisms on all homology groups (this is possible since $H_1 = \pi_1^{\mathrm{ab}}$).

Definition. Consider the perfect normal subgroup $E(R) \subseteq \mathsf{GL}(R)$. Then for $n > 0$,

$$K_n(R) = \pi_n(B\mathsf{GL}(R)^+).$$

Proposition. The functor K_n from rings to abelian groups respects biproducts.

Theorem. We have $K_3(R) = H_3(\mathsf{St}(R), \mathbb{Z})$.

Proof. $K_3(R) \coloneqq \pi_3(B\mathsf{GL}(R)^+) \cong \pi_3(B\mathsf{St}(R)^+) \overset{\star}{\cong} H_3(B\mathsf{St}(R)^+) \cong H_3(\mathsf{St}(R))$, where the $\star$ isomorphism is due to lower homology groups vanishing.

Computation. $K_3(\mathbb{Z}) \cong \mathbb{Z}/48\mathbb{Z}$, and $K_n(\mathbb{Z})$ are related to the **Bernoulli numbers** for $n \equiv 2, 3 \pmod 4$.

Milnor K-Theory

When trying to generalize K-theory, Milnor observed that for fields $R = F$, we have $K_1(F) = F^\times$. The following ad-hoc definition for $n > 0$ turned out to be surprisingly close to K_n, and is now known as Milnor's K-theory.

$$K_n^M(F) = \frac{F^\times \otimes F^\times \otimes \cdots \otimes F^\times}{\langle a_1 \otimes a_2 \cdots \otimes a_n : a_i + a_{i+1} = 1 \rangle}.$$

There exists a natural map $K_n^M(F) \to K_n(F)$ that is an isomorphism for $n \leq 2$. For $n > 2$ this map is no longer necessarily an isomorphism. For example, for $n \geq 2$ we have $K_{2n-1}^M(\mathbb{F}_q) \cong 1$, but $K_{2n-1}(\mathbb{F}_q) \cong \mathbb{Z}/(q^n - 1)\mathbb{Z}$.

Example. One can compute

$$K_2^M(\mathbb{Q}) \cong K_2(\mathbb{Q}) \cong \mathbb{Z}/2\mathbb{Z} \oplus \bigoplus_p \mathbb{Z}/(p-1)\mathbb{Z}.$$

Milnor K-theory ($K_n^M(F)$) is deeply related to arithmetic objects such as the Brauer group $\mathsf{Br}(F)$ via the **Merkurjev-Suslin** Theorem ($n = 2$), the abelianized Galois group $\mathbf{G}_F^{\mathrm{ab}}$ (via **Kato's higher dimensional class field theory**), and more generally ètale cohomology via the **Bloch-Kato conjecture** (n arbitrary), proved by Voevodsky, which says that

$$K_n^M(F)/\ell \cong H_{et}^n(F, \mu_\ell^{\otimes n})$$

for some ℓ invertible in F.

K-theory and Number Theory

Computing K-groups in general is very difficult. For example, we can currently classify $K_n(\mathbb{Z})$ for $n = 0, 4$ or $n \not\equiv 0 \pmod 4$. It is conjectured to be trivial for $4n > 0$.

In 1992, Kurihara showed that this statement is equivalent to p not dividing the class number of the maximal real subfield of $\mathbb{Q}(\zeta_p)$, which is equal to $\mathbb{Q}(\zeta_p + \zeta_p^{-1})$.

In this way, K-theory provides a way to relate purely algebraic results to number theoretic results. Indeed, the K-theory of rings of integers of number fields is always a finitely generated abelian group, and is related to special values of L-functions.

References and Acknowledgements

The author would like to thank the UCSB Mathematics Department for running DRP.

M. John. *Introduction to Algebraic K-Theory*. Princeton University Press, 1971.
M. Kurihara. "Some remarks on conjectures about cyclotomic fields and K-groups of Z". In: *Compositio Mathematica* (1992) p. 223-236
J.L. Loday. *Cyclic homology*. Springer-Verlag New York, Incorporated, 1992.
C. Weibel. *The K-book: An Introduction to Algebraic K-theory*. Graduate Studies in Mathematics 145. American Mathematical Society, 2013.

UC SANTA BARBARA

Primes is in (DR)P

Anikait Mundhra Mentor: Sawyer Dobson

Introduction

The Primes problem is formally defined as: given an integer n, determine whether n is prime. This fundamental decision problem has been of significant interest, with implications for number theory and cryptography.

The complexity of an algorithm is measured in terms of input size. For Primes, this is the $\log n$ bits needed to represent the number n (note we use $\log n$ to represent $\log_2 n$). Many simple algorithms for Primes, such as trial division, take $\mathcal{O}(\sqrt{n}) = \mathcal{O}(2^{\log n/2})$ time, which is exponential in input size. This is impractical for 1024-bit primes used in cryptography, and motivated the search for a polynomial-time algorithm running in $\mathcal{O}(\log^k n)$ for $k \in \mathbf{N}$

In this poster, we trace the evolution from polynomial-time certificates to deterministic polynomial-time algorithms for primality testing.

Complexity Classes: P and NP

The complexity class **P** contains problems solvable in polynomial time.

- Primes is in **P** if and only if there exists a $\mathcal{O}(\log^k n)$ algorithm for some constant k.

The complexity class **NP** contains problems verifiable in polynomial time using certificates, which can be thought of as proofs.

- Primes is in **NP** if and only if there exists an $\mathcal{O}(\log^k n)$ algorithm for some constant k that takes as input a number n, a decision d and a certificate c and outputs if the decision is correct.

A certificate enables verification without solving the original problem. For primality, a certificate allows Bob to verify Alice's claim that n is prime/composite without independently determining primality. We start with showing that primality testing is in **NP**.

Primes is in NP (1975)

To prove Primes $\in$ **NP**, we need certificates for two cases:

- n is composite: let c be any non-trivial factor of n (to verify, check if n is divisible by c)
- n is prime: More complex certificate based on primitive roots

***Theorem* 1** (Fermat's Little Theorem Converse)**.**
For all $n \in \mathbf{N}$, n is prime if and only if there exists a witness $a < n$ such that

1. $a^{n-1} \equiv 1 \pmod n$
2. $a^{\frac{n-1}{q}} \not\equiv 1 \pmod n$ *for all prime factors q of $n-1$*

Note a is a primitive root modulo n.

We inductively construct a certificate for $n > 3$ being prime as follows. As n is prime, there is some primitive root a modulo n. Let the certificate for n be the pair (a, n) alongside the factorization of $n-1 = q_1 \cdots q_k$, where each q_i is prime. We attach the certificate of primality of each q_i to the certificate for n.

if we inductively assume the number of the integers in the certificate for each q_i is at most $5\lceil \log q_i \rceil - 5$, then the number of integers of the certificate for n

$$2 + k + \sum_{i=1}^{k} (5\lceil \log q_i \rceil - 5) \le 2 + 5\lceil \log n \rceil - 3k \le 5\lceil \log n \rceil - 5$$

so the inductive hypothesis holds. Note the size of the certificate is bounded above by $(5\lceil \log n \rceil - 5)\lceil \log n \rceil = \mathcal{O}(\log^2 n)$, and the constants are chosen to satisfy the base case.

Bob can verify this certificate in polynomial time by:

1. Verifying that $\prod_{i=1}^{k} q_i = n - 1$ (polynomial in $\log n$)
2. Verifying $a^{n-1} \pmod n \equiv 1$ (polytime by repeated squaring)
3. For each q_i, verifying $a^{(n-1)/q_i} \pmod n \not\equiv 1$ (polytime by repeated squaring)
4. Verifying the primality of each q_i

The certificate size is $\mathcal{O}(\log^2 n)$ bits. With a similar induction, we can show verification takes $\mathcal{O}(\log^4 n)$ time. Bob can verify primality of n in polynomial time, so Primes $\in$ **NP**.

Probabilistic Primality Tests

Monte Carlo algorithms offer efficient primality testing with one-sided error: always correctly identify primes, but may misclassify composites as primes with bounded probability. The key insight involves randomly selecting witnesses that reveal compositeness. Multiple iterations exponentially reduce error probability, enabling testing of numbers far beyond deterministic methods' capabilities.

Solovay-Strassen Algorithm (1977)

The Solovay-Strassen test uses the Jacobi symbol and the Euler criterion for primality testing. We define $\left(\frac{\cdot}{\cdot}\right)$ to be the Jacobi symbol, where

$$\left(\frac{a}{n}\right) = \begin{cases} 0 & \text{if } \gcd(a,n) > 1 \\ 1 & \text{if } n \text{ is prime and } a \equiv x^2 \pmod n \text{ for some } x \neq 0 \\ -1 & \text{if } n \text{ is prime and } a \not\equiv x^2 \pmod n \text{ for all } x \\ \prod_{i=1}^{k} \left(\frac{a}{p_i}\right)^{e_i} & \text{if } n = \prod_{i=1}^{k} p_i^{e_i} \quad \text{is composite} \end{cases}$$

***Theorem* 2** (Euler Criterion)**.** *Let p be a prime. Then $a^{(p-1)/2} \equiv \left(\frac{a}{p}\right) \pmod p$ for all $a < p$.*

Algorithm 1 Solovay-Strassen (n, k)

```
for i = 1 to k do
    Choose random a ∈ [2, n − 1]
    if gcd(a, n) > 1 then
        return composite
    if a^{(n−1)/2} ≡ (a/n) (mod n) then
        return probably prime
return composite
```

Proof Sketch. For prime n, the algorithm always returns "probably prime" by Euler's criterion. For composite n, let $G = \{a \in \mathbf{Z}_n^\times \mid a^{(n-1)/2} \equiv \left(\frac{a}{n}\right) \pmod n\}$. Note G is a group, and the probability of returning prime when composite is at most $|G|/|\mathbf{Z}_n^\times|$. So, if we show $G \neq \mathbf{Z}_n^\times$, then the probability of error is at most $|G|/|\mathbf{Z}_n^\times| \le 1/2$.

For the general case, assume for contradiction that n is composite (but not a prime power nor a perfect square) and $G = \mathbf{Z}_n^\times$. Then for all $a \in \mathbf{Z}_n^\times$, $a^{(n-1)/2} \equiv \pm 1 \pmod n$. Let $n = r \cdot s$, with $(r, s) = 1$. If $a^{(n-1)/2} \equiv -1 \pmod n$ for some a, then we can find b such that $b \equiv 1 \pmod r$ and $b \equiv a \pmod s$. Then $b^{(n-1)/2} \equiv 1 \pmod r$ and $b^{(n-1)/2} \equiv -1 \pmod s$, a contradiction. So, $a^{(n-1)/2} \equiv 1 \pmod n$ for all $a \in \mathbf{Z}_n^\times$. Yet, this means $\left(\frac{a}{n}\right) = 1$ for all $a \in \mathbf{Z}_n^\times$, which implies n must be a perfect square, a contradiction. □

Using repeated squaring, each iteration of the algorithm runs in time $\mathcal{O}(\log^2 n)$ time, and after k iterations, the one-sided error probability is at most $1/2^k$.

Miller-Rabin Algorithm (1980)

Miller-Rabin improves on Solovay-Strassen with a $1/4$ error probability per iteration. The key insight is that if n is prime, ± 1 are the only square roots of 1 modulo n. If we let $n - 1 = 2^s \cdot d$ (where d is odd), for all $a < n$, either $a^d \equiv 1 \pmod n$ or $a^{2^j d} \equiv -1 \pmod n$ for some $j \in [0, s-1]$. For composite numbers, this condition fails for $3/4$ of possible choices of a.

Algorithm 2 Miller-Rabin (n, k)

```
Write n − 1 = 2^s · d where d is odd
for i = 1 to k do
    Choose random a ∈ [2, n − 1]
    if a^d ≡ 1 (mod n) then continue to next iteration
    for j = 1 to s − 1 do
        if a^{2^j d} ≡ −1 (mod n) then continue to next iteration
▷ If we reach here, then a^d ≢ 1 (mod n) and a^{2^j d} ≢ −1 (mod n) for all j ∈ [0, s − 1]
    return composite
return probably prime
```

Using repeated squaring, each iteration runs in $\mathcal{O}(\log^3 n)$ time, and after k iterations, the error probability is at most $1/4^k$.

AKS Algorithm: Primes is in P (2002)

The AKS algorithm places primality testing in **P** with a deterministic polynomial-time algorithm running in $\mathcal{O}(\log^{11} n)$. It uses a key insight about polynomial identities over finite rings.

Notice for any $n \in \mathbf{N}$, a coprime to n, the binomial expansion $(X+a)^n \equiv X^n + a \pmod n$ if and only if n is prime. So, we can test if n is prime by checking if $(X+a)^n \equiv X^n + a \pmod n$ for our favorite $a < n$. Yet, this takes $\mathcal{O}(n)$ time, evaluating every coefficient of the binomial expansion. Instead, we try to evaluate if

$$(X+a)^n \equiv X^n + a \pmod{X^{r-1}, n}$$

where r is carefully chosen to be polynomial in $\log n$, and large enough to ensure n is composite if and only if the above identity fails.

We define the order $\mathrm{ord}_r(a)$ to be the smallest positive integer k where $a^k \equiv 1 \pmod r$.

Algorithm 3 AKS Primality Test (n)

```
if n is a perfect power then return composite
Find smallest r such that ord_r(n) > 4 log^2 n
for a = 1 to r do
    if gcd(a, n) > 1 then return composite
for a = 1 to ⌊√φ(r) log n⌋ do
    if (X + a)^n ≠ X^n + a (mod X^r − 1, n) then return composite        (∗)
return prime
```

Proof Sketch. If n is prime, then $(X+a)^n \equiv X^n + a \pmod n$ for all $a < n$, thus the algorithm returns prime.

The challenge is proving that composite numbers always fail the test. Let n be a number which passes all the congruences $(*)$ and reaches the end of the algorithm. For sake of contradiction, say n is composite. Let prime $p \mid n$. As $(r, n) = 1$, we define the group

$$I = \{n^i p^j \pmod r \mid i, j \ge 0\}.$$

We have $t := |I| \ge \mathrm{ord}_r(n) > 4\log^2 n$. As p is prime, $(X+a)^{n^i p^j} \equiv X^{n^i p^j} + a \pmod{X^r - 1, p}$ for all $i, j \ge 0$. So, let $h(X)$ be an irreducible factor of $(X^r - 1)/(X - 1)$ over $\mathbb{F}_p$. Let $\ell = \lfloor \sqrt{\phi(r)} \log n \rfloor$. We define the group

$$J = \langle (X+1),\ (X+2),\ \cdots,\ (X+\ell) \pmod{h(X), p} \rangle$$

As n passes $(*)$ for all $a < \ell$, for all $f(X) \in J$, $f(X)^n \equiv f(X^N) \pmod{h(X), p}$. Further analysis using the congruences of I above with the elements of J shows that $|J| \ge 2^{\min \ell, t} > n^{2\sqrt{t}}$ (hint, compare elements of J with degree $\le t$). Yet, with some more work, we can show that $|J| \le n^{\sqrt{t}}$ (hint: find $m_1 > m_2$ such that each $f(X) \in J$ is a root of $Q(Y) = Y^{m_1} - Y^{m_2} \pmod{h(X), p}$, then compare $|J|$ to m_1). This is a contradiction, so n must be prime. □

Conclusions

A solution can be theoretically elegant yet impractical. Though AKS definitively showed that primality testing is in **P**, its $\mathcal{O}(\log^{11} n)$ (later improved to $\mathcal{O}(\log^6 n)$) complexity means it's still too slow for most applications. In practice, Miller-Rabin dominates. Other approaches include Goldwasser-Kilian's elliptic curve primality proving, which creates primality certificates running in expected polynomial time. This method constructs elliptic curves over finite fields whose orders can be factored to recursively prove primality.

Moving forward, similar to how primality testing was not known to be in **P** for a long time, prime factorization is still believed to be outside **P**, and still remains open.

References and Acknowledgments

I would like to give my gratitude and appreciation to my mentor, Sawyer Dobson, for his guidance, patience, and everlasting support throughout this project.

- Pratt, V. (1975). Every prime has a succinct certificate. *SIAM Journal on Computing*.
- Solovay, R., & Strassen, V. (1977). A fast Monte-Carlo test for primality. *SIAM Journal on Computing*.
- Rabin, M. O. (1980). Probabilistic algorithm for testing primality. *Journal of Number Theory*.
- Goldwasser, S., & Kilian, J. (1986). Almost all primes can be quickly certified. *ACM STOC*.
- Agrawal, M., Kayal, N., & Saxena, N. (2004). PRIMES is in P. *Annals of Mathematics*.

Introduction to the Theory of Computation

Harvey Cho, Pranav Hegde, Kai Maeda, Chris Wagner

University of California Santa Barbara

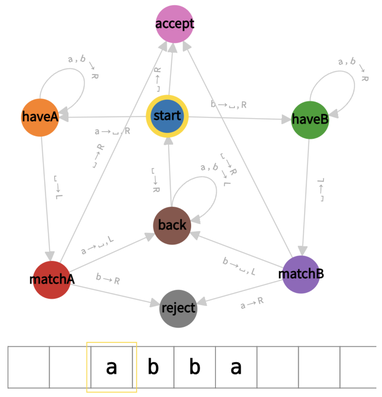
Turing Machines

Definition

- A **Turing machine (TM)** is a model proposed by Alan Turing in 1936. Similar to a general purpose computer but with unlimited and unrestricted memory. An infinite tape is used to represent the infinite memory of a Turing machine. It has a tape head that can read and write symbols and move left and right on the tape. The initial state of the tape is the input string and is blank everywhere else.
- A **configuration** is defined as the current state, current tape contents, and current head location. The information in the configuration is needed to understand how our state machine should behave.
- Each Turing machine has a corresponding state machine and given an input string, it will continue transitioning from configuration to configuration until it reaches an accept or reject state. Otherwise it may never halt.

Example

- Using the diagram below we see an example of a Turing machine that accepts palindromes using the alphabet {a,b}. The circles represent states, the arrows represent transitions, the tape contains the current contents, and the boxed letter is the current head location. The transitions can be defined as $Q \times \Gamma \to Q \times \Gamma \times \{L, R\}$ st Γ is the tape alphabet and Q is the set of states. We can take the transition if the first Γ contains the symbol at the current head. The second Γ represents the symbol we replace the current head with. And the L,R at the end represents whether we should move the tape head left or right.
- Let us consider what would happen on the current configuration. We would take the $a, \to \sqcup, R$ transition. This means we would update our current state to the state "haveA", replace the current head symbol with a blank and move the tape head right.



Decidable versus Recognizable

- The collection of strings that a Turing machine M accepts is defined as the language of M, denoted as L(M).
- A language is A **decidable** if there exists a Turing machine that accepts all strings in A and rejects all strings not in A.
- A language A is **recognizable** if there exists a Turing machine that accepts all strings in A and if the string is not in A, the Turing machine might either reject or loop forever.

Nondeterministic Turing Machine (NTM)

- The transition function for an NTM maps to a *set* of possible moves: $\delta : Q \times \Gamma \to \mathcal{P}(Q \times \Gamma \times \{L, R\})$, where Q is the set of states, Γ the tape alphabet, and $\mathcal{P}$ denotes the power set.
- A computation forms a *tree* of branches, where each branch follows one sequence of nondeterministic choices.
- The machine *accepts* if **any** computation branch reaches the accept state.

Time and Space Complexity of Turing Machines

Big 'O' Notation

Let $f, g : \mathbb{N} \to \mathbb{R}^+$. We say:

$$f(n) = O(g(n)) \iff \exists\, c > 0,\ n_0 \in \mathbb{N},\ \forall n \geq n_0,\ f(n) \leq c \cdot g(n)$$

This means that $g(n)$ is an *asymptotic upper bound* for $f(n)$. Constants and lower-order terms are ignored. **Little-'o' notation:**

$$f(n) = o(g(n)) \iff \lim_{n\to\infty} \frac{f(n)}{g(n)} = 0$$

This means $f(n)$ grows strictly slower than $g(n)$ asymptotically.

Time Complexity

- Let M be a deterministic Turing Machine that halts on all inputs. The **time complexity** of M is the function $f : \mathbb{N} \to \mathbb{N}$, where $f(n)$ is the maximum number of steps M uses on any input of length n.
- Let N be a nondeterministic Turing Machine that halts on all branches. Its time complexity is also a function $f : \mathbb{N} \to \mathbb{N}$, where $f(n)$ is the maximum number of steps taken on any branch of computation on inputs of length n.

Time complexity classes:

- TIME$(t(n)) = \{L \mid$ some TM decides L in $O(t(n))$ time$\}$
- NTIME$(t(n)) = \{L \mid$ some NTM decides L in $O(t(n))$ time$\}$

Space Complexity

- Let M be a deterministic Turing Machine that halts on all inputs. The **space complexity** of M is the function $f : \mathbb{N} \to \mathbb{N}$, where $f(n)$ is the maximum number of distinct tape cells that M scans on any input of length n.
- If M is a nondeterministic TM where all branches halt, then the space complexity is also a function $f : \mathbb{N} \to \mathbb{N}$, where $f(n)$ is the maximum number of tape cells scanned on any branch of computation for inputs of length n.

Space complexity classes:

- SPACE$(f(n)) = \{L \mid L$ is decided by a TM using $O(f(n))$ space$\}$
- NSPACE$(f(n)) = \{L \mid L$ is decided by an NTM using $O(f(n))$ space$\}$

Example Problem: Given a Boolean formula ϕ in conjunctive normal form (CNF) — that is, an AND of OR-clauses over variables and their negations — with n variables and m clauses, decide if there exists a satisfying assignment.

Deterministic TM:

- *Time:* $O(2^n \cdot m)$ — tries all 2^n assignments and checks each.
- *Space:* $O(n + m)$ — input + current assignment + working memory.

Nondeterministic TM:

- *Time:* $O(n + m)$ — guesses one assignment and verifies in linear time.
- *Space:* $O(n + m)$ — similar to DTM.

The P vs NP Problem

Definition: P is the class of languages that are decidable in polynomial time on a deterministic single-tape TM, i.e.

$$\mathsf{P} = \bigcup_k \mathsf{TIME}(n^k).$$

Some examples of languages in P are:

- $PATH = \{\langle G, s, t\rangle \mid G$ is a directed graph that has a directed path from s to $t\}$
- $PRIMES = \{\langle x\rangle \mid x$ is prime$\}$ is in P by the AKS algorithm [1], which is based on a generalization of Fermat's Little Theorem.

Definition: NP is the class of languages that are decidable in polynomial time on a nondeterministic TM, i.e.

$$\mathsf{NP} = \bigcup_k \mathsf{NTIME}(n^k).$$

- $CLIQUE = \{\langle G, k\rangle \mid G$ is an undirected graph with a k-clique$\}$
- $SAT = \{\langle \phi\rangle \mid \phi$ is a satisfiable Boolean formula$\}$

Does P = NP?

- P $\subseteq$ NP because any problem that can be solved in polynomial time can also be verified in polynomial time.
- Currently, we believe that P $\neq$ NP, but this hasn't been proven using current methods.
- If P = NP, every efficiently verifiable problem could also be efficiently solved. This would break systems in many fields such as cryptography, since many rely on problems in NP that are not believed to be in P.

Hierarchy Theorems: Separating Complexity Classes

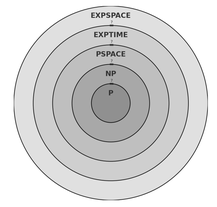
More Time, More Power?

- The **Time Hierarchy Theorem** proves that giving a Turing machine more time allows it to solve strictly more problems, provided the extra time is constructible.
- Here, "constructible" time simply means the machine can keep track of how many steps it is taking, within that same time limit.
- **Time Hierarchy Theorem**:
  For any time constructible function $t(n)$, a language A exists that is decidable in time $O(t(n))$ but not decidable in time $o(\frac{t(n)}{\log t(n)})$.
- In an analogous way, the **Space Hierarchy Theorem** states that more usable space also leads to strictly greater computational power.

Ideas Behind the Proof of the Time Hierarchy Theorem

- Given a time constructible function $t(n)$, we want to show the existence of a language that is decidable in time $O(t(n))$ but not in time $o(\frac{t(n)}{\log t(n)})$. To do this, we define a TM D that, on input of the form <M>10^* where M is a Turing machine, simulates M on <M>10^* for at most $\frac{t(n)}{\log t(n)}$ steps. If M halts within this time limit, D outputs the opposite of M's output. If not, D rejects. This guarantees that D differs from every machine running in $o(\frac{t(n)}{\log t(n)})$ time. The language decided by D, call it A, is therefore decidable in time $O(t(n))$ but not in time $o(\frac{t(n)}{\log t(n)})$.
- Currently, when we simulate M using D, we incur an extra logarithmic time overhead from keeping track of the time taken. If we had an algorithm to simulate any Turing machine for a given number of steps with only a constant-factor slowdown, then the theorem would be strengthened by changing $o(\frac{t(n)}{\log t(n)})$ to $o(t(n))$; but no such efficient simulation is known.

Class Separations from the Hierarchy Theorems



- As direct consequences of the Hierarchy theorems, we get that

  P $\subsetneq$ EXPTIME and PSPACE $\subsetneq$ EXPSPACE

Acknowledgements

We would like to thank our mentor, Sawyer Dobson, for his guidance as well as the UCSB Directed Reading Program for the opportunity to work on this project.

References

[1] Manindra Agrawal, Neeraj Kayal, and Nitin Saxena. "PRIMES is in P". In: *Annals of Mathematics* 160.2 (2004), pp. 781–793. DOI: 10.4007/annals.2004.160.781.

[2] Michael Sipser. *Introduction to the Theory of Computation*. Cengage Learning, 2013.

The Quillen-Suslin Theorem

Nathaniel Hurst Michael Zhou Yue Cao Mentor: Waqar Ali Shah

University of California - Santa Barbara
Department of Mathematics - Directed Reading Program 2025

Serre's Problem

In 1955, Jean-Pierre Serre asked whether every finitely generated projective module over the ring $k[x_1, \ldots, x_n]$, where k is a field, is free. The geometric motivation behind this question was that the affine scheme underlying $k[x_1, \ldots, x_n]$ is the affine n-space $\mathbb{A}^n_k$ and algebraic vector bundles over $\mathbb{A}^n_k$ correspond to finitely generated projective modules over $k[x_1, \ldots, x_n]$. Since the real affine space $\mathbb{R}^n$ is contractible, every topological (and even smooth) vector bundle over it is trivial. A similar argument shows that $\mathbb{C}^n$ admits no non-trivial holomorphic vector bundles. Thus, if Serre's question had an affirmative answer, it would imply the analogous claim in the algebraic setting. It took 21 years, but the statement was eventually proved by Daniel Quillen and Andrei Suslin, independently of each other.

Projective Modules

Definition: Let A be a commutative ring with identity. An A-module P is called **projective** if every short exact sequence

$$0 \longrightarrow K \longrightarrow V \xrightarrow{\varphi} P \longrightarrow 0$$

splits. In other words, every surjective morphism $\varphi : V \twoheadrightarrow P$ admits a *section map*, i.e., a homomorphism $\psi : P \to V$ such that $\varphi \circ \psi$ is the identity map on P.

A consequence of this definition is that projective modules are precisely those that arise as direct summands of free modules. Indeed, if P is projective, the obvious surjection $\varphi : A^{\oplus x \in P} \twoheadrightarrow P$ induces the identification $A^{\oplus x \in P} \cong \ker(\varphi) \oplus P$. The converse is apparent.

Stably Free

In 1957, Serre took the first step toward answering his question by showing that every finitely generated projective module over $k[x_1, \ldots, x_n]$ is, in a sense, "almost" free.

Definition: An A-module U is **stably free** if $U \oplus A^n \cong A^m$ for some $n, m \in \mathbb{Z}_{\geq 0}$.

This definition resembles that of projective modules; however, it additionally requires that the complement in the direct sum decomposition of the free module is itself free. By definition, every stably free module is projective. To show that every projective module over $k[x_1, \ldots, x_n]$ is stably free, we need to define the notion of finite free resolution.

Definition: We say that an A-module M admits a **finite free resolution** if there exists an exact sequence of finite length

$$0 \to E_n \to \cdots \to E_0 \to M \to 0$$

such that each E_i is free of finite rank.

Proposition: *Let M be a projective A-module. Then M is stably free if and only if M admits a finite free resolution.*

The forward direction of this theorem is trivial. The idea of the converse is to induct on the length of the finite free resolution, taking the kernel of the map $E_1 \to E_0$ and constructing a new exact sequence where the kernel has a shorter finite free resolution than M.

Theorem: Let R be a commutative noetherian ring. If every finite R-module admits a finite free resolution, then every finite $R[x]$-module admits a finite free resolution.

This proof is long, but once we have this result we can use induction to show that every projective module over $k[x_1, \ldots, x_n]$ is stably free. See [1, Chapter XXI].

The question that now arises is when stably free modules are actually free. It turns out that if a ring satisfies certain elementary matrix-theoretic conditions, which we will elaborate on shortly, then all stably free modules over that ring are automatically free. Thus the challenge was to prove that the polynomial ring $k[x_1, \ldots, x_n]$ satisfied such matrix-theoretic properties. While such properties were known for polynomial rings over local domains, extending it to arbitrary rings required new ideas.

Acknowledgments

We would like to thank our mentor Waqar Ali Shah for his insight and guidance throughout this project. Additionally we would like to thank the DRP for organizing this project.

Quillen-Suslin & $k[x_1, \ldots, x_n]$

The jump from stably free to free took significantly more time and is by no means straightforward. However, a completely elementary version of the proof of the Quillen–Suslin Theorem was given by Leonid Vaserštein, and it is this proof that we will discuss. We follow the exposition given in [1, Chapter XXI].

Definition: Let A be a commutative ring with identity. We call a vector $(f_1, \ldots, f_n) \in A^n$ **unimodular** if its elements generate the unit ideal in A. Additionally, we say that the vector has the **unimodular extension property** if there exists an invertible square matrix with entries in A whose first column is $(f_1, \ldots, f_n)^T$. We say that two unimodular vectors f and g are **equivalent** if there exists an invertible matrix M such that $f = Mg$ and we write $f \sim g$ to denote this.

The first step in Vaserštein's proof is to use a specific result about the unimodular extension property for a polynomial ring over local domains. This is known as Horrocks' theorem.

Theorem: *Let ω be a local ring and let f be a unimodular vector in $\omega[x]^n$ such that some component of f is monic. Then f has the unimodular extension property.*

The proof uses the relation $\sum g_i f_i = 1$ where $g_i \in \omega[x]$ and elementary row operations to induct down on the highest degree of a monic entry in f to show that any unimodular vector in $\omega[x]^n$ is equivalent to the first standard basis vector. An immediate corollary of this is that $f \sim f(0)$ over $\omega[x]$. The next step is to globalize Horrocks' result.

Proposition: *Let R be an integral domain and let f be a unimodular vector in $R[x]^n$ such that some component of f is monic. Then $f \sim f(0)$ over $R[x]$.*

This proof uses a lemma in which we go from the local result of Horrocks to two variables by shifting $x \mapsto x + cy$, where the set of all possible c's which we can shift turns out to be the whole of R. We use this fact to get a global result that $f \sim f(0)$ over $R[x]$. Finally, we have the full Quillen-Suslin theorem.

Theorem: *Let k be a field. Then every finitely generated projective module over the polynomial ring $k[x_1, \ldots, x_n]$ is free.*

This result follows by establishing that $k[x_1, \ldots, x_n]$ has the unimodular column extension property for all unimodular vectors, not just those with some component monic. This is achieved via a clever substitution of variables. More specifically, we make the substitution $x_n = y_n$, and $x_i = y_i - y_n^{r_i}$ for some collection of sufficiently large r_i. This coupled with the fact that any finitely generated projective module is stably free gives us the full result. To elaborate, every finitely generated projective module M over $A := k[x_1, \ldots, x_n]$ is stably free by the work of Serre. One now appeals to the unimodular extension property of A to show that if $M \oplus A^n \cong A^m$ for some $m, n \in \mathbb{Z}_{\geq 0}$, then $m \geq n$ and $M \cong A^{m-n}$.

Spec($k[x_1, \ldots, x_n]$) & Affine Space

Suppose k is an algebraically closed field. Then the prime spectrum of the ring $A := k[x_1, \ldots, x_n]$ can be identified with a topological space that is essentially the affine space $\mathbb{A}^n_k$, equipped with the so-called *Zariski topology*. This follows from **Hilbert's Nullstellensatz**, which shows that the maximal ideals of $k[x_1, \ldots, x_n]$ are all of the form $(x_1 - a_1, \ldots, x_n - a_n)$ for some $a_1, \ldots, a_n \in k$. See [2, §15.3] for details. We can now identify the maximal ideal $(x_1 - a_1, \ldots, x_n - a_n)$ with the point $(a_1, \ldots, a_n)$ in the affine n-space, i.e.,

$$k^n \xrightarrow{\sim} \mathrm{MaxSpec}(k[x_1, \ldots, x_n]) \subseteq \mathrm{Spec}(k[x_1, \ldots, x_n]).$$

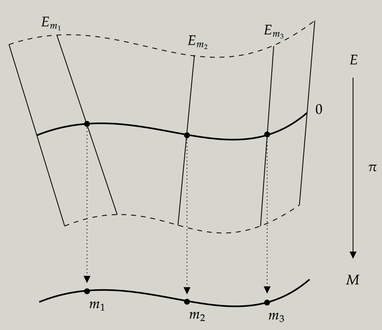
It can be shown that any finitely generated projective module M over A is finite locally free; that is, for each $\mathfrak{p} \in \mathrm{Spec}(A)$, there exists an element $f \in A \setminus \mathfrak{p}$ such that M_f is free over A_f of fixed finite rank. This fact admits a geometric interpretation: M corresponds to an *algebraic vector bundle* in the Zariski topology, and the freeness of M_f over A_f reflects the property of *local triviality*, which we now explain.

References

[1] Serge Lang. *Algebra*. New York, NY: Springer, 2012.
[2] David S. Dummit and Richard M. Foote. *Abstract Algebra*. Wiley, NJ, 2003.
[3] Ravi Vakil. *The Rising Sea: Foundations of Algebraic Geometry*. PUP, NJ, 2023.
[4] Karen E. Smith et al. *An Invitation to Algebraic Geometry*. Springer, NY, 2003.

Vector Bundles

Vector bundles generalize the idea of families of vector spaces. Formally, a vector bundle is a map $\pi : E \to M$, where E is the "total space", and M is a topological space (e.g., the prime spectrum of a ring), called the "base space". The fiber $\pi^{-1}(\{m\})$ over a point $m \in M$ is the vector space associated with m, denoted E_m. The fibers over points in M vary continuously. These collections are *locally trivial*, which means that around every point $m \in M$, there is some open set U containing m such that the fibers $\pi^{-1}(U)$ look locally like $U \times k^n$, for some field k and natural number n. See [3, §14.1] for details.



A trivial vector bundle

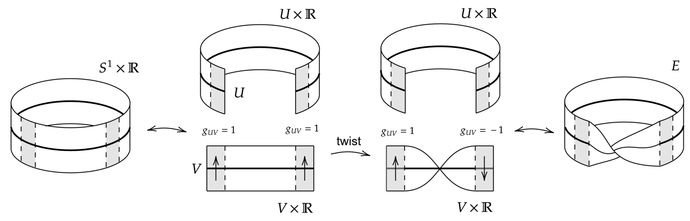
However, a vector bundle is not necessarily globally trivial. Typically, it is constructed by taking locally trivial collections of fibers and gluing them together via *transition functions* along overlaps of local trivializations. Take two trivial collections of fibers $V \times k^n$, and $U \times k^n$. We define a transition function as a map $g_{UV} : U \cap V \to GL_n(k)$. This function "glues" the two collections together by enforcing agreement on their overlap. Moreover, this agreement is consistent with any third such collection, provided the transition functions satisfy the so-called *cocycle condition*. If the base space is an algebraic variety, then an algebraic vector bundle is one whose local trivializations are glued together using algebraic transition functions.

Over affine spaces, all algebraic vector bundles are trivial, as classified by the Quillen–Suslin theorem. The classification of algebraic vector bundles over other algebraic varieties, especially over *projective spaces*, remains an active area of research. A **projective space** of dimension n over k is the set of one dimensional subspaces of $\mathbb{A}^{n+1}_k$. Intuitively, this is just the set of lines through the origin of the vector space k^{n+1}. A geometric interpretation of projective space is just taking affine space and adding a "point at infinity".

One obvious way to define a one dimensional vector bundle (also called a *line bundle*) on projective space is that of the **tautological line bundle**, $\mathcal{O}_{\mathbb{P}^n}(-1)$. This bundle assigns to each point in $\mathbb{P}^n$ the one-dimensional vector space it is associated with in $\mathbb{A}^{n+1}$. See [4, §8.4]

Classifying general vector bundles over projective n-space is quite difficult. For projective space of dimension one, a celebrated theorem of Alexander Grothendieck asserts that every vector bundle splits as a direct sum of line bundles obtained by "twisting" the tautological bundle. However, as the dimension increases, vector bundles become more intricate, and there exist non-split vector bundles on $\mathbb{P}^n$ for all $n \geq 2$. On the other hand, a famous conjecture of Hartshorne from the 1970s asserts that all rank two vector bundles over $\mathbb{P}^n$ for $n \geq 7$ split as the direct sum of two line bundles. As of 2025, this conjecture remains open.

Example: To illustrate how non-trivial vector bundles can arise, let us consider the **Möbius strip** as a line bundle on the real projective space $\mathbb{RP}^1$. This can be constructed by taking an affine line, gluing its ends to shape it into a circle and inserting a twist in the middle.
More specifically, we take two local trivializations $U \times \mathbb{R}$ and $V \times \mathbb{R}$. The globally trivial line bundle is obtained by gluing these two charts using the identity transition function $g_{UV} = 1$. The Möbius strip, on the other hand, arises by using the transition function $g_{UV} = -1$. This construction is essentially that of the tautological line bundle $\mathcal{O}(-1)$ on $\mathbb{RP}^1$.



Construction of the Möbius strip

Chopping up three dimensional spaces

Andrew Sylvester Mentor: Rhea Palak Bakshi
Directed Reading Program 2025

1. Manifolds

An n-manifold is a space such that at every point, one can travel in n independent directions. There are multiple ways to rigorously define manifolds.

A **topological n-manifold** M is a (Hausdorff and second countable) topological space that is locally Euclidean of dimension n, i.e., each point in M has an open neighborhood that is homeomorphic to $\mathbb{R}^n$. A map between topological manifolds is a continuous map between the underlying topological spaces. Two topological manifolds are considered equivalent if there is a homeomorphism between them.

Given an open set $U \subseteq M$ and a homeomorphism $\varphi\colon U \to \mathbb{R}^n$, the pair (U, φ) is called a chart for M. A collection $\{(U_i, \varphi_i)\}_{i\in I}$ of charts such that the U_i cover M is called an atlas for M. Given two charts (U, φ) and (V, ψ), the map $\psi|_{U\cap V} \circ \varphi^{-1}\colon \varphi(U \cap V) \subseteq \mathbb{R}^n \to \mathbb{R}^n$ is called the transition map from (U, φ) to (V, ψ).

A **smooth n-manifold** is a topological n-manifold with an atlas such that all transition maps are differentiable as functions from open subsets of $\mathbb{R}^n$ to $\mathbb{R}^n$. This allows for a well-defined notion of differentiable maps between smooth manifolds.

There also exist **piecewise linear (PL) n-manifolds** with appropriately defined PL maps. Their definition is combinatorial in nature.

Theorem (Moise, Bing, Hamilton): Every topological 3-manifold has a unique PL structure. **Theorem (Whitehead):** Every PL 3-manifold has a unique smooth structure. Together, this implies that the three categories of topological, smooth, and PL manifolds are equivalent for 3-dimensional manifolds. Since manifolds are topological spaces, we can require them to be compact or connected, and we can combine them by taking products, disjoint unions, and gluing along specific maps.

2. Prime Decomposition

The connected sum $M\#N$ of two connected oriented n-manifolds is obtained by removing an open n-ball from each manifold and gluing the resulting manifolds along the $(n-1)$-spheres in their boundaries via an orientation reversing homeomorphism. The 3-sphere behaves as the identity element of this operations.

A 2-sphere embedded in a 3-manifold is called separating if $M \setminus S^2$ has two components. It is a non-separating 2-sphere if $M \setminus S^2$ is connected. If X and Y are the two components of $M \setminus S^2$ for a separating 2-sphere, then $M = X\#Y$.

A manifold is called prime if it cannot be written as the connected sum of two manifolds that are not S^3. A manifold is called irreducible if every embedded 2-sphere bounds an embedded 3-ball. The only oriented 3-manifold that is prime but not irreducible is $S^2 \times S^1$ because it has a non-separating 2-sphere.

Theorem (Kneser, Haken): Every compact orientable 3-manifold factors uniquely into a connected sum of prime 3-manifolds.

3. Handle Decomposition

For $0 \le p \le n$, we can write $D^n \cong D^p \times D^{n-p}$. This decomposes the boundary into

$$\partial(D^p \times D^{n-p}) = (\partial D^p \times D^{n-p}) \sqcup (D^p \times \partial D^{n-p})$$
$$\cong (S^{p-1} \times D^{n-p}) \sqcup (D^p \times S^{n-p-1}).$$

An n-ball with this decomposition is called a p-handle. The subset of the boundary corresponding to $S^{p-1} \times D^{n-p}$ is called the attaching region.

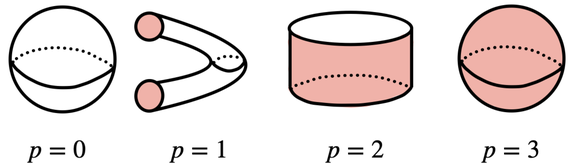


Figure 1: All four types of handles for a 3-ball with their attaching regions shaded in red.

Given a 3-manifold M and an embedding $f\colon S^{p-1} \times D^{3-p} \to M$, we can add a p-handle to M along f, giving $M \cup_{S^{p-1}\times D^{3-p}} D^3$, this is called p-handle addition.

Starting with an empty manifold, we can attach 0-handles (add disjoint 3-balls), then attach 1-handles, then attach 2-handles, and attach 3-handles to create a 3-manifold possibly with boundary. A manifold constructed from successive handle addition is said to have a handle decomposition.

A handlebody H_g of genus g is the 3-manifold obtained from attaching g 1-handles to one 0-handle. Note that ∂H_g is the surface of genus g.

Theorem (Smale): Every closed, oriented, smooth 3-manifold has a handle decomposition. Notice that thickening a triangulation of a manifold obtains a handle decomposition.

4. Lens Spaces

Lens spaces are a class of manifolds which can be easily described by their handle decomposition. Start with a 0-handle and attach a 1-handle to form a solid torus X. Given a simple closed curve γ in ∂X, attach a 2-handle to an annular neighborhood of γ on ∂X. The resulting manifold has boundary S^2, to which we attach a 3-handle forming a closed 3-manifold.

This process is determined only by the homotopy type of γ in ∂X. Any such γ can be written as $p\lambda + q\mu$ where λ is the longitude of X, μ is the meridian of X, and p and q are relatively prime integers. The lens space resulting from $\gamma = p\lambda + q\mu$ is called $L(p, q)$.

Example: $L(1,0) \cong S^3$, $L(0,1) \cong S^1 \times S^2$, and $L(2,1) \cong \mathbb{R}P^3$. Lens spaces $L(p_1, q_1)$ and $L(p_2, q_2)$ are homeomorphic if and only if $p_1 = p_2$ and $q_1 \equiv \pm q_2^{\pm 1} \bmod p_1$. They are homotopy equivalent if and only if $p_1 = p_2$ and $q_1 q_2 \equiv \pm n^2 \bmod p_1$ for some $n \in \mathbb{N}$. Lens spaces are a historic example showing that homotopy groups and homology groups are insufficient for classifying 3-manifolds.

4. Heegaard Decomposition

The 3-sphere can be decomposed into two closed 3-balls which intersect only on their boundary. This can be seen from the definition of $S^3 \subset \mathbb{R}^4$:

$$S^3 = \{(x, y, z, w) \in \mathbb{R}^4 \colon x^2 + y^2 + z^2 + w^2 = 1\}.$$

We can split this into two closed sets based on the values of w:

$$N = \{(x, y, z, w) \in S^3 \colon w \ge 0\},\ S = \{(x, y, z, w) \in S^3 \colon w \le 0\}$$

The sets N and S are homeomorphic to D^3. In their intersection, $w = 0$ so $E = N \cap S = \{(x, y, z, w) \colon x^2 + y^2 + z^2 = 1\} \cong S^2$.

The 3-sphere can also be split into two solid tori that overlap on their boundary. Consider $S^3 \subset \mathbb{C}^2$ as

$$S^3 = \{(z, w) \in \mathbb{C}^2 \colon |z|^2 + |w|^2 = 1\}.$$

Then

$$X = \{(z, w) \in S^3 \colon |z| \le 1/2\},\ Y = \{(z, w) \in S^3 \colon |z| \ge 1/2\}$$

gives two solid tori which overlap on their boundary which is homeomorphic to $S^1 \times S^1$.

A Heegaard splitting of a 3-manifold M is a decomposition of M along an embedded surface called the Heegaard surface into two handlebodies of neccesarily equal genus. A manifold with a Heegaard splitting can be recreated by gluing two disjoint handlebodies via an orientation-reversing homeomorphism of their boundaries.

The genus of a Heegaard splitting is defined to be the genus of the Heegaard surface. The Heegaard genus of a manifold is the minimum Heegaard genus of all Heegaard splittings of the manifold.

Theorem (Moise): Any 3-manifold can be given a Heegaard splitting using a handle decomposition. The union of the 0 and 1 handles form one handlebody and the union of the 2 and 3 handles form another handlebody.

5. Acknowledgements

Thank you to Rhea Palak Bakshi for mentoring me and my fellow students. Thank you to Rhea Palak again and to Melody Molander for organizing the UCSB Topology seminar. Thank you to the organizers of the Directed Reading Program.

6. References

1. A. T. Fomenko, S. V. Matveev, Algorithmic and Computer Methods for Three-Manifolds, Springer Dordrecht.
2. J. H. Przytycki, R. P. Bakshi, D. Ibarra, G. Montoya-Vega and D. E. Weeks, Lectures in Knot Theory: An Exploration of Contemporary Topics, Springer Universitext (Springer International Publishing, 2024).
3. Jennifer Schultens, Introduction to 3-Manifolds, AMS Graduate Studies in Mathematics.

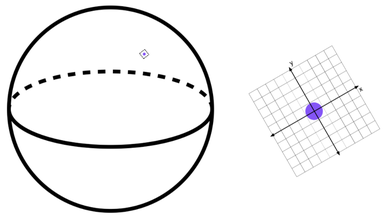
Data Analysis on Manifolds in Geomstats

Sneha Cheenath
University of California Santa Barbara

What is a manifold?

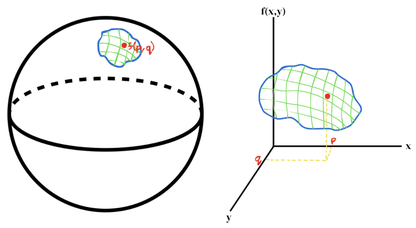
If any of the following three conditions are satisfied, a smooth surface can be considered a manifold:

1. **Local parametrization:** At any point on the manifold, an open set centered on that point can be mapped onto a cartesian plane. This is nicely illustrated by a globe: the overall surface is curved, but it appears flat locally.



2. **Implicit Function:** All the points on the manifold verify an implicit function. For example, the unit sphere in $\mathbb{R}^3$, which is a manifold, can be described as the set $\{(x,y,z)|x^2+y^2+z^2=1\}$

3. **Local graph:** Each region of the manifold can be described as a function $f(x_1,...,x_d)$. In other words, $x_{d+1}=f(x_1,...,x_d)$. This function does not have to be the same for the whole manifold, however, because manifolds do not have the same properties as functions – in particular, the same set of inputs can return several outputs.



Real-world data often falls on a manifold. If the shape of that manifold can be determined and quantified, then discerning trends and making predictions about data becomes much easier. For example, data about weather or location of cities falls on a spherical manifold, because the Earth is a globe.

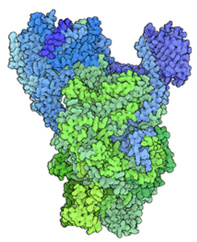


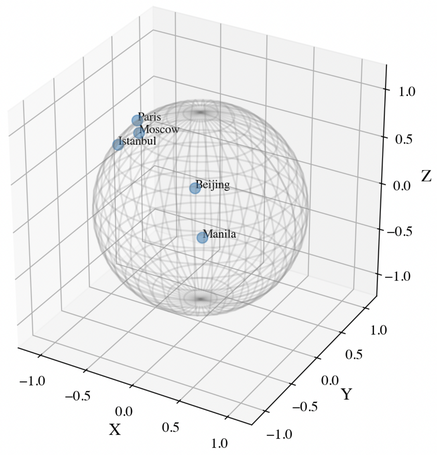
Figure 1: RCSB Protein Bank's molecule of the month for May, 2025 [3]

There are also bioinformatic uses for data analysis on manifolds [2]. For example, protein imaging can be used to analyze the relationship between the physical characteristics and the biological function of each protein. This is a huge bank of three-dimensional data, there are over 100,000 difference protein structures available online. Protein structures would generally fulfill the first definition of manifold because they are naturally occurring, so they should be smooth without any sharp corners.

A great motivation for the data analysis on manifolds is that it offers the ability to find a mean that lies on the surface. If the mean is determined to be the center of cartesian points, then it often will not lie on the surface, which is less useful. For example, to find the midpoint between two cities on the globe, it would not be helpful to find what location in the core of the Earth is between them.

Analyzing Manifolds in Geomstats

Geomstats is a Python library that provides tools to analyze data that lies on a manifold [1]. It can't help determine what manifold the data is on, but if the data's shape is already defined then it gives tools to analyze it. The parent class, *Manifold*, is abstract – because so many different surfaces can be considered manifolds, there are no universal functions to analyze them all. Instead, it just contains a skeleton for the attributes and methods consistent across the subclasses. For example, the attribute *dim* defines how many coordinates are necessary to fully describe a point on the manifold. It also has methods such as *belongs(), is_tangent(), random_point(),* and *random_tangent_vector()*.



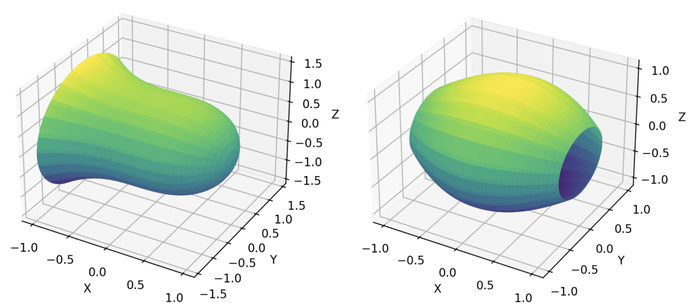
The Hypersphere subclass can be used with matplotlib to display where a city is on the globe, as shown to the right [1].

Some of the subclasses directly correspond to one of the conditions of being a manifold. The subclass *VectorSpaceOpenSet* corresponds with the first condition, as it is built for any manifold that can be described as an open set. The subclass *LevelSet* is to analyze functions where a function is set equal to a constant, which directly corresponds to the seconds condition.

The data in the manifold can be described intrinsically or extrinsically, and it can switch from one to the other. Extrinsic data is described with Euclidean coordinates, while intrinsic data is from the manifold's inherent topological space.

Class for Rotationally Symmetric Surfaces

A rotationally symmetric surface is one that is unchanged when rotated around an axis. Any curve f rotated around the x-axis is rotationally symmetric and has the parametrization $g(x,\theta)=(x,f(x)\cos(\theta),f(x)\sin(\theta))$. For example, below are visualizations of the functions $f(x)=\sqrt{1-x^3}$ and $f(x)=\cos(x)$ rotated around the x-axis on [-1,1].



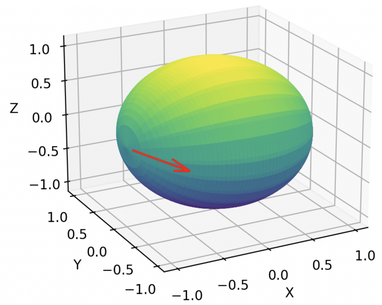
These figures are smooth surfaces that can be locally mapped to a Euclidean grid, so they are manifolds. Thus, we developed a new class to analyze three-dimensional rotationally symmetric surfaces. Using the properties of these surfaces, we were able to write methods that determine if a point belongs to a manifold, output a randomly generated point, output a randomly generated tangent vector at a point, and project any vector onto the tangent space of the surface. The class had two attributes: the function $f(x)$ being rotated and the domain [p,q].

Implementing the Methods

belongs(self, point): This function determines if a point (a,b,c) belongs to the manifold. It first confirms that a is in the given domain of the function. Then, since the parametric equation is $g(x,\theta)=(x,f(x)\cos(\theta),f(x)\sin(\theta))$, it is possible to check if (a,b,c) is on the curve using the following:

$$b^2+c^2=(f(x)\cos\theta)^2+(f(x)\sin\theta)^2=f(x)^2(\sin^2\theta+\cos^2\theta)=f(x)^2$$

Thus, the function returns true if $f(a)^2-(b^2+c^2)$ is zero.



random_point(self): This function randomly generates a point on the manifold. It creates temp, a random float value between [0,1], and maps it to the entire domain [p,q] by using the equation rand = (q-p)*temp+p. Then, it multiplies another random float across 2π to find a random θ, and plugs both those values into a parametric equation to return a final point. In the figure to the right, the function outputs point [-0.9903, -0.1379, 0.01873] to the rotation of $f(x)=\sqrt{1-x^2}$ around [-1,1].

random_tangent_vector(self, base_point): Similarly to random_point(), this function generates a random tangent vector at the base point (a,b,c). First, it ensures that the given point belongs on the surface – if it doesn't the function returns *None*. Then, it takes the derivative of the function and evaluates it at a. It can then solve for θ because $f(a)\cos\theta=b$, and $f(a)$, b are given by the base point. Thus, $\theta=\arccos(\frac{b}{f(a)})$. However, this will invoke an error if $f(a)=0$, so before then it's important to hard code that case. It is sufficient to make $\theta=0$ because if $f(a)=0$, then any terms containing θ will come out to 0 regardless of θ. Then, the function finds the partial derivative of $g(x)$ with respect to a and θ. These two vectors span the tangent vector space. Then, the function sums them with random coefficients and returns the product – thereby returning a random linear combination of the vectors in the basis. The above figure shows [8, -56.43, 7.525], a random vector tangent to the point outputted by *random_point(self)*.

to_tangent(self, base_point, vector): This function projects any vector onto the tangent space. Using the same method as *random_tangent_vector()*, this function finds θ. Then, it takes the partial derivative of $g(x)$ with respect to a and θ. It then projects the given vector onto each of these partial derivatives and returns the pairwise sum of the two projections.

Acknowledgements

I would like to thank Jihye Lee for her continued support, encouragement, and guidance. I would also like to thank Adele Myers for her help with Geomstats documentation. And finally, thank you to the UC Santa Barbara Directed Reading Program for the opportunity to work on this project.

References

[1] Adele Myers Anna Calissano Nina Miolane. *Geomstats latest documentation*. `https://geomstats.github.io/notebooks/01_foundations__manifolds.html`. Accessed: 2025-05-12.

[2] Victor Patrangenaru and Leif Ellingson. *Nonparametric statistics on manifolds and their applications to object data analysis*. CRC Press, Taylor & Francis Group Boca Raton, 2016.

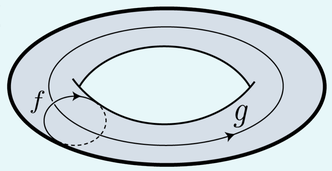
[3] *Research Collaboratory for Structural Bioinformatics Protein Data Bank*. `https://www.rcsb.org/`. Accessed: 2025-05-13.

A Classical formulation of the Seifert-Van Kampen Theorem

Max Welter
University of California Santa Barbara

Paths and Homotopies

Let X be a topological space, which moving forward we will simply call a space for brevity. We define a **path** from x_0 to x_1 in X to be a continuous map from the unit interval I to X, such that $f(0) = x_0$ and $f(1) = x_1$. If, perchance, $x_0 = x_1$, then we call this a **loop** based at x_0. We call a space X **path-connected** if there exists a path between every pair of points in X. In the event we are given two paths f and g in X who miraculously satisfy the equation $f(1) = g(0)$ we define the **product** of paths $f \cdot g$, as



Loops f and g on the torus

$$f \cdot g(s) = \begin{cases} f(2s) & 0 \leqslant s \leqslant 1/2 \\ g(2s-1) & 1/2 \leqslant s \leqslant 1. \end{cases}$$

We call this the **product path** of f and g. Intuitively, the product path is the path created by "adjoining" the paths f and g.

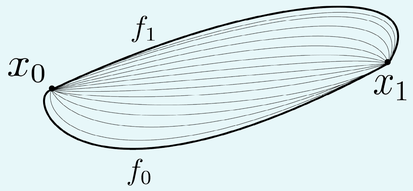


Figure 1: A homotopy from f_0 to f_1

A **homotopy of paths** is a family $f_t : I \to X$, for $0 \leqslant t \leqslant 1$, of paths from x_0 to x_1 such that the function $F : I \times I \to X$ defined by $F(s,t) = f_t(s)$ is continuous. If two paths f_0 and f_1 are connected by such a family we call f and g **homotopic**, and denote this by $f_0 \simeq f_1$. Being homotopic is an equivalence relation on paths and we call the associated equivalence class of any path f the **homotopy class** of f, denoted $[f]$. Using the product of paths we let $[f][g] := [f \cdot g]$ be the product of homotopy classes [1].

Fundamental Groups

The **fundamental group** of X based at x_0, denoted $\pi_1(X, x_0)$, is the set of all homotopy classes of loops in X with base point x_0, equipped with the product of homotopy classes as its operation. Algebraic topologists are often tasked with distinguishing spaces from each other, and as such, keeping track of the invariants of spaces is important. This gives us one of the main reasons we even care about $\pi(X, x_0)$, namely because it is a topological invariant of X! Beyond that, $\pi(X, x_0)$ allows us to use algebraic tools to help shed light on various topological properties of spaces, giving us even greater insights of problems originating in topology.

Change of Basepoint:
If a space X is path-connected we can transform any loop based at x_0 into a loop with a base point x_1, for any x_1 in X. Using this one can show that $\pi_1(X, x_0) \cong \pi_1(X, x_1)$. This means that for X path-connected, we usually omit a base point from the notation and write $\pi_1(X)$. We also say a space is **simply connected** if it is path-connected and has a trivial fundamental group.[1].

Induced Homomorphisms:
Let X and Y be spaces and take $\varphi : X \to Y$ to be a continuous map taking a base point $x_0 \in X$ to $y_0 \in Y$. Then we call the map

$$\begin{aligned} \varphi_* : \pi_1(X, x_0) &\to \pi_1(Y, y_0) \\ [f] &\longmapsto [\varphi f] \end{aligned}$$

the **homomorphism induced** by φ [1].

A Few Words on Free Products of Groups

Let X and Y be spaces with base points x_0 and y_0, respectively, and set $G_1 := \pi_1(X, x_0)$ and $G_2 := \pi_1(Y, y_0)$. We define a **word** to be the empty tuple or any n-tuples whose coordinates are non-trivial loops in either G_1 or G_2. A word $(g_1, \ldots, g_n)$ is said to be **reduced**, if it is the empty word, which is the empty tuple, or if no two adjacent loops g_{j-1} and g_j lie in the same G_i. Define the **concatenation** of the words $w = (a_1, \ldots, a_n)$ and $w' = (b_1, \ldots, b_n)$ by

$$(w, w') \mapsto (a_1, \ldots, a_n, b_1, \ldots, b_n)$$

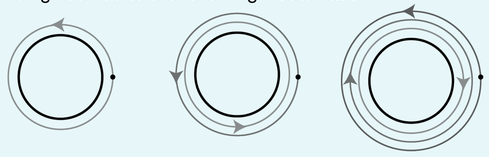
Given a word $w = (a_1, \ldots, a_n)$ and $a_{j-1}, a_j \in G_i$, we put

$$w_1 = \begin{cases} (a_1, \ldots, a_{j-1}a_j, \ldots a_n) & \text{if } a_{j-1}a_j \neq 1 \\ (a_1, \ldots, a_{j-2}, a_{j+1}, a_n) & \text{if } a_{j-1}a_j = 1 \end{cases}$$

We call $w \to w_1$ an **elementary reduction** and a sequence $w \to w_1 \to \cdots \to w_r$ a **reduction** if w_r is reduced. The **free product** of G_1 and G_2, written $G_1 * G_2$, is the set of all reduced words endowed with concatenation followed by reduction as an operation. If, in addition, $G_1 = \langle a \rangle$ and $G_2 = \langle b \rangle$, then we call $G_1 * G_2$ a **free group** with system of free generators a and b [2][3].

Fundamental Group of the Circle

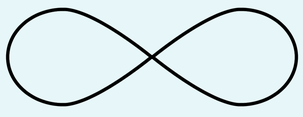
To help elucidate the ideas behind both fundamental and free groups, we will describe the fundamental group of the unit circle, which we will denote S^1. Since S^1 is path-connected, we really only need to consider the fundamental group of S^1 at a *single* arbitrary point x_0. For the rest of the section all loops discussed will be assumed to have the base point x_0. Begin by noticing that any loop in with S^1 is homotopic to a loop that winds around the circle n times, for some $n \in \mathbb{Z}$, which gives rise to the following visualization:



We can therefore create a function that assigns every loop a unique integer corresponding to it, called its winding number. Observe that the product of loops f and g, with winding numbers n and m respectively, is homotopic to a loop with winding number $n + m$. By this reasoning the product of f with the constant loop must be homotopic to f and the product of f with a loop with winding number $-n$ is homotopic to the constant loop. Thus, we see that $\pi_1(S^1)$ is isomorphic to the additive group $\mathbb{Z}$ and the free group on one generator a.

The Motivation for Seifert-van Kampen

Suppose you were tasked with computing the fundamental group of the following space:



You might notice that this infinity loop, which we will denote as R, is topologically equivalent to taking two copies of S^1 adjoined at a single point. However, instead of naively duplicating your calculation of $\pi_1(S^1)$ you remember that you are an algebraic topologist (in case you forgot). Accordingly, you ruminate on possible generalizations of the problem at hand.

Pondering Time:
You think to yourself: "Spaces are often, in a sense, decomposable into smaller, well-studied, *simpler* spaces, as is the case here. In these cases, it would be incredibly convenient if I could use our knowledge about the fundamental groups of the simpler spaces to help us compute the fundamental group of the larger space at hand." Well-done! You have found your way to the essence of the Seifert–van Kampen theorem.

The Seifert-van Kampen Theorem

Theorem 1 (Classical Formulation)**.** *Let $X = U \cup V$, where U and V are open in X; assume U, V, and $U \cap V$ are path connected; let $x_0 \in U \cap V$. Set $i_1 : \pi_1(U \cap V, x_0) \to \pi_1(U, x_0), i_2 : \pi_1(U \cap V, x_0) \to \pi_1(V, x_0), j_1 : \pi_1(U, x_0) \to \pi_1(X, x_0), j_2 : \pi_1(V, x_0) \to \pi_1(X, x_0)$ to be the homomorphisms induced by inclusion. Let*

$$\varphi : \pi_1(U, x_0) * \pi_1(V, x_0) \longrightarrow \pi_1(X, x_0)$$

be the homomorphism defined by the natural extension of j_1 and j_2. Then φ is surjective, and its kernel is the least normal subgroup N of the free product that contains all elements represented by words of the form

$$i_1(g)^{-1} i_2(g)$$

for $g \in \pi_1(U \cap V, x_0)$[2].

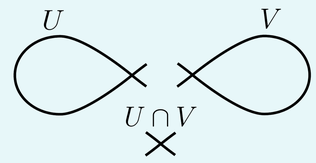
This theorem tells us that if two path-connected spaces with path-connected intersection, U and V form a space X, then together $\pi_1(U, x_0)$ and $\pi_1(V, x_0)$ can recover important algebraic data of $\pi_1(X, x_0)$.

Corollary 1. *Assume the hypotheses of the Seifert-Van Kampen theorem. If $U \cap V$ is simply connected, then*

$$\pi_1(U, x_0) * \pi(V, x_0) \cong \pi(X, x_0).$$

We can think about this corollary as saying that since $\pi_1(U \cap V)$ is trivial it is, in a sense, negligible. For this reason, we only have to think about how the elements of $\pi_1(U, x_0)$ and $\pi_1(V, x_0)$ combine in order to completely characterize $\pi_1(X, x_0)$.

Applications:
Let us return to R. Take open balls A and B so that they contain the ellipse on the left and right, respectively. Define U and V to be the intersection of A and B with their respective ellipses. Then we have



By construction, U and V are open, path connected, non-empty, and $U \cap V$ is simply connected. Additionally, we can "squish down" U and V to be S^1. Therefore, if we apply corollary 1, we deduce that $\pi_1(R)$ is isomorphic to $\mathbb{Z} * \mathbb{Z}$ and the free group on two generators.

Acknowledgements

I would like to express my most profound gratitude to Mary Zimmerman. Her patient guidance during challenging periods and responsive support during productive phases were invaluable to the completion of this work and my growing love of Algebraic Topology. I am beyond grateful for the opportunity to study under her tutelage.

References

[1] A. Hatcher. *Algebraic Topology*. Algebraic Topology. Cambridge University Press, 2002. ISBN: 9780521795401. URL: https://books.google.com/books?id=BjKs86kosqgC.

[2] J. Munkres. *Topology*. Pearson Modern Classics for Advanced Mathematics Series. Pearson, 2017. ISBN: 9780134689517. URL: https://books.google.com/books?id=51n8MAAACAAJ.

[3] J.J. Rotman. *Advanced Modern Algebra*. Graduate Studies in Mathematics. American Mathematical Society, 2015. ISBN: 9781470415549. URL: https://books.google.com/books?id=SugUCwAAQBAJ.

How to Distinguish the Left-Handed and Right-Handed Trefoil Knots

Sam Wang Runxin Shen

University of California Santa Barbara

Brief Introduction

In the film " 君の名は。(Your Name)," released in 2016, the protagonist's name shares the same character in Japanese as the trefoil knot. The movie often uses "knot" as a symbol, echoing the "connection between people" in Japanese culture. Here, we explore several methods to distinguish the left- and right-handed trefoil knots.

Figure 1. character name: 三葉 (Mitsuha) trefoil knot in Japanese: 三葉結び目

Definition of Knot and link

A **knot** is an embedding of the circle S^1 into three-dimensional Euclidean space $\mathbb{R}^3$ (or the 3-sphere S^3):

$$K : S^1 \to \mathbb{R}^3.$$

Two knots are considered equivalent if they are **ambient isotopic**, meaning there exists a continuous deformation of $\mathbb{R}^3$ taking one embedding to the other without cutting or self-intersections.

A **link** is an embedding of a disjoint union of finitely many circles into $\mathbb{R}^3$:

$$L : \bigsqcup_{i=1}^{n} S^1 \to \mathbb{R}^3,$$

where each S^1 is mapped to a smooth, simple closed curve, and the images are pairwise disjoint.

What is a Trefoil Knot?

The trefoil knot is the simplest **nontrivial knot**. Unlike the unknot, which is a simple loop, the trefoil knot can not be untangled without cutting the loop. There are two trefoil knots up to isotopy: the left-handed trefoil and its mirror image, the right-handed trefoil.

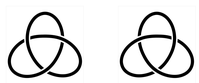


Figure 2. Left-handed and Right-handed Trefoil Knots

These two versions look similar but are fundamentally different. You can not twist or deform one into the other without cutting or passing through itself. That is because they are not equivalent under ambient isotopy.

Reidemeister Moves

A knot is called **amphicheiral** if it is ambient-isotopic to its mirror image—equivalently, if a finite sequence of Reidemeister moves turns the diagram into its mirror. There are exactly three kinds of moves:

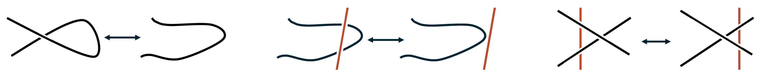


Figure 3. Reidemeister moves I, II and III.

The trefoil knot is not amphicheiral, but there is no known proof in terms of Reidemeister moves. An example of an amphicheiral knot is the figure-eight.

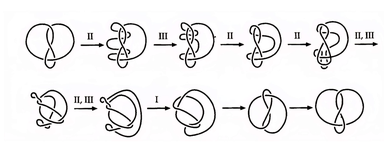


Figure 4. The figure-eight knot is equivalent to its mirror image.

Tricolorability

A projection of a knot or link is tricolorable if each strand in the projection can be colored one of three different colors, so that at each crossing, either three different colors come together or all the same color comes together.

Figure 5. Tricolorability of Left-handed and Right-handed Trefoil Knots

Both the left- and right-handed trefoil knots are tricolorable, and any tricolorable knot is necessarily distinct from the unknot. However, tricolorability alone does not distinguish between the left- and right-handed trefoils.

Dowker Notation

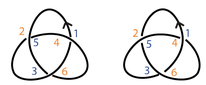


Figure 6. Left-handed and Right-handed Trefoil Knots

Result: they share the same Dowker Notation, 4 6 2. Thus we cannot distinguish between the two trefoils by Dowker Notation.

The Very First Polynomial Link Invariant

Introduced by J. W. Alexander in 1928, it is usually presented as a Laurent polynomial in t:

$$\Delta : \{\text{oriented link diagrams}\} \to \mathbb{Z}[t, t^{-1}]$$

Before the calculation of the Alexander polynomial, we better need to clarify a definition:

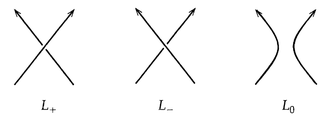


Figure 7. Three links that are identical except at this crossing

It can be calculated by two rules, which was shown by John Conway:

$$\text{Rule 1:}\quad \Delta\left(\bigcirc\right) = 1$$

$$\text{Rule 2:}\quad \Delta(L_+) - \Delta(L_-) + \left(t^{1/2} - t^{-1/2}\right)\Delta(L_0) = 0$$

We can **treat the link as L_+ or L_- based on one of its crossings**, and then simplify it step by step to arrive at the final answer.

Let's work on the Alexander polynomial of left-handed and right-handed trefoils separately.

Left-handed

$\Delta\left(\text{LH}\right) - \Delta\left(\text{LH}\right) + (t^{1/2} - t^{-1/2})\Delta\left(\text{LH}\right) = 0$

where $\Delta\left(\text{LH}\right) = \Delta\left(\bigcirc\right) = 1$

and $\Delta\left(\text{LH}\right) - \Delta\left(\text{LH}\right) + (t^{1/2} - t^{-1/2})\Delta\left(\text{LH}\right) = 0$

By the theorem that the Alexander polynomial of a splittable link is always 0, $\Delta\left(\text{LH}\right) = 0$, so $\Delta\left(\text{LH}\right) = t^{1/2} - t^{-1/2}$ and $\Delta\left(\text{LH}\right) = (t^{1/2} - t^{-1/2})^2 + 1 = t - 1 + t^{-1}$

Right-handed

$\Delta\left(\text{RH}\right) - \Delta\left(\text{RH}\right) + (t^{1/2} - t^{-1/2})\Delta\left(\text{RH}\right) = 0$

where $\Delta\left(\text{RH}\right) = \Delta\left(\bigcirc\right) = 1$

and $\Delta\left(\text{RH}\right) - \Delta\left(\text{RH}\right) + (t^{1/2} - t^{-1/2})\Delta\left(\text{RH}\right) = 0$

Again, by the theorem that the Alexander polynomial of a splittable link is always 0, $\Delta\left(\text{RH}\right) = 0$, so $\Delta\left(\text{RH}\right) = -t^{1/2} + t^{-1/2}$ and $\Delta\left(\text{RH}\right) = (t^{1/2} - t^{-1/2})^2 + 1 = t - 1 + t^{-1}$

Result: we cannot distinguish between the two trefoils by the Alexander polynomial.

References

[1] Colin Conrad Adams. *The Knot Book: An Elementary Introduction to the Mathematical Theory of Knots.* Providence, R.I., American Mathematical Society, 2010.

Kauffman Bracket Polynomial

The Kauffman bracket polynomial is defined for **unoriented** knots and links, and provides the foundation for constructing the Jones polynomial.

Rule 1

$$\langle \bigcirc \rangle = 1$$

Rule 2

$$\left\langle \times \right\rangle = A \left\langle \;)(\right\rangle + A^{-1} \left\langle \asymp \right\rangle$$
$$\left\langle \times \right\rangle = A \left\langle \asymp \right\rangle + A^{-1} \left\langle \;)(\right\rangle$$

Rule 3

$$\langle L \cup \bigcirc \rangle = (-A^2 - A^{-2})\langle L \rangle$$

Right-handed Trefoil

$$\begin{aligned}
\left\langle \text{RH} \right\rangle &= A \left\langle \text{RH} \right\rangle + A^{-1} \left\langle \text{RH} \right\rangle \\
&= A \left\langle \bigcirc\bigcirc \right\rangle + A^{-1} \left(A \left\langle \text{RH} \right\rangle + A^{-1} \left\langle \text{RH} \right\rangle \right) \\
&= A(-A^4 - A^{-4}) + A^{-1} \left(A(-A^{-3}) + A^{-1}(-A^2 - A^{-2})(-A^{-3}) \right) \\
&= A^{-7} - A^5 - A^{-3}
\end{aligned}$$

Left-handed Trefoil

$$\begin{aligned}
\left\langle \text{LH} \right\rangle &= A^{-1} \left\langle \text{LH} \right\rangle + A \left\langle \text{LH} \right\rangle \\
&= A^{-1} \left\langle \bigcirc\bigcirc \right\rangle + A \left(A^{-1} \left\langle \text{LH} \right\rangle + A \left\langle \text{LH} \right\rangle \right) \\
&= A^{-1}(-A^4 - A^{-4}) + A \left(A^{-1}(-A^3) + A(-A^2 - A^{-2})(-A^3) \right) \\
&= A^7 - A^3 - A^{-5}
\end{aligned}$$

Result: The Kauffman bracket distinguishes the two trefoils.

The X Polynomial and Writhe

Introduce an **orientation** on a knot or link projection. The *writhe* $w(L)$ is defined as the difference between the number of positive and negative crossings in the projection.

We define a new polynomial called the X-polynomial. It is a polynomial of oriented links and is defined as:

$$X(L) = (-A^3)^{-w(L)}\langle L \rangle$$

where $w(L)$ is the writhe of the oriented link diagram L, and $\langle L \rangle$ is the Kauffman bracket.

Now suppose we perform a Type I Reidemeister move that introduces a positive twist, so that:

$$w(L') = w(L) + 1$$

$$X(L') = (-A^3)^{-(w(L)+1)}\langle L' \rangle = (-A^3)^{-w(L)-1}(-A^3\langle L \rangle) = (-A^3)^{-w(L)}\langle L \rangle = X(L)$$

The Jones Polynomial

The Jones polynomial is a Laurent polynomial in the variable q, assigned to an oriented knot or link such that:

$$V_L : \{\text{oriented link diagrams}\} \to \mathbb{Z}[q, q^{-1}]$$

We can calculated it in **two different ways**

Given two rules,

$$\text{Rule 1:}\quad V\left(\bigcirc\right) = 1$$

$$\text{Rule 2:}\quad q^{-1}V(L_+) - qV(L_-) + (q^{-1/2} - q^{1/2})V(L_0) = 0$$

Or, it can be directly obtained ***from the X polynomial by replacing each A by*** $\boldsymbol{q^{-1/4}}$

So, left-handed: $-A^{16} + A^{12} + A^4 \to -(q^{-1/4})^{16} + (q^{-1/4})^{12} + (q^{-1/4})^4 \to q^{-1} + q^{-3} - q^{-4}$

So, right-handed: $-A^{-16} + A^{-12} + A^{-4} \to -(q^{-1/4})^{-16} + (q^{-1/4})^{-12} + (q^{-1/4})^{-4} \to q + q^3 - q^4$

Result: Finally, we can distinguish between the two trefoils by the Jones Polynomial, and it was the first invariant polynomial that distinguishes between a knot and its mirror image.

Acknowledgements

Thanks to our mentor Rhea Palak Bakshi. She opened the doors of topology for us, thanks for all her dedication. Thanks to the other seniors in our group who helped us a lot. And thanks to DRP for this opportunity.

The Kauffman Bracket Skein Module

Miranda Jiang [1] Mentored by Rhea Palak Bakshi [1]

[1]University of California - Santa Barbara
Department of Mathematics - Directed Reading Program 2025

Introduction

We often use polynomial invariants to detect if two knots are different. In fact, this familiar method could be elevated to 3-manifolds that a link is sitting in.
The **Kauffman bracket skein module** is a powerful invariant that is able to detect both the links in them, and the 3-manifolds themselves.

Background

Let M be a 3-manifold with boundary ∂M. A p-**handle** is a 3-disk with a specific product structure $D^3 = D^p \times D^{3-p}$. Its boundary is $\partial D^3 = (\partial D^p \times D^{3-p}) \cup (D^p \times \partial D^{3-p})$.
A p-**handle addition** is a modification of M by gluing a p-handle $D^3 = D^p \times D^{3-p}$ to ∂M along the boundary of the p-handle.

0-handle: adding a disjoint 3-ball, since $\partial D^0 \times D^3 = \varnothing$.
1-handle: gluing a solid cylinder $D^1 \times D^2$ along $S^0 \times D^2$, that is, along its two boundary disks.
2-handle: gluing a "plate" $D^2 \times D^1$ along the annular boundary $S^1 \times D^1$.
3-handle: gluing a 3-ball along $S^2 \times D^0$, often visualized as capping off a hole.

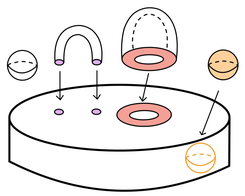


Figure 1. p-handle additions to M for $p = 0, 1, 2, 3$

The 2- and 3- handle additions are essential to our study of 3-manifolds because any compact oriented 3- manifold can be obtained from a handlebody by adding 2- and 3-handles to it, hence we can compute the skein module by computing the skein module of the handlebody first, then quotient by the submodule generated by 2-handle sliding relations. Note that 3-handle additions induce isomorphisms.

Kauffman Bracket Skein Module (KBSM)

The **Kauffman bracket skein module**, $S_{2,\infty}(M)$, of a 3-manifold M is defined as the quotient of the free R-module spanned by ambient isotopy classes of unoriented framed links (including the empty link $\varnothing$) in M modulo the following (local) skein expressions:

(i) $L_+ - AL_0 - A^{-1}L_\infty$,
(ii) $L \sqcup \bigcirc + (A^2 + A^{-2})L$,

where A is a fixed invertible element of the free-R-module, $\bigcirc$ denotes the trivial framed knot and the *skein triple* (L_+, L_0, L_∞) denotes three framed links in M, which are identical except in a small 3-ball in M where they look like the following:

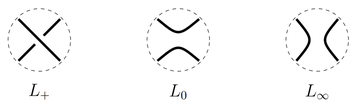
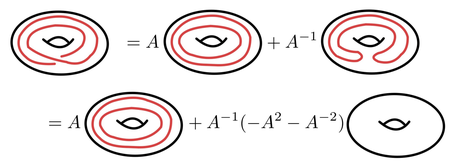


Figure 2. Skein Triple (L_+, L_0, L_∞)

The **Kauffman Bracket Skein Module (KBSM)** is the quotient

$$S_{2,\infty}(M; R, A) = R\mathcal{L}^{fr}/S^{sub}_{2,\infty}.$$

Example: An Element of KBSM of *Ann*$\times[0,1]$



$$= Ax^2 + (-A - A^{-3}) \cdot \varnothing$$

Figure 3. Element of KBSM corresponding to a nontrivial link

Properties of KBSM

1. Functoriality An orientation-preserving embedding of 3-manifolds $i : M \hookrightarrow N$ yields a homomorphism of skein modules

$$i_* : S_{2,\infty}(M; R, A) \longrightarrow S_{2,\infty}(N; R, A)$$

This gives a functor from the category of 3-manifolds and orientation-preserving embeddings (up to ambient isotopy) to the category of R-modules with a specified invertible element $A \in R$.

2. 3-handle addition If N is obtained from M by adding a 3-handle to M (i.e., capping off a hole), and $i : M \hookrightarrow N$ is the associated embedding, then

$$i_* : S_{2,\infty}(M; R, A) \longrightarrow S_{2,\infty}(N; R, A)$$

is an isomorphism.

3. 2-handle addition Let N be the 3-manifold obtained from M by adding a 2-handle to M along a simple closed curve γ in ∂M, then the inclusion $i : M \hookrightarrow N$ induces an epimorphism of the skein modules:

$$i_* : S_{2,\infty}(M; R, A) \longrightarrow S_{2,\infty}(N; R, A),$$

whose kernel is generated by relations yielded by 2-handle slidings. That is,

$$S_{2,\infty}(N; R, A) = S_{2,\infty}(M; R, A)/\mathcal{J}.$$

where $\mathcal{J}$ is generated by relations of the form $L - sL_\gamma(L)$

4. Disjoint union If $M_1 \sqcup M_2$ is the disjoint sum of 3-manifolds M_1 and M_2, then

$$S_{2,\infty}(M_1 \sqcup M_2; R, A) = S_{2,\infty}(M_1; R, A) \otimes_R S_{2,\infty}(M_2; R, A).$$

5. The Universal Coefficient Property Let $r : R \to R'$ be a ring homomorphism between commutative rings with unity R and R'. Then the identity map on $\mathcal{L}_{fr}$ induces an isomorphism between the R- and R'-modules:

$$\bar{r} : S_{2,\infty}(M; R, A) \otimes_R R' \longrightarrow S_{2,\infty}(M; R', r(A)).$$

6. Connected Sums Let $M \# N$ denote the connected sum of two compact, oriented 3-manifolds M and N, and let $A^k - 1$ be invertible in R for any $k > 0$. Then,

$$S_{2,\infty}(M \# N; R, A) = S_{2,\infty}(M; R, A) \otimes S_{2,\infty}(N; R, A).$$

Note: This result does not hold for $\mathbb{Z}[A^{\pm 1}]$.

Acknowledgements

Many many many thanks to my mentor, Rhea Palak Bakshi, for her continuous support both inside and outside of the program itself. Math is hard, but Rhea's enthusiasm to the subject and her patience to my stupidities really made the experience of learning topology joyful.
Many many thanks to my group mates Cole, Alex, Andrew, Sam and Runxin! It's the sharing of knowledge made learning extra enjoyable.
Many many thanks to the DRP committee for making all of these possible!

Examples: KBSM of Different 3-Manifolds

The study of skein modules of 3-manifolds involves both the algebraic structure of the module, and the 3-manifold itself under various operations. We here introduce a few of the key theorems and examples:

KBSM of Surface I-bundles

Theorem *(Przytycki): The Kauffman bracket skein module of $\Sigma \times I$ is freely generated by the empty link and links in Σ without crossings and trivial components.*
This result in particular applies to a ***handlebody****, because $H_n = \Sigma_{0,n+1} \times I$.*

Examples: Thickened Surfaces.

Let $\Sigma_{g,n}$ be an oriented surface with genus g and n boundary components.

1. Thickened Annulus
$S_{2,\infty}(S^1 \times D^2; R, A) \cong S_{2,\infty}(\Sigma_{0,2} \times I; R, A)$ is free and infinitely generated by the curves $\{x^i\}_{i=0}^{\infty}$, where x denotes the homotopically nontrivial curve on the annulus and x^0 denotes the empty link $\varnothing$.

2. Thickened Torus
$S_{2,\infty}(T^2 \times I; R, A)$ is a free R-module generated by the empty link $\varnothing$, all (p, q)-curves, and their parallel copies on the torus. These are simple closed curves that wrap along the torus p times in the longitudinal direction and q times in the meridional direction. Here $\gcd(p, q) = 1$.

3. Thickened Pair of Pants
$S_{2,\infty}(\Sigma_{0,3} \times I; R, A)$ is free and infinitely generated by the monomials $\{x^i y^j z^k\}_{i,j,k \geq 0}$. Here, x, y, and z denote the homotopically nontrivial curves in $\Sigma_{0,3}$. Note that the empty link is represented by $x^0 y^0 z^0$. Additionally,

$$S_{2,\infty}(\Sigma_{1,1} \times I; R, A) \cong S_{2,\infty}(\Sigma_{0,3} \times I; R, A).$$

KBSM of Lens Spaces

Theorem *(Przytycki): $S_{2,\infty}(L(p, q))$ is a free $\mathbb{Z}[A^{\pm 1}]$-module and it has $\lfloor \frac{p}{2} \rfloor + 1$ free generators.*

KBSM of $S^1 \times S^2$

Theorem *(Przytycki): $S_{2,\infty}(S^1 \times S^2)$ is an infinitely generated $\mathbb{Z}[A^{\pm 1}]$-module. More precisely,*

$$S_{2,\infty}(S^1 \times S^2) = \mathbb{Z}[A^{\pm 1}] \oplus \bigoplus_{i=1}^{\infty} \frac{\mathbb{Z}[A^{\pm 1}]}{1 - A^{2i+4}}.$$

KBSM of Prime 3-Manifolds

Conjecture*: The KBSM of a closed oriented prime 3-manifold has a decomposition into free and cyclic modules, just like $S^1 \times S^2$.*

Connections

Skein modules is a rich area of research that connects to many fields of mathematics and physics. We can use the theory of skein modules to build 3-manifold invariants in which distinguish two different 3-manifolds. Many of the 3-manifold invariants derived from KBSM are widely used in the study of quantum topology and especially the building of Topological Quantum Field Theories (TQFT's). The Witten–Reshetikhin–Turaev (WRT) invariant is one of the most studied invariants derived from KBSM. KBSM is also connected to algebraic geometry via the $SL_2(\mathbb{C})$ character variety.

References

[1] Józef H. Przytycki, Rhea Palak Bakshi, Dionne Ibarra, Gabriel Montoya-Vega, and Deborah Weeks. *Lectures in Knot Theory: An Exploration of Contemporary Topics*. Springer Universitext, 2024.

[2] Józef H. Przytycki. Fundamentals of kauffman bracket skein modules, 1998.

[3] Józef H. Przytycki. Kauffman bracket skein module of a connected sum of 3-manifolds, 1999.

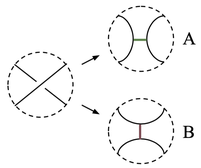
Knot Your Average Homology

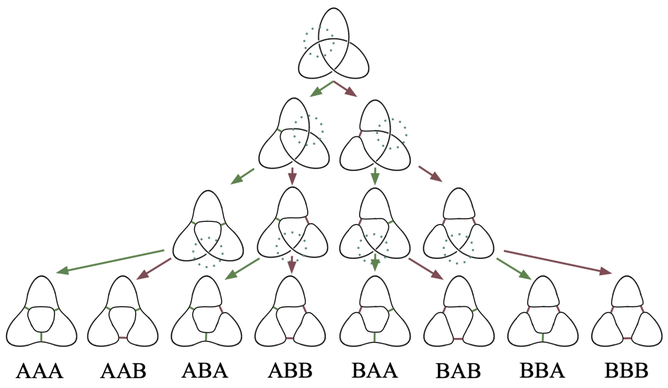
Alex Gaither
University of California, Santa Barbara

The Definition of Khovanov Homology

Khovanov homology is a powerful algebraic link invariant that categorifies the Jones Polynomial.
Theorem (Kronheimer & Mrowka): Khovanov homology detects the unknot.

A **Kauffman state** (KS), denoted s, is a particular sequence of smoothings (**A** or **B**), one at each crossing, on a link. An **enhanced Kauffman state** (EKS), denoted S, assigns a positive or negative sign to each cycle in a KS. Below are all the KS for the right-handed trefoil. There are $2^3 = 8$ since the knot has 3 crossings.





Definition of the Khovanov Chain Complex:

- Choose an ordering of the crossings on the link. What choice is irrelevant.
- $\sigma(s)$: # A markers minus # B markers in a KS
- $\tau(S)$: # positive assigned cycles minus # negative assigned cycles in an EKS
- **Bidegree** on the EKS of a link diagram D:

$$S_{a,b}(D) = \{S \in EKS | a = \sigma(s), b = \sigma(s) + 2\tau(S)\}$$

- **Chain groups:** $C_{a,b} = \mathbb{Z}S_{a,b}$
- **Boundary map:** $\partial_{a,b} : C_{a,b} \to C_{a-2,b}$ by

$$\partial_{a,b}(S) = \sum_{S' \in S_{a-2,b}} (-1)^{t(S,S')}(S, S')S'$$

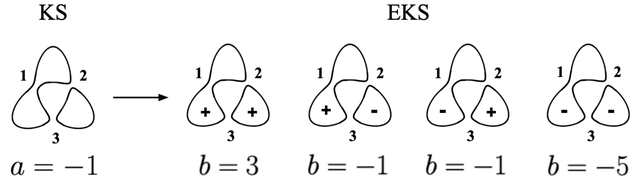
where S' is an EKS in $C_{a-2,b}$. We must and do have $\partial_{a,b} \circ \partial_{a+2,b} = 0$ as ∂ is a boundary map. $(S, S') = 1$ if the two conditions below hold. Else, $(S, S') = 0$:

1. S and S' only differ at one crossing v, where in S it is smoothed with an A marker and in S' it is smoothed with a B marker
2. $\tau(S') = \tau(S) + 1$ and only cycles involving v change their sign

- $t(S, S')$: # B markers in S after the crossing v in our ordering
- **Khovanov homology groups:** $H_{a,b} = \frac{ker(\partial_{a,b})}{im(\partial_{a+2,b})}$. These are well defined because of the above requirement for the boundary map.

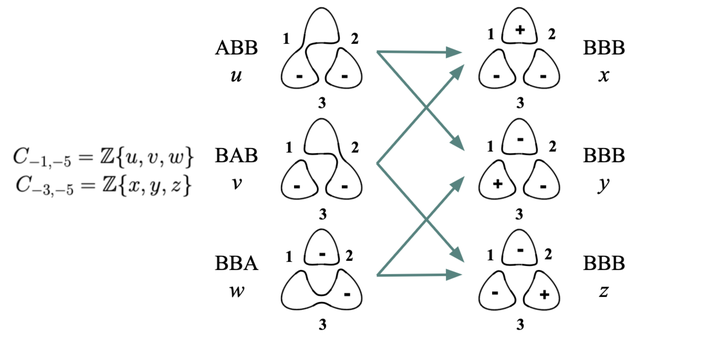
An Example: The Right-Handed Trefoil

Below are all of the EKS for the ABB KS for the right-handed trefoil. Every EKS is assigned a bigrading a, b. Notice they all have the same a, which is inherited from the ABB KS. There are $2^2 = 4$ states since there are 2 cycles.



Examine below the chain complex for $b = -5$ in the right-handed trefoil, which contains the groups of the form $C_{a,-5}$. The EKS below each chain group are the chain group's generators. The green arrows depict each generator's image under the boundary map, with each arrow representing two adjacent EKS ($(S, S') = 1$).

$$\ldots \xrightarrow{\partial_{3,-5}} 0 \xrightarrow{\partial_{1,-5}} C_{-1,-5} \xrightarrow{\partial_{-1,-5}} C_{-3,-5} \xrightarrow{\partial_{-3,-5}} 0$$



Khovanov homology of the Chiral Trefoils:

Left-Handed

b\a	−3	−1	1	3
9				$\mathbb{Z}$
5			$\mathbb{Z}_2$	
1			$\mathbb{Z}$	
−3	$\mathbb{Z}$			
−7	$\mathbb{Z}$			

Right-Handed

b\a	−3	−1	1	3
7				$\mathbb{Z}$
3				$\mathbb{Z}$
−1		$\mathbb{Z}$		
−5	$\mathbb{Z}_2$			
−9	$\mathbb{Z}$			

Notice that Khovanov homology distinguishes the chiral trefoils!

We calculate $H_{-3,-5}$ for the right-handed trefoil from the figure above:

$$H_{-3,-5} = \frac{ker(\partial_{-3,-5})}{im(\partial_{-1,-5})}$$
$$\partial_{-1,-5}(u) = x + y$$
$$\partial_{-1,-5}(v) = -x - z$$
$$\partial_{-1,-5}(w) = y + z$$
$$ker(\partial_{-3,-5}) \cong \mathbb{Z}\{x, y, z\}$$

$$\implies \mathrm{H}_{-3,-5} \cong \frac{\mathbb{Z}\{x,y,z\}}{\mathbb{Z}\{x+y,-x-z,y+z\}} \cong \mathbb{Z}_2 \text{ as } (x+y)+(-x-z)+(y+z) = 2y \equiv 0$$

Framed vs. Unframed, Euler Characteristic

The version of Khovanov homology presented in this poster uses Oleg Viro's notation and is for framed links. We may translate into the unframed version through the following change of variables:

$$H^{i,j}(\vec{D}) = H_{a,b}(D) = H^{\frac{w(\vec{D})-a}{2}, \frac{3w(\vec{D})-b}{2}}(\vec{D}),$$

where $w(\vec{D})$ is writhe. We can get the Kauffman bracket polynomial (KBP) of a link from its KS using the Kauffman bracket state sum formula (KBSS):

$$[D] = \sum_{S \in EKS} (-1)^{|D_s|} A^{\sigma(s)+2\tau(S)},$$

where $|D_s|$ is the number of cycles in a KS and A is the variable in the KBP. From algebraic topology, we see the KBP is an Euler characteristic of Khovanov homology–the generating Poincaré polynomial for the chain groups is:

$$\sum_{a,b} rank(C_{a,b}) x^a y^b$$

Substituting $x = -i$ and $y = iA$ gives the KBSS through a parity argument.

Torsion

The Khovanov homology of links often contains torsion, which is information not found in the Jones Polynomial. The topological meaning of this torsion is an active area of research.
Conjecture (Shumakovitch): The Khovanov homology of every link, excepting the unknot, the Hopf Link, and connected sums or disjoint unions of the two contains $\mathbb{Z}_2$ torsion.
Attempts to find links with high order torsion are ongoing. In 2020, Sujoy Mukherjee found $\mathbb{Z}_{81}$ torsion with the link $T(2,3)\#_4(\sigma_1\sigma_2\sigma_3)^4\sigma_1\sigma_2$. The previously highest known odd order was $\mathbb{Z}_7$.

Acknowledgments & References

Thank you to my mentor Rhea Palak Bakshi for her enthusiasm, unending patience, and knowledge. Thank you to the UCSB DRP for this opportunity.

1. M. Khovanov, A categorification of the Jones polynomial, Duke Math. J. 101 (3) (2000) 359–426
2. O. Viro, Khovanov homology, its definitions and ramifications, Fund. Math. 184 (2004) 317–342.
3. J. H. Przytycki, R. P. Bakshi, D. Ibarra, G. Montoya-Vega and D. E. Weeks, Lectures in Knot Theory: An Exploration of Contemporary Topics, Springer Universitext (Springer International Publishing, 2024)
4. Shumakovitch, A. N.: Torsion of Khovanov homology. Fund. Math. 225, 343–364 (2014) Zbl 1297.57022 MR 3205577
5. S. Mukherjee, On odd torsion in even Khovanov homology. Exp. Math. 0(2020), 1–7.
6. Kronheimer, P. B., Mrowka, T. S.: Khovanov homology is an unknot-detector. Publ. Math. Inst. Hautes Études Sci. 113, 97–208 (2011) Zbl 1241.57017 MR 2805599

Symmetry in Molecules: Group Representation Meets Chemistry

Xi He

University of California, Santa Barbara
Department of Mathematics, Directed Reading Program 2025

Why Study Molecular Symmetry?

A molecule has symmetry if it looks the same after certain movements (e.g., rotating 180°, flipping across a plane). Group theory provides tools to analyze this symmetry and predict properties like energy levels, light absorption, and chemical behavior.

From Symmetry to Group Representations

A **representation** of a group G on a vector space V over $\mathbb{C}$ is a group homomorphism:

$$\rho : G \to \mathrm{GL}(V).$$

This means each group element acts as an invertible linear transformation on V, preserving the group structure.

To study group representations systematically, we highlight several key notions:

- **G-invariant subspace**: A subspace $W \subset V$ is called *G-invariant* if $\rho(g)(w) \in W$ for all $g \in G$ and $w \in W$.
- **Irreducible representation**: A representation is *irreducible* if it has no nontrivial G-invariant subspace.
- **Maschke's Theorem**: Every finite group representation over $\mathbb{C}$ is a direct sum of irreducible representations.
- **Character of a representation**: The *character* χ of a representation ρ is the function

$$\chi(g) = \mathrm{tr}(\rho(g)).$$

- **Orthogonality relations**: The irreducible characters form an orthonormal basis with respect to the inner product $\langle \chi_i, \chi_j \rangle = \frac{1}{|G|} \sum_{g \in G} \chi_i(g) \overline{\chi_j(g)}$.
- **Character table**: A table with rows for irreducible representations and columns for conjugacy classes.
- **Example: Character table of S_3**

	$[1]$	$[(12)]$	$[(123)]$
χ_{triv}	1	1	1
χ_{sign}	1	-1	1
χ_{std}	2	0	-1

where $[1] = \{1\}$, $[(12)] = \{(12), (13), (23)\}$, and $[(123)] = \{(123), (132)\}$.

What can you observe from the character table?

Symmetry Operations and Point Groups

- A **symmetry operation** is an action that leaves an object in a position that looks exactly the same as before.
- A **point group** is the set of all symmetry operations that leave at least one point fixed and map the molecule onto itself.

Operator	Description
$\boldsymbol{E}$	Identity operation (no change to the molecule)
$\boldsymbol{C}_n$	Rotation by $\frac{2\pi}{n}$ about an axis of symmetry
$\boldsymbol{\sigma}$	Reflection through a mirror plane
$\boldsymbol{S}_n$	Improper rotation: $\boldsymbol{C}_n$ followed by reflection in a plane perpendicular to that axis
$\boldsymbol{i}$	Inversion through the center of the molecule

Table: Fundamental symmetry operations in molecular point groups

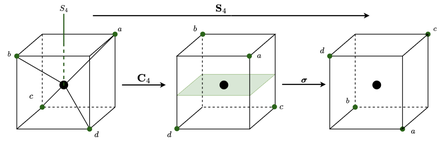
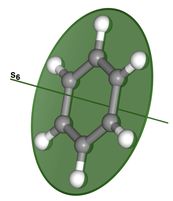


Figure: Improper rotational axis of Methane CH_4

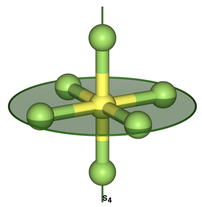
Molecular Point Groups and Their Symmetry Elements

The table below lists several commonly encountered point groups along with their defining symmetry elements.

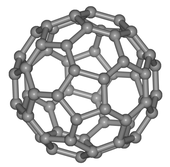
Point Group	Essential Symmetry Element(s)	Example Molecule
$\mathcal{C}_s$	One symmetry plane	Formic acid HCOOH
$\mathcal{C}_i$	Center of inversion	Ethene C_2H_4 (planar)
$\mathcal{C}_2$	One 2-fold axis	Hydrogen peroxide H_2O_2
$\mathcal{C}_{2v}$	$\boldsymbol{C}_2$ + two vertical planes $\boldsymbol{\sigma}_v$	Water H_2O
$\mathcal{D}_{6h}$	High symmetry: $\boldsymbol{C}_6$, $\boldsymbol{\sigma}_h$, etc.	Benzene C_6H_6
$\mathcal{T}_d$	Tetrahedral symmetry	Methane CH_4
$\mathcal{O}_h$	Octahedral symmetry	Sulfur hexafluoride SF_6
$\mathcal{I}_h$	Icosahedral symmetry	Buckminsterfullerene C_{60}



C_6H_6



SF_6



C_{60}

Basis Functions in Character Tables

In chemistry, the character table often includes a *Basis Functions* column.

The water molecule belongs to the point group $\mathcal{C}_{2v}$, whose symmetry elements and corresponding character table are shown below.

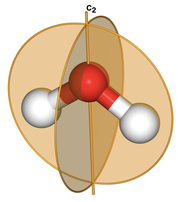


Figure: C_2 and σ_v of H_2O

$\mathcal{C}_{2v}$	$\boldsymbol{E}$	$\boldsymbol{C}_2$	$\boldsymbol{\sigma}_v(xz)$	$\boldsymbol{\sigma}_{v'}(yz)$	Basis Functions
A_1	1	1	1	1	z, x^2, y^2, z^2
A_2	1	1	-1	-1	R_z, xy
B_1	1	-1	1	-1	x, R_y, xz
B_2	1	-1	-1	1	y, R_x, yz

Table: Character table of $\mathcal{C}_{2v}$ with basis functions

- A **basis function** is a function that spans an invariant subspace under the action of a symmetry group and transforms according to a specific irreducible representation.
- In $\mathcal{C}_{2v}$ character table, the basis functions are the Cartesian axes x, y, z, the Cartesian products $x^2, y^2, z^2, xy, zx, yz$, and the rotations R_x, R_y, R_z.
- For example, the function x transforms as $(x, -x, x, -x)$ under the symmetry operations $\boldsymbol{E}, \boldsymbol{C}_2, \boldsymbol{\sigma}_v(xz), \boldsymbol{\sigma}_{v'}(yz)$, which corresponds to the irreducible representation B_1.
- Basis functions are essential for analyzing molecular vibrations and understanding how molecules interact with light.
- We can use character tables and basis functions to:
  - Predict which vibrations are IR or Raman active.
  - Classify vibrational modes based on symmetry

Group Theory Predicts IR and Raman Activity

- A vibration is **Infarred (IR) active** if it causes a change in dipole moment. This corresponds to basis functions x, y, or z.
- A vibration is **Raman active** if it changes the polarizability tensor. This corresponds to quadratic basis functions such as x^2, xy, etc.
- By matching vibrational motions with the irreducible representations in the character table, we can determine their activity.

Molecular Symmetry in Cyclohexane and Ferrocene

Cyclohexane C_6H_{12}: Chair Symmetry

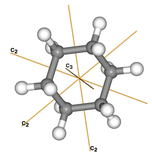


Figure: Front view with symmetry axes

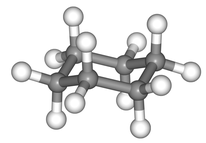


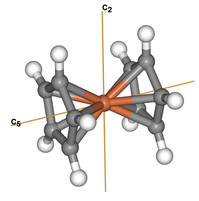
Figure: Side view of chair conformation

$\mathcal{D}_{3d}$	$\boldsymbol{E}$	$2\boldsymbol{C}_3$	$3\boldsymbol{C}_2$	$\boldsymbol{i}$	$2\boldsymbol{S}_6$	$3\boldsymbol{\sigma}_d$	Basis Functions
A_{1g}	1	1	1	1	1	1	$x^2 + y^2, z^2$
A_{2g}	1	1	-1	1	1	-1	R_z
E_g	2	-1	0	2	-1	0	(R_x, R_y), $(x^2 - y^2, xy)$, (xz, yz)
A_{1u}	1	1	1	-1	-1	-1	
A_{2u}	1	1	-1	-1	-1	1	z
E_u	2	-1	0	-2	1	0	(x, y)

Table: Character table of $\mathcal{D}_{3d}$ with basis functions

Ferrocene $Fe(C_5H_5)_2$: A Metal Sandwich with 5-Fold Symmetry

- Ferrocene ($Fe(C_5H_5)_2$) is a prototypical *sandwich compound*, consisting of an iron atom coordinated between two cyclopentadienyl (Cp) rings.
- In its **eclipsed conformation**, ferrocene belongs to the high-symmetry point group $\mathcal{D}_{5h}$, which features a fivefold rotation axis $\boldsymbol{C}_5$, a horizontal mirror plane $\boldsymbol{\sigma}_h$, and multiple vertical mirror planes.



- This high symmetry gives rise to degenerate molecular orbitals, which can be systematically classified using group theory and labeled by irreducible representations.
- These orbitals explain ferrocene's exceptional electronic stability and aromatic character.

Acknowledgment

I would like to sincerely thank my mentor Qing Zhang and the DRP for their invaluable guidance and support. Special thanks to my friend Jane for her patient and insightful explanations of chemistry concepts.

References

[1] M. Artin, *Algebra*, 2nd ed., Pearson, 2011.
[2] D. M. Bishop, *Group Theory and Chemistry*, Oxford University Press, 1993.
[3] R. Bacher, *Symotter – Symmetry Group Otter*, `https://symotter.org/`.

Persistent Homology

Luke Crossley

University of California Santa Barbara

Introduction to Homology

Homology is, fundamentally, the study and identification of holes in topological spaces. Each of these spaces can be created out of different dimensional objects called **simplicial complexes**. These complexes are constructed out of a union of points, line segments, planes, and higher dimensional polytopes, which are call n-simplicies depending on the dimension of the object. Every simplicial complex, denoted as K, is a collection of subsets of the vertices of the object of interest. As discussed in [1], we will study collections of formal sums of For simplicity, we will study simplicial complexes as vector spaces over $\mathbb{F}_2$, the field with two elements $\{0, 1\}$. Let $C_n(X)$ denote the vector space with basis given by the n-simplices of X. Thus, we can define the **boundary map** for $n = 1, 2, ...$ as

$$\partial_n : C_n(X) \to C_{n-1}(X)$$

which sends each n-simplex to its boundary. We also see that the boundary of a boundary is empty; therefore, the composition of any two consecutive boundary maps is the zero map. From here, we can define what is known as a **chain complex**

$$0 \xrightarrow{\partial_{n+1}} C_n(X) \xrightarrow{\partial_n} C_{n-1}(X) \xrightarrow{\partial_{n-1}} ... \xrightarrow{\partial_2} C_1(X) \xrightarrow{\partial_1} C_0(X) \xrightarrow{\partial_0} 0$$

To compute the homology vector space, $H_n(X)$ from the chain complex we find the quotient

$$H_n(X) = \frac{\ker(\partial_n)}{\text{im}(\partial_{n+1})}$$

To explain in simplest terms, the quotient ensures each hole is counted exactly once, identifying any two cycles of the same hole as the same cycle in the space $H_n(X)$. The dimension of $H_n(X)$ gives the number of n-dimensional holes present, called the nth **Betti number** of X.

Filtered Simplicial Complexes and Persistent Homology

Persistent Homology attempts to study the homology of a space over some notion of time given a metric to determine the distance or dissimilarity between any two points in the space. To study this further, we introduce the concept of the filtered simplicial complex. Given a finite simplicial complex X and finite sequence of subcomplexes $X_1 \subset X_2 \subset ... \subset X_k \subset X$, we call X a filtered simplicial complex. While we can compute the homology for each dimension present within each of the subcomplexes, we will discuss an algorithm that provides a more efficient computation and visualization of the significant features present in the simplicial complex. The filtered simplicial complex is created, in most cases, from a point cloud data set using something known as the **Vietoris-Rips complex** [2]. The VR complex defines some $\epsilon > 0$ and adds a 1-simplex between any two points that have a pairwise distance of less than 2ϵ based on the defined metric. Higher dimensional simplices are added as the distances between closed loops of n-simplices are less than 2ϵ. A simple way of viewing this is assuming that each point has some $\epsilon-$neighbourhood defined around it and when the neighbourhoods of any two vertices intersects, a 1-simplex is added connecting the points. This concept be seen in the following figure, although the metric used does not follow strictly from an $\epsilon-$neighbourhood.

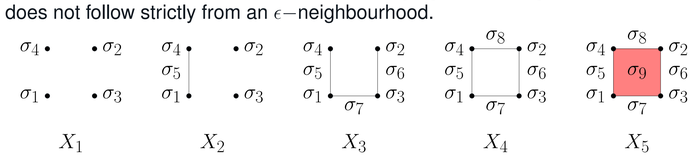


Figure 1: Visual Example of a Filtered Simplicial Complex

The Standard Algorithm and Barcode Plots

To create a graphic that displays birth and death of different features, we seek to create the barcode plot. First, we define the **boundary matrix**. Let n be the number of simplices in the simplicial complex X and denote the simplices as $\sigma_1, ..., \sigma_n$ ordered in this manner. We construct the $n \times n$ boundary matrix $B = \{b_{i,j}\}$ where each element $b_{i,j}$ is either a 1 if σ_i is a subset of the vertices of σ_j or a 0. In [2], they discuss what they call the **Standard Algorithm** to reduce the boundary matrix. First, they define $low(j)$ to be the largest value of i such that $b_{i,j} \neq 0$ where $i, j \in \{1, ..., n\}$. If the jth column contains only values of 0, then $low(j)$ is undefined. Now, we can discuss the standard algorithm:

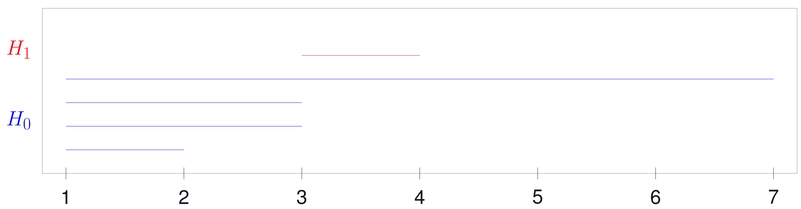
for $j = 1$ *to* n **do**
  while *there exists* $i < j$ *and* $low(i) = low(j)$ **do**
    add column i to column j
  end
end

Algorithm 1: The Standard Algorithm for Reducing the Boundary Matrix

With a reduced boundary matrix B', we can determine the **birth and death** of the features that are present throughout the filtration steps of the filtered complex. If $low(j) = i$, then we say σ_j is paired with σ_i meaning that a feature is born when σ_i enters the filtration and dies or disappears when σ_j enters into the filtration. If $low(j)$ is undefined, then σ_j entering the filtration causes a feature to be born. If there exists an l such that $low(l) = j$, then the feature born due to the entrance of σ_j dies with the entrance of σ_l into the filtration. With no such l, then the feature generated by σ_j's entrance remains throughout the entire filtration. Given the filtered complex laid out in Figure 1, we can create the boundary matrix and follow the standard algorithm to reduce the boundary matrix. Thus, we have the boundary matrix, B, and the reduced boundary matrix, B', from the standard algorithm (in $\mathbb{F}_2$):

$$B = \begin{pmatrix} 0&0&0&0&1&0&1&0&0 \\ 0&0&0&0&0&1&0&1&0 \\ 0&0&0&0&0&1&1&0&0 \\ 0&0&0&0&1&0&0&1&0 \\ 0&0&0&0&0&0&0&0&1 \\ 0&0&0&0&0&0&0&0&1 \\ 0&0&0&0&0&0&0&0&1 \\ 0&0&0&0&0&0&0&0&1 \\ 0&0&0&0&0&0&0&0&0 \end{pmatrix} \qquad B' = \begin{pmatrix} 0&0&0&0&1&0&1&0&0 \\ 0&0&0&0&0&1&1&0&0 \\ 0&0&0&0&0&1&0&0&0 \\ 0&0&0&0&1&0&0&0&0 \\ 0&0&0&0&0&0&0&0&1 \\ 0&0&0&0&0&0&0&0&1 \\ 0&0&0&0&0&0&0&0&1 \\ 0&0&0&0&0&0&0&0&1 \\ 0&0&0&0&0&0&0&0&0 \end{pmatrix}$$

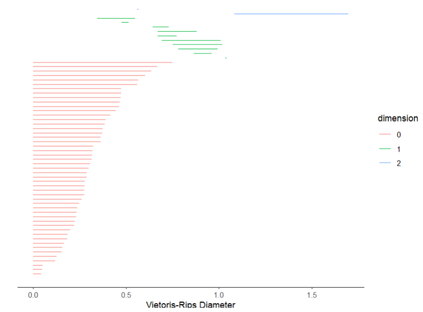
Thus, we get the intervals $[1, \infty)$ for σ_1 unpaired, $[1, 3)$ for σ_2, which is paired with σ_7, $[1, 2)$ for σ_4, which is paired with σ_5, $[4, 5)$ for σ_8 paired with σ_9, and $[1, 3)$ for σ_3 which is paired with σ_6. These were found following the process defined in [2].
From Figure 1, we can create the **barcode plot** by reading from the intervals of the reduced boundary matrix.



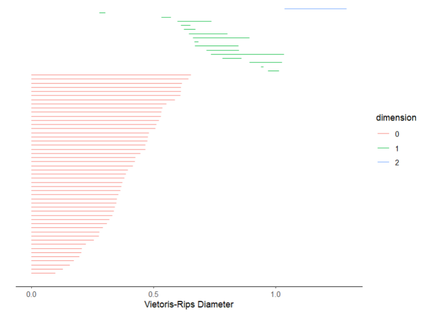
The above barcode for Figure 1 shows the presence of 0 and 1-dimensional holes in the filtered simplicial complex as well as how long these holes persist. The 1-dimensional hole persists only between 3 and 4, indicating that it is not a particularly significant feature of the space X. However, barcode plots are not dependent on which stage of the filtered complex each simplex appears at but rather the chosen metric, i.e., the value of ϵ. Thus, it is possible that ϵ increased a great deal more between 3 and 4 than between 1 and 2. Since we did not choose a specific metric, we displayed the barcode based on the subcomplexes.

Barcode Plots of Different Spheres

We now examine the differences between two different barcode plots based on two slightly differing filtered complexes. Both complexes are created by generating 50 points along the surface of S^2. In Case 1, the points lie exactly along the surface of S^2, while in Case 2 each point is slightly perturbed, only roughly being along the surface of S^2. These plots were generated using the TDAstats package in R. First, we look at the barcode plot for Case 1.



We see that the longest bar in dimension 2 suggests that there is a persistent generator of H_2. Since there are no other significant generators in either dimension 1 or 0, we can conclude that the space we are in is homologically S^2. We can view this as pointing to the existence of the hollow inside of S^2 before the interior is filled in by a 3-simplex. Next, we look at the barcode plot of Case 2.



We see slight differences mainly in a greater dispersal of the H_1 generators with two of those generators persisting long enough that they seem like significant features that define the space we are in. While we have a persistent H_2 generator similar to that of the previous barcode plot, it is about as persistent as the H_1 generators. Thus, the perturbation of the points suggests the space we are in is no longer homologically S^2.

Acknowledgements

I thank Andrew Ortegaray for his guidance as well as the UCSB Directed Reading Program for the opportunity to work on this project.

References

[1] Allen Hatcher. *Algebraic Topology*. Cambridge University Press, 2001.

[2] Nina Otter et al. "A roadmap for the computation of persistent homology". In: *EPJ Data Science* (2017).

General Stokes' Theorem

Author: Hunter Lin Mentor: Yusen Xia
University of California, Santa Barbara

Goals

In vector calculus, we learn **Stokes' Theorem** in familiar forms like **Green's Theorem**, **the classical Stokes' Theorem**, and **the Divergence Theorem**, which relate integrals over a region to integrals over its boundary. Our goal is to explore and understand the deeper structure behind these results by generalizing Stokes' Theorem to smooth manifolds using **differential forms**. This unified, coordinate-free formulation reveals the theorem as a powerful tool in modern geometry and physics — one that applies across any dimension or shape.

General Stokes' Theorem

Let M be an orientable smooth n-manifold, and let ω be a compactly supported smooth $(n-1)$-form on M. Then, we have:

$$\int_M d\omega = \int_{\partial M} \omega$$

Here if Ω is a chosen to be an orientation of M, then we will take $i_\eta \Omega$ to be the orientation of ∂M where η is an outward-point normal vector of ∂M. In particular, if $\partial M = \emptyset$, then $\int_M d\omega = 0$.

Smooth Manifolds

A n-dimensional topological manifold M is said to be an n-dimensional topological manifold, if there is a collection A of local parametrizations $F_\alpha : U_\alpha \to O_\alpha$ such that

- $\bigcup_{\alpha \in A} O_\alpha = M$, i.e. these local parametrizations cover all of M; and
- all transition maps $F_\alpha^{-1} \circ F_\beta$ are smooth (i.e C^∞) on their domains.

Differential Forms

In vector calculus, we integrate vector fields over surfaces or curves using tools like line integrals and flux. These make sense because the domains — like regions in $\mathbb{R}^2$ or $\mathbb{R}^3$— are covered by a single coordinate system. However, by the definition of manifolds, we might have many overlapping coordinates chart which makes the vector field make no sense to have different value on the same coordinates. Thus, we need to introduce Differential Forms. Let M be a smooth manifold. A smooth differential k-form ω on M is a map $\omega_p : T_pM \times T_pM \times \cdots \times T_pM \to \mathbb{R}$ at each $p \in M$ such that under any local parametrization $F(u_1, \ldots, u_n) : U \to M$, it can be written in the form:

$$\omega = \sum_{i_1,\ldots,i_k=1}^{n} \omega_{i_1 \ldots i_k} du^{i_1} \wedge \ldots du^{i_k}$$

where $\omega_{i_1 \ldots i_k}$'s are smooth scalar functions locally defined in $F(u)$, and they are commonly called the local components of ω. The vector space of all smooth differential k-forms on M is denoted by $\wedge^k T^*M$

Exterior Derivative

We define the exterior derivative for a differential form. Let M be a smooth manifold. Let M^n be a smooth manifold and $(u_1, \ldots, u_n)$ be local coordinates on M. Given any (smooth) k-form

$$\omega = \sum_{j_i,\ldots,j_n=1}^{n} \omega_{j_1 \ldots j_k} du^{j_1} \wedge \cdots \wedge du^{j_k}$$

We define:

$$d\omega := \sum_{j_i,\ldots,j_n=1}^{n} \sum_{i=1}^{n} \frac{\partial \omega_{j_1 \ldots j_k}}{\partial u_i} du^i \wedge du^{j_1} \wedge \cdots \wedge du^{j_n}$$

In particular, if ω is an n-form (where $n = dimv$), we have $d\omega = 0$. Like the normal derivative, if we take a derivative to the first derivative, we will get k+1 form if we take external derivative toward k-form. It satisfies $d^2 = 0$ and obeys a graded Leibniz rule.

Manifolds with Boundary

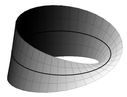
Now we need to define what is the boundary for a manifolds. A n-dimensional topological manifold M is said to be an n-dimensional topological manifold, or in short a topological n-manifold, if there is a collection $\mathcal{A}$ of local parametrizations $F_\alpha : \mathcal{U}_\alpha \to O_\alpha$ such that

- $\bigcup_{\alpha \in \mathcal{A}} O_\alpha = M$, i.e. these local parametrizations cover all of M; and
- all transition maps $F_\alpha^{-1} \circ F_\beta$ are smooth (i.e C^∞) on their domains.

Since $d \circ d = 0$ for all differential forms, we also get $\partial(\partial M) = \emptyset$.

Orientable Manifolds

Like vector calculus, we need to define the orientation of a manifold. For example, Möbius Strip is not an orientable manifold.



A smooth manifold M is said to be orientable if there exists a family of local parametrizations $F_\alpha : \mathcal{U}_\alpha \to M$ covering M such that for any F_α and F_β in the family with $F_\beta(\mathcal{U}_\beta) \cap F_\alpha(\mathcal{U}_\alpha) \neq \emptyset$, we have:

$$\det\ D(F_\alpha^{-1} \circ F_\beta) > 0 \ on\ F_\beta^{-1}(F_\beta(\mathcal{U}_\beta) \cup F_\alpha(\mathcal{U}_\alpha)).$$

In this case, we call the family $\mathcal{A} = \{F_\alpha : \mathcal{U}_\alpha \to M\}$ of local parametrizations to be an oriented atlas of M.

We also define the orientation of the manifold. Given an orientable manifold M^n, a non-vanishing global n-form Ω is called an orientation of M. A basis of tangent vectors $\{T_1, \ldots, T_n\} \in T_pM$ is said to be Ω-oriented if $\Omega(T_1, \ldots, T_n) > 0$. A local coordinate system $(u_1, \ldots, u_n)$ is said to be Ω-oriented if $\Omega(\frac{\partial}{\partial u_1}, \ldots, \frac{\partial}{\partial u_n})$

Proof of General Stokes' Theorem

Step 1 We first show that for the special case where $supp\ \omega$ is contained inside a single parametrization chart of interior type.

$$\begin{aligned}\int_M d\omega &= \int_{\mathcal{U}} \sum_{i=1}^{n} (-1)^{i-1} \frac{\partial \omega_i}{\partial u_i} du^1 \ldots du^n \\ &= \sum_{i=1}^{n} (-1)^{i-1} \int_{-R}^{R} \cdots \int_{-R}^{R} [\omega_i]_{u_i=-R}^{u_i=R} du^1 \ldots du^n\end{aligned}$$

Since ω_i's vanish at the boundary of the rectangle $[-R, R]^n$, we have $\int_M d\omega = 0$. Since $supp\ \omega$ is contained in a single parametrization chart of the interior type, we have $\omega = 0$ on the boundary ∂M. Thus, we have:

$$\int_M d\omega = 0 = \int_{\partial M} \omega$$

References

[1] Anish Athalye. Gemini: A modern latex beamerposter theme, 2016.

[2] Frederick Tsz-Ho Fong. *Differentiable Manifolds & Riemannian Geometry*.

Step 2 We then show that for the special case where $supp\ \omega$ is contained inside a single parametrization chart of boundary type.

$$\begin{aligned}\int_M d\omega &= \int_{\mathcal{V}} \sum_{i=1}^{n} (-1)^{i-1} \frac{\partial \omega_i}{\partial u_i} du^1 \ldots du^n \\ &= \sum_{i=1}^{n-1} (-1)^{i-1} \int_0^R \int_{-R}^R \cdots \int_{-R}^R (-1)^{i-1} \frac{\partial \omega_i}{\partial u_i} du^1 \ldots du^n + (-1)^{n-1} \int_0^R \int_{-R}^R \cdots \int_{-R}^R \frac{\partial \omega_i}{\partial u_i} du^1 \ldots \\ &= (-1)^{n-1} \int_0^R \int_{-R}^R \cdots \int_{-R}^R \frac{\partial \omega_i}{\partial u_i} du^1 \ldots du^n \\ &= (-1)^{n-1} \int_{-R}^R \cdots \int_{-R}^R \omega_n(u_1, \ldots, u_{n-1}, 0) du^1 \ldots du^n \\ \int_{\partial M} \omega &= (-1)^n \int_{\mathcal{V} \cap \{u_n = 0\}} \omega_n(u_1, \ldots, u_{n-1}, 0) du^1 \ldots du^n \\ &= (-1)^{n-1} \int_{-R}^R \cdots \int_{-R}^R \omega_n(u_1, \ldots, u_{n-1}, 0) du^1 \ldots du^n\end{aligned}$$

Thus, we have:

$$\int_M d\omega = \int_{\partial M} \omega$$

Step 3 Finally, we "glue" the previous two steps together and deduce the general case. Let $\mathcal{A} = \{F_\alpha : \mathcal{U}_\alpha \to M\}$ be an atlas of M where all local coordinates are Ω-oriented. Suppose $\{\rho_\alpha : M \to [0, 1]\}$ is a partition of unity of subordinate to $\mathcal{A}$. Then we have:

$$\begin{aligned}\int_{\partial M} \omega &= \int_{\partial M} \sum_\alpha \rho_\alpha \omega \\ &= \sum_\alpha \int_{\partial M} \rho_\alpha \omega \\ &= \sum_\alpha \int_M d(\rho_\alpha \omega) \\ &= \int_M d(\sum_\alpha \rho_\alpha) \wedge \omega + (\sum_\alpha \rho_\alpha) \wedge d\omega \\ &= \int_M d\omega\end{aligned}$$

Application to Vector Calculus

Green's Theorem

Let R be a closed and bounded smooth 2-submanifold in $\mathbb{R}^2$ with boundary ∂R. Given any smooth vector field $V = (P(x, y), Q(x, y))$ defined in R, then we have:

$$\oint_{\partial R} V \cdot dl = \int_R (\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y})\ dx\ dy$$

The line integral on the LHS is oriented such that $\{\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\}$ has the same orientation as $\{\mu, T\}$ where μ is the outward-pointing normal of R, and T is the velocity vector of the curve ∂R.

Consider the 1-form $\omega := P\ dx + Q\ dy$ defined on R, then we have:

$$d\omega = (\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y})\ dx \wedge dy$$

Suppose we fix an orientation $\Omega = dx \wedge dy$ for R so that the order of coordinates is (x, y), then by generalized Stokes' Theorem we get:

$$\underbrace{\oint_{\partial R} P\ dx + Q\ dy}_{\oint_{\partial R} \omega} = \underbrace{\int_R (\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y})\ dx \wedge dy}_{\int_R d\omega} = \int_R (\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y})\ dx\ dy$$

Tannaka Duality

Yunlong Xu (Mentor: Quinn Kolt)

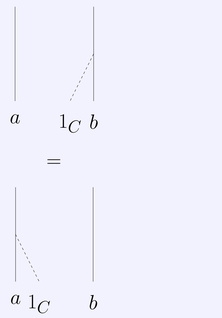
Department of Mathematics, University of California, Santa Barbara

Monoidal Category

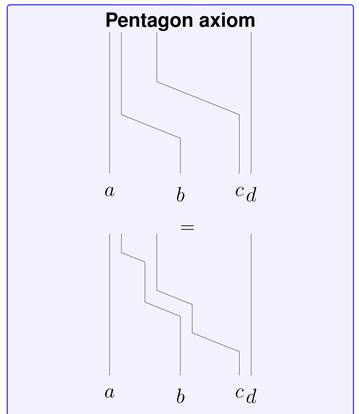
Definition 1.1 A Monoidal category is a tuple $(C, \otimes, 1_C, \alpha_{a,b,c}, \lambda_a, \rho_a)$ where $- \otimes - : C \times C \to C$ is a bifunctor called the **tensor product**, 1_C is an object called the **unit object** in C, $\alpha_{(a,b,c)} : (a \otimes b) \otimes c \to a \otimes (b \otimes c)$ is a natural isomorphism called the **associator**, $\lambda_a : 1_C \otimes a \to a$ and $\rho_a : a \otimes 1_C \to a$ are natural isomorphism called **unitors**.

This data must satisfy the following two axioms:

Triangle axiom



Pentagon axiom



Examples:

- ($G-$graded vector spaces) Let G be a finite group and Vec(G) be the category of all finite dimensional $G-$graded complex vector space $V = \bigoplus_{g \in G} V_g$ with grading-preserving linear map, i.e. , if $T : V \to W$, then $T(V_g) \subseteq W_g$. We endow Vec(G) with the structure of tensor product as follows:

$$(V \otimes W)_g := \bigoplus_{hk=g} V_h \otimes W_k.$$

  The associator moves the parentheses, and the unit object is the field $\mathbb{C}$. Under this definition, Vec(G) is a monoidal category.

- (Representation for G) Let G be a finite group. **Rep**$_k(G)$ is a category of all representations of G over complex field $\mathbb{C}$. The objects in this category are pairs (V, π_V), where $\pi_V : G \to GL(V)$ is the homomorphism for the representation V. We define the tensor product of pairs as:

$$(V, \pi_V) \otimes (W, \pi_W) = (V \otimes W, \pi_{V \otimes W}), \pi_{V \otimes W}(g) := \pi_V(g) \otimes \pi_W(g).$$

  The unit is the trivial representation $1 = k$, and the associator moves parentheses. Then **Rep**$_k(G)$ is a monoidal category. A similar statement holds for the category Rep$_k(G)$ of all finite-dimensional representations of G.

Theorem 1.2 (Coherence) Any monoidal category is monoidally equivalent to a strict monoidal category.

Tensor Category

Definition 2.1 Rigidity is the generalization of duality. The dual of an object $X \in C$ is an object $X^\vee$ with evaluation and coevaluation morphisms $\mathrm{ev}_c : c^\vee \otimes c \to 1_c$ and $\mathrm{coev}_c : 1_c \to c \otimes c^\vee$ satisfying the following relation

$$(\mathrm{id}_X \otimes \mathrm{ev}_X) \circ (\mathrm{coev}_X \otimes \mathrm{id}_X) = \mathrm{id}_X, \qquad (\mathrm{ev}_X \otimes \mathrm{id}_{X^\vee}) \circ (\mathrm{id}_{X^\vee} \otimes \mathrm{coev}_X) = \mathrm{id}_{X^\vee}.$$

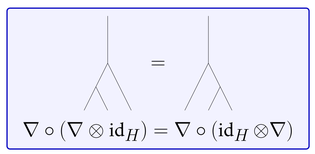
Definition 2.2 A **finite tensor category** is a locally finite k-linear abelian rigid monoidal category such that there are enough projective covers of simple objects, finitely many isomorphism classes of simple objects, and 1_C is simple.

Hopf Algebras

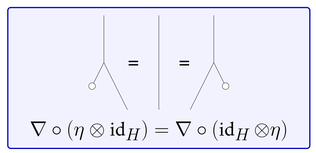
Definition 3.1 A Hopf algebra is a sextuple $(H, \nabla, \eta, \Delta, \varepsilon, S)$ where $\nabla : H \otimes H \to H$ is **multiplication**, $\eta : \mathbb{F} \to H$ is **unit**, $\Delta : H \to H \otimes H$ is **comultiplication**, $\varepsilon : H \to \mathbb{F}$ is **counit**, and $S : H \to H$ is **antipode**. All those maps are linear.

This data must satisfy the following axioms:

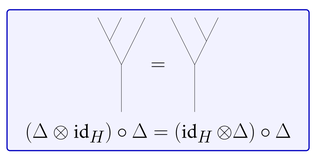
1 **Associativity axiom**;



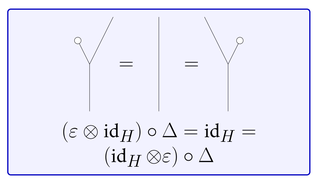
2 **Unital axiom** ;



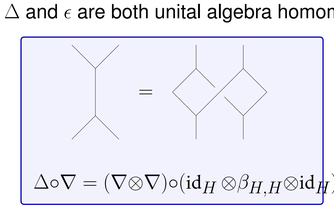
3 **Coassociativity axiom**;

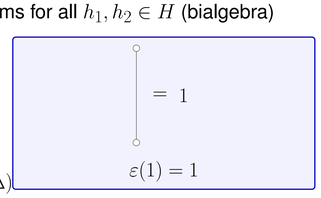


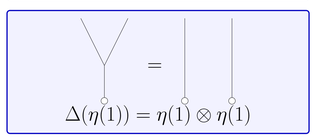
4 **Counital**;

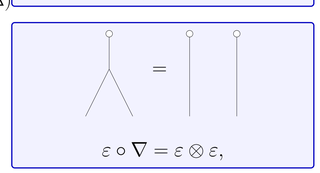


5. Δ and ϵ are both unital algebra homomorphisms for all $h_1, h_2 \in H$ (bialgebra)

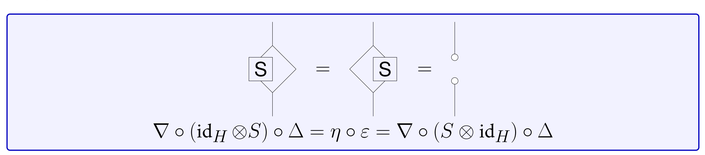








6 **Antipode axiom**.



Examples:

- (Trivial Hopf algebra) Let $\mathbb{F}$ be a field. It is also a Hopf algebra with multiplication and unit of the field, comultiplication $\Delta(a) = a \otimes 1_{\mathbb{F}}$, counit $\varepsilon(a) = a$, and antipode $S(a) = a$

- (Group Algebra) Let G be any group and $\mathbb{F}$ be a field. Considering the group ring $F[G]$. We define the multiplication $\nabla : F[G] \otimes F[G] \to F[G]$ by extending the usual group multiplication $g \circ h = g *_G h$ for $g, h \in G$. $e \in G$ is the unit object under this multiplication.

  The comultiplication $\Delta : F[G] \to F[G] \otimes F[G]$ is generated by $g \to g \otimes g$ for $g \in G$. By the counital axiom, $\varepsilon(g)g = g$, we could conclude that $\varepsilon(g) = 1_F$ for all $g \in G$. Antipode is defined as $S(g) = g^{-1}$ for all $g \in G$. All together makes $F[G]$ a Hopf algebra.

Tannaka Reconstruction

Definition 4.1 A **Fiber functor** is a exact faithful linear functor

$$F : C \to \mathrm{Vec}$$

preserving identity. It is also equipped with a natural isomorphism:

$$J_{X,Y} : F(X) \otimes F(Y) \to F(X \otimes Y), \text{ for } X, Y \in \mathrm{Obj}(C)$$

such that $J : F(\mathbb{1}_C) \to \mathbb{C}$

Theorem 4.2 Every finite tensor category equipped with a fiber functor is realized as the category of finite-dimensional representations of finite-dimensional Hopf algebras over the field $\mathbb{C}$. In particular,

$$(C, F) \to H := \mathrm{End}(F), H \to (\mathrm{Rep}(H), F)$$

are mutually bijections up to tensor equivalence.

Proof: Generally speaking, the proof is just checking that End(F) is a Hopf alegbra, (Rep(H), F) becomes a finite tensor category, and Rep(End(F)) $\cong C$
End(F) is equipped with an algebra structure. The crucial point for the End(F) is to construct the comultiplication, counit, and antipode. Comultiplication and axiom 5 make it into a bialgebra, while the Antipode and axiom 6 make it into a Hopf algebra. Comultiplication is a homomorphism from End(F) to End(F) $\otimes$ End(F). Consider the Delign's tensor product of End(F) with itself, there are canonical isomorphism α between End(F) $\otimes$ End(F) and End($F \boxtimes F$). For $X, Y \in C$, we have that $(F \boxtimes F)(X \boxtimes Y) = F(X) \otimes F(Y)$. Hence, for a natural transformation $a = \{a_Z : F(Z) \to F(Z) | \text{for all } Z \in C\} \in \mathrm{End}(F)$, we define a new natural transformation

$$\tilde{\Delta}_{X \boxtimes Y} = J^{-1}_{X \boxtimes Y} \circ \phi_{X \boxtimes Y} \circ J_{X \boxtimes Y}; \ (F \boxtimes F)(X \boxtimes Y) \to (F \boxtimes F)(X \boxtimes Y)$$

Pulling $\tilde{\Delta}$ back through α^{-1} lands in End(F) $\otimes$ End(F) and is defined as $\Delta(a)$. Coassociativity and the algebra map property of Δ then follow from the monoidal functor axioms for F and the multiplicativity of α. We then define the counit $\varepsilon : \mathrm{End}(F) \to \mathbb{C}$ by $\varepsilon(a) = a$ where $a_1 \in \mathrm{End}(F(1_C)) = \mathrm{End}(1) = \mathbb{C}$. Overall, we have that H is a bialgebra. Transporting the evaluation and coevaluation maps through the fibre functor F will obtain the usual linear duality maps on the vector spaces. One could define the antipode as $S(a_X) = a^*_{X^*}$. Hence, it makes H into a Hopf algebra.

On the other hand, $C = \mathrm{Rep}_{\mathbb{C}}(H)$ is a category whose objects are all finite representations of H over the field $\mathbb{C}$. The tensor product of H-modules X, Y is defined as the usual tensor product of vector spaces $X \otimes Y$. And the $H-$ action is defined as

$$h \cdot (x \otimes y) = \Delta(h)(x \otimes y) = (h^{(1)} \cdot x) \otimes (h^{(2)} \cdot y)$$

By the coassociativity and counital axioms, the tensor product here is associative and unital up to isomorphism, and the unit object is the trivial representation. Pentagon and triangular axioms hold because the underlying category is Vec. Every object in C is dualizable by the antipode axiom. The remaining axioms for a Finite Tensor Category are easy to verify.

And intuitively, $H = \mathrm{End}(F)$ contains every way that the objects of C can act on themselves inside vector spaces. And those actions could become modules that recreate all of C, giving the desired isomorphism of categories Rep(End(F)) $\cong C$

Acknowledgements and References

I would like to thank my mentor Quinn Kolt for her support and inspiration for this project, as well as the UCSB Directed Reading Program for this invaluable experience.

[1] Tensor categories by Etingof, Gelaki, Nikshych and Ostrik [2] https://people.math.osu.edu/penneys.2/8800//Notes/FusionCats.pdf
[3] https://arxiv.org/pdf/math/0401246

A Solution to Post-Quantum Security: NTRU

Marcus Nguyen [1] Nicholas Caputo [2]

[1]University of California - Santa Barbara
Department of Mathematics - Directed Reading Program 2025

UC **SANTA BARBARA**

Introduction

Cryptography is an area of applied mathematics that draws on disciplines such as number theory, abstract algebra, probability. These cryptosystems allow users to communicate securely across the globe, by basing their security on underlying mathematical problems which are hard to solve, that become easy once additional information is known. The most widely used system is called RSA which relies on the following problem. Given a large composite number n, is one able to find two large prime number p and q such that the product of these two prime numbers is n. Currently problems such as prime integer factorization are too difficult to solve given known algorithms and computing power. However with the power of quantum computing and known techniques such as Schor's Alogrithm (1994) we are able to solve these problems in realistic time frames.

This is where lattices come into the picture. We define a lattice as such

Definition. Let $v_1, \ldots, v_n \in \mathbb{R}^m$ be a set of linearly independent vectors. The *lattice* L generated by $v_1, \ldots, v_n$ is the set of linear combinations with coefficients in $\mathbb{Z}$,

$$L = \{a_1 v_1 + a_2 v_2 + \cdots + a_n v_n : a_1, a_2, \ldots, a_n \in \mathbb{Z}\}.$$

A *basis* for L is any set of independent vectors that generates L. Any two such sets have the same number of elements. The *dimension* of L is the number of vectors in a basis for L.

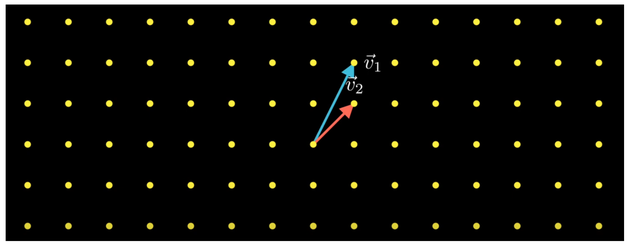


Figure 1. Visual of a 2-dimensional lattice.

Lattices provide a different approach to cryptography by inviting a new type of hard problem that are quantum resistant for the time being. This means that there are no known quantum algorithms that can efficiently solve hard lattice problems. This brings us to **NTRUEncrypt** a public key cryptosystem which is based on lattices. It can be shown that the security of **NTRU** is equivalent to solving the shortest vector problem.

The Shortest Vector Problem

The Shortest Vector Problem (SVP) can be described as such: Find a shortest nonzero vector in a lattice L, i.e., find a nonzero vector $v \in L$ that minimizes the Euclidean norm $||v||$.

The **SVP** can be intuitively thought of as such. Imagine we are standing at the origin of a lattice, what is the closest point to the origin that you can find and what is the length of the vector that describes this point. While in 2 or 3 dimensions this problem seems fairly simple to solve, as the number of dimensions grow, the difficulty of this problem increases as well.

Building off of this. The **apprSVP** is a modified version of the **SVP** in which we bound the euclidian norm of a vector $v \in L$ by the product of a scalar function n and the shortest vector in the lattice called $v_{shortest}$, where n is the dimension of our lattice.

$$||v|| \leq \psi(n)||v_{shortest}||$$

It should be noted that if we have a basis of a lattice where the vectors are pairwise orthogonal, then it is easy to solve both SVP and CVP (Closest Vector Problem).

NTRUencrypt

NTRUencrypt (N-th degree Truncated polynomial Ring Units) has three parts to it. Key creation, encryption and decryption. These three parts utilize two important mathematical objects, lattices and polynomial rings. We begin by fixing an integer $N \geq 1$ and two moduli p and q (where p and q are coprime), and we let R, R_p, and R_q be the convolution polynomial rings

$$R = \frac{\mathbb{Z}[x]}{(x^N - 1)}, \quad R_p = \frac{(\mathbb{Z}/p\mathbb{Z})[x]}{(x^N - 1)} \quad R_q = \frac{(\mathbb{Z}/q\mathbb{Z})[x]}{(x^N - 1)}.$$

In the equations above R is the name of the ring, $\mathbb{Z}[x]$ tells us that the coefficients of x are integers (in the case of $(\mathbb{Z}/p\mathbb{Z})[x]$ it would imply the coefficients are of p) and $x^N - 1$ indicates the greatest possible degree of the polynomial. In other words the highest degree will be (N-1). These rings will be the playground that our polynomials live in. Next we shall choose our private keys from the space of ternary polynomials $\mathcal{T}(d_1,d_2)$,

$$\mathcal{T}(d_1, d_2) = \left\{ \mathbf{a}(x) \in R : \begin{array}{c} \mathbf{a}(x) \text{ has } d_1 \text{ coefficients equal to } 1, \\ \mathbf{a}(x) \text{ has } d_2 \text{ coefficients equal to } -1, \\ \mathbf{a}(x) \text{ has all other coefficients equal to } 0 \end{array} \right\}.$$

Where d obeys the inequality $q > (6d{+}1)$. Ternary polynomials allow us to encrypt and decrypt at lighting speeds reducing our multiplication to simple arithmetic. Our private key consists of two randomly chosen polynomials

$$f(x) \in \mathcal{T}(d+1, d) \qquad \text{and} \qquad g(x) \in \mathcal{T}(d, d).$$

We compute the following inverse,

$$F_q(x) = f(x)^{-1} \text{ in } R_q \qquad \text{and} \qquad F_p(x) = f(x)^{-1} \text{ in } R_p$$

Our published key will be the following,

$$h(x) = F_q \star g(x)$$

Where $\star$ denotes the multiplication of polynomials with a reduced mod N of power. Using our public key $h(x)$ we can encrypt any message we like.

The NTRU Lattice

Let

$$h(x) = h_0 + h_1 x + ...h_{N-1} x^{N-1}$$

be an NTRU public key. The NTRU lattice of $h(x)$ is the $2N$-dimensional lattice spanned by the rows of the matrix

$$M_h^{\text{NTRU}} = \left(\begin{array}{cccc|cccc} 1 & 0 & \cdots & 0 & h_0 & h_1 & \cdots & h_{N-1} \\ 0 & 1 & \cdots & 0 & h_{N-1} & h_0 & \cdots & h_{N-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & h_1 & h_2 & \cdots & h_0 \\ \hline 0 & 0 & \cdots & 0 & q & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & q & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & q \end{array} \right).$$

Assuming $f(x) \star h(x) \equiv g(x) \pmod q$, let $u(x) \in \mathbb{R}$ be the polynomial satisfying,

$$f(x) \star h(x) = g(x) + qu(x).$$

Then

$$(f, -u) M_h^{\text{NTRU}} = (f, g),$$

so the vector (f, g) is in the NTRU lattice L_h^{NTRU}. In other words if one were to search for a private key to be able to decrypt a message in NTRU, then it is equivalent to solving the shortest vector problem. This is a difficult task given the shortest vector is not unique, the gap between the shortest and second shortest vector might be small, and in higher dimensions it becomes infeasible to brute-force search.

The LLL Algorithm

From the **SVP** section we know that given a basis of a lattice where the vectors are pairwise orthogonal, then it is easy to solve both **SVP** and **CVP**. This naturally gives birth to the question, given a basis of a lattice are we able to find a way to transform this basis into one where the vectors are as short and orthogonal as possible. Can we use this basis to solve SVP or CVP?

This is where the **LLL** comes into play. The **LLL Algorithm** is a basis reduction algorithm in which we reduce a given basis to a new basis where each vector is as short as possible starting with the shortest vector as our first and each following vector increases until we reach the last vector. Additionally we would also like these vectors to be as orthogonal as possible so **LLL** algorithm uses the Gram-Schmidt process in order to orthogonalize the basis vectors to our best ability.

Conditions of LLL

We can apply the following definition to a basis of a lattice

Definition: Let $B = \{v_1, v_2, \ldots, v_n\}$ be a basis for a lattice L and let $B^* = \{v_1^*, v_2^*, \ldots, v_n^*\}$ be the associated Gram-Schmidt orthogonal basis. The basis B is said to be *LLL reduced* if it satisfies the following two conditions:

1. (Size Condition) $|\mu_{i,j}| = \left| \frac{v_i \cdot v_j^*}{||v_j^*||^2} \right| \leq \frac{1}{2}$ for all $1 \leq j < i \leq n$.
2. (Lovász Condition) $||v_i^*||^2 \geq \left(\frac{3}{4} - \mu_{i,i-1}^2 \right) ||v_{i-1}^*||^2$ for all $1 < i \leq n$.

With these conditions being met this ensures that our new basis will be as small and orthogonal as possible. It should be noted that **LLL** runs in polynomial time and is guaranteed to terminate.

Solving Hard Lattice Problems

What a basis reduction algorithm like **LLL** tries to do is make a basis orthogonal enough so that we can solve the **apprCVP** or **apprSVP**. In lower dimensions the **LLL** algorithm works quite effectively. For instance **LLL** is able to solve apprSVP within a factor of $2^{\frac{n-1}{2}}$, however as the number of dimensions increase the **LLL** algorithm becomes quite ineffective and hence not feasible for large lattices. In general the security of lattice based cryptosystems depends on the inability of **LLL** and other lattice reduction algorithms to efficiently solve apprSVP and apprCVP to within a factor of, $O(\sqrt{n})$.

Unlike established cryptographic attack algorithms such as sieves or Pollard's ρ method, the performance characteristics of standard lattice reduction algorithms (e.g., BKZ-LLL) are less theoretically understood. This makes it challenging to precisely predict their effectiveness against specific lattice structures. Hence, for the time being the security of lattice based cryptosystems like NTRU must be determined experimentally.

Acknowlegments

We'd like to thank our mentor Jeremy Khoo for guiding us and opening our world to a new area of mathematics. We'd also like to thank the organizers of DRP for giving us the opportunity to present our findings.

References

[1] Jeffrey Hoffstein, Jill Pipher, and Joseph H. Silverman. Ntru: A ring-based public key cryptosystem. In Joe P. Buhler, editor, *Algorithmic Number Theory: Proceedings of the Third International Symposium (ANTS-III)*, volume 1423 of *Lecture Notes in Computer Science*, pages 267–288. Springer, 1998.

[2] Jeffrey Hoffstein, Jill Pipher, and Joseph H. Silverman. *An Introduction to Mathematical Cryptography*. Undergraduate Texts in Mathematics. Springer, 2008.

Knot Theory: You Will Knot Want To Miss This Poster!

Alexis Alfaro Naranjo, Elizabeth Paxtor, and Hannah Brard
University of California Santa Barbara, Department of Mathematics
May 15, 2025

Section 1: What Is Knot Theory?

What is a Knot

Knot theory is a subfield of the mathematical study of topology that focuses on knots, which can essentially be thought of as a knotted string that has no thickness and whose ends are connected. Additionally, these knots do not intersect each other at any point, they merely overlap one another, which are called crossings.

What are some Classifications

Knots largely fall into one of three categories: a torus knot, a satellite knot, and a hyperbolic knot. A torus is one that lies on an unknotted torus (visualize a doughnut), without crossing over or under themselves. They have two different types of curves: a meridian curve, which runs along the torus the "short way," and a longitude curve, which runs along the torus the "long way." We define a (p, q) torus knot to be one who intersects p times meridionally and q times longitudinally.

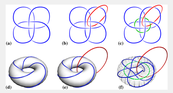


Figure 1: Torus Knots

A satellite knot is when we take a torus knot, with its own unique knot on the inside, and knot the torus itself. The knotting of the solid torus is called the companion knot, which is nontrivial.

Figure 2: Satellite Knot

Finally, a hyperbolic knot is a knot embedded on a hyperbolic surface, or surfaces that have a curvature of -1. Most knots tend to be hyperbolic knots. For example, think of a hyperbolic surface as the inside of the unit sphere. Any knot projected onto this surface is a hyperbolic knot. These knots have very interesting properties, such as the sum of angles of a triangle in this space is less than the standard 180 degrees.

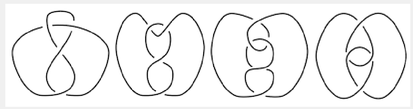


Figure 3: Hyperbolic Knots

Some Examples

Here are some examples of knots:

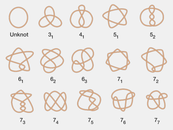


Figure 4: Examples of Famous Knots

Section 2: Properties of Knots

Reidemeister Moves

There are three ways in which we can change a projection of a knot. Reidemeister moves allow us to change the projection of a knot in one of three ways. The first Reidemeister move permits us to take out or put in a twist in the knot. The second Reidemeister move permits us to either add two crossings or remove two crossings.The third Reidemeister move permits us to slide a strand of the knot from one side of a crossing to the other side of the crossing.

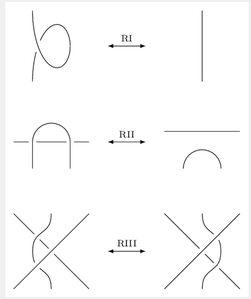


Figure 5: Reidemeister Moves

Two knot diagrams represent the same knot if and only if they can be transformed into each other through a sequence of Reidemeister moves and planar isotopies.

Tricolorability

A projection of a link or a knot is tricolorable when each of the strands in the projection can be colored one of three colors, so that at each crossing, either three different colors come together or all the same color comes together.

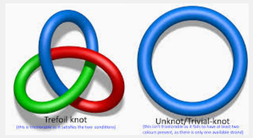


Figure 6: Tricolorability

Tricolorability is a knot invariant - if a knot is tricolorable, any diagram of that knot is also tricolorable.

Unknotting, Bridge, and Crossing Numbers

The unknotting number n of a knot K is determined by the fewest number of crossing changes that are required to transform the knot into the unknot. For example, the unknotting number of the unknot is 0 because there are zero crossing changes to be made in order to achieve the unknot. A composite knot is not able to be unknotted with just one crossing change.
The bridge number of the knot projection is determined by the number of maximal overpasses in that projection. An overpass is a sub arc of the knot that goes over at least one crossing, but never goes under any crossing. A maximal overpass is an overpass that cannot be made any longer.
The crossing number of any knot is the least number of crossings that the projection of the knot can have without breaking a part of any section of it. This is achieved by completing Reidemeister moves to remove any unnecessary crossings.

Section 3: Applications in Biology and Chemistry

DNA

Look at the tangling of DNA in the nucleus. The enzymes that interact with the DNA can be seen as completing Reidemeister moves that further knot some DNA. They do this by taking the ends of a type of DNA structure called a linear duplex and knotting them, connecting the ends,forming a cyclic duplex DNA:

Figure 7: Cyclic DNA Ribbon

Given the twist of the ribbon, or how much the ribbon twists around its axis, and the writhe, or how much the axis is contorted, we can figure out the linking number by adding the two. If the ribbon is supercoiled, that is to say it has too high of a twist, this twist will lower and compensate for this loss by increasing the writhe, causing the space to get more deformed. An enzyme will cause this increase in twist, causing the reaction.

Synthesis of Knotted Molecules

While knots have been observed in large and intricate biological molecules like DNA, simpler molecules can also form knotted or linked structures. Even when two molecules have the same atoms bonded in the same sequence, their configurations can differ so they show distinct physical/chemical properties due to chirality being non-superimposable mirror images. Look at the following:

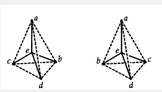


Figure 8:

Although they have the same molecular graph—making them homeomorphic—their mirror imaging prevents them from transforming into each other in 3-D space—denying them the label of isotopic. These pairs are called topological stereoisomers. Chemists have shown a great interest in synthesizing these, because they represent new types of substances. They have developed sophisticated methods to create not only linked molecules but also truly knotted molecular structures. These advances open the door to synthesizing a vast new family of molecular architectures. Chemists use templates to guide the formation of knots, giving way to the first successful synthesis of a molecular trefoil knot.

Chirality of Molecules

A molecular graph in space that can't be deformed through space to its mirror image is called chiral, while a molecular graph in space that can be deformed to its mirror image is called achiral. Knots and links are amphicheiral if they could be deformed to its mirror image. Hence, for knots, amphicheiral and achiral mean the same thing. When chemists search for pairs of topological stereoisomers, they need to know which knots are chiral and which are achiral. A given molecule may be achiral but not chemically achiral. However, a chiral molecule must be chemically chiral. A Möbius ladder has been proven to always be chiral when it has four or more rungs. A Möbius ladder with three rungs however has been shown to be achiral. The first knot to be proven to be chiral was the trefoil knot.

Adjoint Functors and their Applications

Brandon Jensen
University of California Santa Barbara

Mentored by **Jitendra Rathore**
Directed Reading Program (DRP)

Introduction

"Adjoint functors arise everywhere"
- Saunders Mac Lane, *Categories for the Working Mathematician*

Category theory provides a universal language for studying structural similarities between different mathematical objects. Categories were first axiomitized by Eilenberg and Mac Lane in 1945 as an auxiliary definition to make precise the notions of functors and natural transformations [3]. As they put it, *"Category theory starts with the observation that many properties of mathematical systems can be unified and simplified by a presentation with diagrams of arrows"* [2]. Daniel Kan was one of the first to define the notion of an "Adjoint Functor" in 1958, a very special class of functors which are precisely the topic of this poster [3]. After reading this poster, hopefully you will have a better idea of why we care so much about adjoint functors. What problems in mathematics do they help us to solve?

Categories and Functors

A **category** $\mathcal{C}$ consists of two collections, the objects of $\mathcal{C}$ denoted as $Obj(\mathcal{C})$ and the morphisms of $\mathcal{C}$ denoted as $Mor(\mathcal{C})$. These collections include the following data:

1. Every morphism $f \in Mor(\mathcal{C})$ has a specified "source" (domain) and "target" (codomain).
2. Every object $X \in Obj(\mathcal{C})$ has a distinguished identity morphism $1_X : X \to X$.
3. If $f : Y \to Z$ and $g : X \to Y$ are morphisms, then there is a morphism $f \circ g : X \to Z$ called the composite morphism.

In addition, there are two axioms that must be satisfied:

1. UNITALITY: Given any morphism $f : X \to Y$, $f = 1_Y \circ f = f \circ 1_X$.
2. ASSOCIATIVITY: Given three morphisms $h : W \to X$, $g : X \to Y$, and $f : Y \to Z$, we have the equality $f \circ (g \circ h) = (f \circ g) \circ h$.

Definition 1. *A **functor** $F : \mathcal{C} \to \mathcal{D}$ consists of a choice of object $F(c) \in Obj(\mathcal{D})$ for each $c \in Obj(\mathcal{C})$, as well as a choice of morphism $F(g) : F(X) \to F(Y) \in Mor(\mathcal{D})$ for each $g : X \to Y \in Mor(\mathcal{C})$, satisfying the following conditions:*

1. *F preserves composition, i.e. $F(f \circ g) = F(f) \circ F(g)$.*
2. *F preserves identities, i.e. $F(id_X) = id_{F(X)}$ for every object $X \in Obj(\mathcal{C})$.*

Two categories $\mathcal{C}$ and $\mathcal{D}$ are said to be **equivalent** if there exists a functor $F : \mathcal{C} \to \mathcal{D}$ such that F is fully faithful (induced functions on homsets are bijective) and F is essentially surjective (every object in $\mathcal{D}$ is isomorphic to an object in the image of F). In this case we say that F defines an **equivalence of categories**.

What is an Adjoint Functor?

Definition 2. *Given a pair of functors $F : \mathcal{C} \to \mathcal{D}$ and $G : \mathcal{D} \to \mathcal{C}$, we say that F is **left adjoint** to G (denoted $F \dashv G$) if for each pair of objects $X \in Obj(\mathcal{C})$ and $Y \in Obj(\mathcal{D})$ we have a bijection of hom-sets:*

$$Hom_{\mathcal{D}}(F(X), Y) \cong Hom_{\mathcal{C}}(X, G(Y))$$

For those with knowledge in natural transformations, these isomorphisms should assemble into a natural isomorphism (natural in both arguments).

Free and Forgetful Functors

Many adjoint functors tend to fall into a class called "Free/Forgetful Adjunctions". These types of adjunctions are, roughly, defined as follows:

Definition 3. *Suppose we are given an adjoint pair of functors $F \dashv G$ such that $F : \mathcal{C} \to \mathcal{D}$ and $G : \mathcal{D} \to \mathcal{C}$. Informally, we say that this adjunction is a **free-forgetful adjunction** when:*

- *Objects in $\mathcal{D}$ have "more structure" than objects in $\mathcal{C}$.*
- *G is a functor "ignoring" the structure of objects in $\mathcal{D}$, and F endows objects in $\mathcal{C}$ with a free $\mathcal{D}$-structure.*

In many cases, the free functor is left adjoint to the forgetful functor.

As always, a definition would be meaningless without examples. As such, here are some important examples of free-forgetful adjunctions:

1. The Free Group Functor $F(-) : \mathbf{Set} \to \mathbf{Grp}$ is left adjoint to the forgetful functor $U_{\mathbf{Grp}} : \mathbf{Grp} \to \mathbf{Set}$.
2. The Polynomial Ring Functor $\mathbb{Z}[-] : \mathbf{Set} \to \mathbf{CRing}$ is left adjoint to the forgetful functor $U_{\mathbf{CRing}} : \mathbf{CRing} \to \mathbf{Set}$.
3. The Abelianization Functor $(-)^{ab} : \mathbf{Grp} \to \mathbf{Ab}$ is left adjoint to the forgetful functor $U_{\mathbf{Ab}} : \mathbf{Ab} \to \mathbf{Grp}$.
4. The Grothendieck Group Functor $K_0[-] : \mathbf{CMon} \to \mathbf{Ab}$ is left adjoint to the forgetful functor $U_{\mathbf{CMon}} : \mathbf{Ab} \to \mathbf{CMon}$. This functor is used to define the zeroth algebraic K-group $K_0(R)$ of a ring R.

$$\mathbf{CMon} \underset{U_{\mathbf{CMon}}}{\overset{K_0[-]}{\rightleftarrows}} \mathbf{Ab} \underset{U_{\mathbf{Ab}}}{\overset{(-)^{ab}}{\rightleftarrows}} \mathbf{Grp} \underset{U_{\mathbf{Grp}}}{\overset{F(-)}{\rightleftarrows}} \mathbf{Set} \underset{U_{\mathbf{CRing}}}{\overset{\mathbb{Z}[-]}{\rightleftarrows}} \mathbf{CRing}$$

The Optimization Problem

So why do we care about free and forgetful functors? One reason is they allow us to solve certain mapping problems. Let X be a set, and G be a group. Let's say we have a set map $f : X \to U_{\mathbf{Grp}}G$. Now we might wonder, *"How can we extend this set map to a group homomorphism?"*. For simplicity, take $X = \{1\}$ and consider the following diagrams:

$$\begin{array}{ccc} U_{\mathbf{Grp}}(\mathbb{Z}) & \overset{\overline{f}}{\dashrightarrow} & U_{\mathbf{Grp}}(G) \\ {\scriptstyle i}\uparrow & \nearrow_{f} & \\ X & & \end{array} \qquad \begin{array}{ccc} U_{\mathbf{Grp}}(\mathbb{Z}\times\mathbb{Z}) & \overset{\overline{g}}{\dashrightarrow} & U_{\mathbf{Grp}}(G) \\ {\scriptstyle (1,0)}\uparrow & \nearrow_{g} & \\ X & & \end{array}$$

$\overline{f}$ is defined uniquely, since $\mathbb{Z}$ is the free group on one generator. Thus $\overline{f}$ is a "solution" to the problem: how can we extend $i : X \to U_{\mathbf{Grp}}G$ to a group homomorphism $\mathbb{Z} \to G$.

Define $\overline{g}(n, m) = \pm\prod_{i=1}^{n} g(1) \prod_{i=1}^{m} a$ for any $a \in G$ (depending on the sign of n and m). There is no unique group homomorphism from $\mathbb{Z} \times \mathbb{Z} \to G$, yet it is a "solution" to the problem.

Notice that the most efficient solution to this problem is $\mathbb{Z}$, which is exactly $F(X)$, the image of X under the free group functor. The fact that the Free Functor gives us such a solution is an important characteristic of the functor. Later, we will see the **Solution Set Condition**. Instead of one object, we have a set of objects which we can choose to factor through. Without knowing the free functor, these sets might be easier to characterize than finding the "free" object of that category.

What is Preserved by Adjoints?

Another important property of adjoint functors is the structures they preserve:

Preserved by Left Adjoint	Preserved by Right Adjoint
Colimits	*Limits*
1. Coproducts - Direct Sums, Disjoint Union, Free Product, etc.	1. Products - Cartesian Product and Direct Products
2. Coequalizers - Cokernels, Quotients	2. Equalizers - Kernels, Subobjects
3. Pushouts	3. Pullbacks
4. Direct Limits	4. Inverse Limits

This explains why $F(X) * F(Y) \cong F(X \coprod Y)$.

Examples of Adjoint Pairs

1. **Tensor/Hom** - Given $- \otimes_R N : \mathbf{Mod}_R \to \mathbf{Ab}$ and $Hom_{\mathbf{Ab}}(N, -) : \mathbf{Ab} \to \mathbf{Mod}_R$, we have $- \otimes_R N \dashv Hom_{\mathbf{Ab}}(N, -)$. So $- \otimes_R N$ is right exact, and $Hom_{\mathbf{Ab}}(N, -)$ is left exact.
2. **Suspension/Loop-Space** - Given the Suspension Functor $S : \mathbf{hTop} \to \mathbf{hTop}$ and the Loop Space functor $\Omega : \mathbf{hTop} \to \mathbf{hTop}$, we have $S \dashv \Omega$.
3. **Direct Image/Inverse Image** - Let $f : X \to Y$ be a continuous function. Then we have the direct/inverse image functors $f_* : \mathbf{Sh}(X) \to \mathbf{Sh}(Y)$ and $f^{-1} : \mathbf{Sh}(Y) \to \mathbf{Sh}(X)$, and we have $f^{-1} \dashv f_*$. Here we denote $\mathbf{Sh}(-)$ to be the category of sheaves valued in **Ab**.
4. **Group Ring/Group of Units** - The Group of Units Functor $(-)^{\times} : \mathbf{Ring} \to \mathbf{Grp}$ admits a left adjoint, namely the Group Ring Functor $\mathbb{Z}[-] : \mathbf{Grp} \to \mathbf{Ring}$, and we have $\mathbb{Z}[-] \dashv (-)^{\times}$.

Adjoint Functor Theorem

Since adjoint functors preserve limits/colimits, one might wonder: if a functor preserves all limits, does it admit a left adjoint? This leads us to a fundamental result in category theory:

Theorem 1. ***ADJOINT FUNCTOR THEOREM*** *[1] Let $\mathcal{C}$ be locally small and complete. Given any other category $\mathcal{D}$, and a limit preserving functor $U : \mathcal{C} \to \mathcal{D}$, the following are equivalent:*

1. *U has a left Adjoint.*
2. *For each $D \in Obj(\mathcal{D})$, the functor U satisfies the **Solution Set Condition**: There exists a set of objects $(S_i)_{i \in I}$ in $Obj(\mathcal{C})$ such that for any object $C \in \mathcal{C}$ and arrow $f : X \to U(C)$, there exists $i \in I$ and arrows $i : X \to U(S_i)$ and $\overline{f} : S_i \to C$ such that $f = U(\overline{f}) \circ i$*

Notice that $- \otimes_R N$ preserves cokernels, but it does *not* generally preserve kernels. But one might wonder, how far is $- \otimes_R N$ from being left exact? The family of functors, $\mathbf{Tor}^{\mathbf{i}}_{\mathbf{R}}$, answers this question exactly (no pun intended), the derived functors of $- \otimes_R N$. Similarly, $\mathbf{Ext}^{\mathbf{i}}_{\mathbf{R}}(\mathbf{N}, -)$ are the derived functors of $Hom_{\mathbf{Mod}_R}(N, -)$ [4]. These examples have led to the broader study of derived categories and functors.

References

[1] Steve Awodey. *Category Theory*. New York, NY: Oxford University Press Inc., 2010.

[2] Saunders Mac Lane. *Category Theory for the Working Mathematician*. New York, NY: Springer, 1978.

[3] Jean-Pierre Marquis. *Category Theory*. Stanford Encyclopedia of Philosphy, 2019. URL: https://plato.stanford.edu/entries/category-theory/.

[4] Charles Weibel. *An Introduction to Homological Algebra*. New York, NY: Press Syndicate of the University of Cambridge, 1994.

Lie Theory in Physics

Siyu Chen, Mentored by Arthur Jiang
University of California Santa Barbara

Lie Theory in Physics

In physics, Lie theory provides a framework to describe continuous symmetry, or invariant properties of systems under continuous transformations. Such transformations are described by a **Lie group** G, which is a group and a manifold. Here, we restrict discussion to a semi-simple, compact, connected, finite-dimensional Lie group with dimension n.

Lie Group & Lie Algebra

For a manifold M, we generalize the notion of tangent vectors via the following: for $p \in M$, define curves $\gamma : (-\epsilon, \epsilon) \to M, \gamma(0) = p$. We define an equivalence relation on such curves: $\gamma_1 \sim \gamma_2 \Leftrightarrow (\phi \circ \gamma_1)'(0) = (\phi \circ \gamma_2)'(0)$, for some **coordinate chart** $\phi : U \to \mathbb{R}^n$ with $p \in U$ open subset of M. The equivalence class $[\gamma]$ represent a **tangent vector**.
The **tangent space** of p is defined as $T_p(M) \coloneqq \{[\gamma] \mid \gamma(0) = p\}$.
A **Lie algebra** $(\mathfrak{g}, [\cdot, \cdot])$ of a Lie group is the tangent space at identity for G, denoted as $T_e(G)$. It has the following properties:

- If a Lie group G is finite dimensional, then $\dim \mathfrak{g} = \dim G = n$.
- The **Lie bracket** $[\cdot, \cdot]$ is bilinear, anti-symmetric and satisfies Jacobi identity.

$$\begin{aligned} &[X, aY + bZ] = a[X, Y] + b[X, Z], \quad [X, Y] = -[Y, X], \\ &[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0 \end{aligned} \tag{1}$$

In matrix representations, we will use the commutator $[A, B] = AB - BA$.

Exponential Map & Generators

In physical context, we view Lie algebra elements as infinitesimals that exponentiate to Lie group elements. For example, in the defining representation of $SO(2)$, the 2-by-2 rotation matrices, $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathfrak{so}(2)$ is an element of the Lie algebra. We exponentiate this matrix with $\theta \in \mathbb{R}$ to recover rotation by θ.

$$e^{\theta A} = \sum_{n=0}^{\infty} \frac{(\theta A)^n}{n!} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \tag{2}$$

We can formalize this exponentiation into **exponential maps**. Define:

$$\exp : \mathfrak{g} \to G, \quad \exp(X) = \gamma_X(1) \tag{3}$$

where $\gamma_X : [0, 1] \to G$ is a curve called **one-parameter subgroup** such that

$$\gamma_X(0) = e, \quad \gamma_X(s)\gamma_X(t) = \gamma_X(s + t), \quad \frac{\mathrm{d}}{\mathrm{d}t}\gamma_X(t)\big|_{t=0} = X. \tag{4}$$

The exponential map has following properties:

- $\exp(s + t)X = \exp sX \exp tX, \quad (\exp X)^{-1} = \exp -X, \quad (\exp X)^n = \exp nX$
- For compact, connected Lie groups, $\exp$ is surjective. $\exp(\mathfrak{g}) = G$.

We define the set of **generators** $\{X^a\}_{a=1,2,\ldots,n} \subset \mathfrak{g}$ as a basis of the Lie algebra that satisfies the following:

$$[X^a, X^b] = i \sum_{c=1}^{n} f^{ab}{}_c X^c \tag{5}$$

The $f^{ab}{}_c$ are the **structure constants**. They are anti-symmtric in the first two indices, $f^{ab}{}_c = -f^{ba}{}_c$ from anti-symmetry of Lie bracket.. The presence of i is a physics convention resulted from $X^a \coloneqq i\tilde{X}^a$ where $\tilde{X}^a$ are the actual generators. We then exponentiate iX^a instead of $\tilde{X}^a$. This imposes X^a to be Hermitian, $(X^a)^\dagger = X^a$, instead of anti-Hermitian, $(X^a)^\dagger = -X^a$ in the math convention.

Adjoint Representation & Cartan-Killing Form

A **Lie group representation** $\rho : G \to \mathrm{GL}(V)$ induces a **Lie algebra representation**, which is a **Lie algebra homomorphism**. For notation, we may omit dρ map: $Xv \coloneqq d\rho(X)v$.

$$d\rho : \mathfrak{g} \to \mathcal{L}(V), \quad d\rho([X, Y]) = [d\rho(X), d\rho(Y)] \tag{6}$$

We introduce the **adjoint representation** of a Lie algebra. $Ad : \mathfrak{g} \to \mathcal{L}(\mathfrak{g})$ s.t

$$Ad_{X^a}(X^b) \coloneqq [X^a, X^b] = i \sum_{c=1}^{n} f^{ab}{}_c X^c \tag{7}$$

We may have this representation in matrix form. Define $\{T^a\}_{a=1,2,\ldots,n} \subset \mathrm{GL}(n, \mathbb{R})$.

$$(T^a)^b_c \coloneqq -if^{ab}{}_c \quad \Rightarrow \quad [T^a, T^b] = i \sum_{c=1}^{n} f^{ab}{}_c T^c \tag{8}$$

Here we have n matrices indexed by a that are n-by-n indexed by b, c. The right side of eq (8) concludes that T^a indeed represent X^a in adjoint representation.
We define the **Cartan-Killing form** to use index notations. Define the bilinear form

$$g^{ab} \coloneqq \mathrm{tr}\Big(\mathrm{ad}(X^a)\,\mathrm{ad}(X^b)\Big) = \mathrm{tr}\Big(T^a T^b\Big) \tag{9}$$

By the Cartan criterion, semi-simple Lie groups have isomoprhic dual spaces by this metric. I will use full index notation for rest of the poster. Contraction of upper and lower indices indicates a sum. The δ^a_c is the Kronecker delta which gives 1 if $a = c$ and 0 otherwise.

$$X_a Y^a \coloneqq \sum_{a=1}^{n} X_a Y^a, \quad g^{ab} g_{bc} = \delta^a_c, \quad X_a = g_{ab} X^b \tag{10}$$

Root Spaces & Weights

From $\{X^a\}$ we select the maximal commuting subset $\{H^i\}, i = 1, 2, \ldots l$, known as the **Cartan subalgebra** $\mathfrak{h}$ with **rank** l, each represented by commuting matrices $\{T^i\}$. We simultaneously diagonalize them with real diagonal elements $-\beta^i(a)$, a is the column number.

$$(T^i)^a_b = -\mathrm{diag}(\beta^i(1), \beta^i(2), \ldots, \beta^i(a), \ldots, \beta^i(n)) = -\beta^i \delta^a_b \tag{11}$$

Notably, while X^a were Hermitian, under the diagonalization of H^i, the off diagonal elements become complex linear combination of the Hermitian X^a, and is no longer Hermitian. We call them E^a. Acting H^i on E^a, we have

$$\mathrm{ad}_{H^i}(E^a) = [H^i, E^a] = if^{ia}{}_b E^b = -(T^i)^a{}_b E^b = \beta^i(a)\delta^a_b E^b = \beta^i(a) E^a \tag{12}$$

This shows that each E^a is an eigenvector of H^i with real eigenvalue β^i. Then we identify each E^a as **root vectors** E_β with **root** $\vec{\beta} \in \mathbb{R}^l$ as the following:

$$\mathrm{ad}_{H^i}(E_\beta) = [H^i, E_\beta] = \beta^i E_\beta, \quad \vec{\beta}(a) = (\beta^1(a), \beta^2(a), \ldots, \beta^l(a)) \tag{13}$$

Taking the complex conjugate of eq (12), we show $E^\dagger_\beta = E_{-\beta}$ and split root vectors into positive and negative pairs. This implies that $n - l$ is even, so n, l must have same parity.

$$\mathrm{ad}_{H^i}(E^\dagger_\beta) = [H^i, E^\dagger_\beta] = -\beta^i E^\dagger_\beta \tag{14}$$

We now move away from adjoint to general representation $d\rho : \mathfrak{g} \to \mathcal{L}(V)$. Elements of $\mathfrak{h}$ share eigenvectors: $\{v_w \in V \mid H^i v_w = w^i v_w, \forall H^i \in \mathfrak{h}\}$, named **weight vectors**.

$$w : \mathfrak{h} \to \mathbb{R}, \quad w(H^i) = w^i \tag{15}$$

And above, we define the **weight** is the map that takes elements of $\mathfrak{h}$ to their eigenvalue when acting on eigenvector v_w. Notably, root vectors connect weights by the following:

$$H^i \cdot E_\beta \cdot v_w = E_\beta \cdot (H^i \cdot v_w) + [H^i, E_\beta] \cdot v_w = (w + \beta)^i E_\beta \cdot v_w \tag{16}$$

By these properties, Cartan generators are often associated with physical observables.

Examples in Physics

Lie groups $SU(2), SU(3)$ are great examples of Lie theory in physics.
SU(N). The **defining representation** of $SU(N)$ is the N-by-N special unitary matrix.

$$\rho : SU(N) \to \mathrm{GL}(N, \mathbb{C}), \quad \det(U) = 1, \quad U^\dagger U = UU^\dagger = I \tag{17}$$

Plug in the speical and unitary conditions to exponential, we conclude that the generators are N-by-N traceless Hermitian matrices.
SU(2). In the defining representation of $SU(2)$, we have the generators:

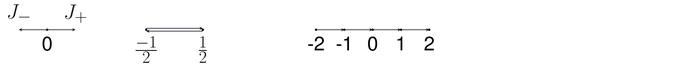
$$J^x = \frac{\sigma^x}{2} = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad J^y = \frac{\sigma^y}{2} = \frac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad J^z = \frac{\sigma^z}{2} = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{18}$$

The σ^μ are Pauli matrices, $[J^i, J^j] = \epsilon^{ij}{}_k J^k$, where ϵ is the totally antisymmetric Levi-Civita symbol. J^z is the only Cartan generator and it is associated with spin.

$$J_\pm = J^x \pm iJ^y, \quad [J_z, J_\pm] = \pm J_\pm \tag{19}$$

$J_\pm$ are the $E_{\pm\beta}$ root vectors, with roots of ± 1. Physically, any weight diagram of $SU(2)$ into representation $SU(N)$ describe particle of spin $j = (N - 1)/2$.

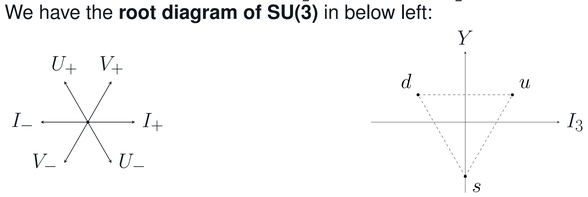
Root diagram **Spin $\frac{1}{2}$ (Defining rep)** **Spin 2 (SU(5))**



Above, we have the defining rep $SU(2), j = \frac{1}{2}$ and $SU(2) \to SU(5), j = 2$.
SU(3) represent the strong interaction in the Standard Model. The generators in the defining rep are the 8 traceless Hermitian 3-by-3 Gell-Mann matrices.

$$\begin{aligned} &\lambda^1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda^2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda^3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda^4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \\ &\lambda^5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \quad \lambda^6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \lambda^7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad \lambda^8 = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \end{aligned} \tag{20}$$

The diagonal λ_3, λ_8 are the 2 Cartan generators associated with I_3, Y, isospin and hypercharge in Quantum Chromodynamics. Rest of the 6 root vectors divides into 3 pairs. Those are $I_\pm = \frac{1}{2}(\lambda^1 \pm i\lambda^2), U_\pm = \frac{1}{2}(\lambda^6 \pm i\lambda^7), V_\pm = \frac{1}{2}(\lambda_4 \pm \lambda_5)$.
We have the **root diagram of SU(3)** in below left:



In above right, the **weight diagram of defining representation of** $SU(3)$, the weights correspond to up, down and strange quarks.
Other interesting Lie groups to study in physics include $U(1)$ for electric charge, the (not connected) Lorentz group $SO(3, 1)$ in special relativity.

Acknowledgements & References

I would like to thank my mentor Arthur Jiang for his guidance, my supportive DRP peer Bryan, Professor Zee for his lectures and textbooks on group theory.
John M Lee, *Introduction to Differentiable Manifolds, 2nd Edition*.
Anthony Zee, *Group Theory in a Nutshell for Physicists*.

TYING TOGETHER SKEINS AND CHARACTERS

Cole Ellison
University of California Santa Barbara

Skein Modules and Skein Algebras

Classically, knot theory is concerned with embeddings of S^1 into either $\mathbb{R}^3$ or S^3. Here, we can compute algebraic invariants such as knot polynomials by working with projections onto the plane, where all crossings can be resolved. However, if we are willing to forgo the necessity of such a global projection, we can consider embeddings into arbitrary 3-manifolds, instead resolving tangles locally by taking projections in neighborhoods diffeomorphic to $\mathbb{R}^3$. When our manifold has interesting topology, we may end up with untangled yet homotopically nontrivial knots and link components.

Let M be an oriented 3-manifold, $\mathcal{L}^{fr}$ be the set of unoriented framed links (including the empty link $\varnothing$) in M up to ambient isotopy, R a commutative ring with unity, $A \in R^\times$, and $R\mathcal{L}^{fr}$ the free R-module with basis $\mathcal{L}^{fr}$. Let $S^{sub}_{2,\infty} \subseteq R\mathcal{L}^{fr}$ be the submodule

$$S^{sub}_{2,\infty} = \langle \underbrace{L_+ - AL_0 - A^{-1}L_\infty}_{\text{(skein relation)}}, \underbrace{L \sqcup \bigcirc + (A^2 + A^{-2})L}_{\text{(trivial component relation)}} \rangle$$

where $\bigcirc$ denotes the trivial framed knot and L_+, L_0, L_∞ are identical except in a small 3-ball in M where they differ as follows:

L_+ L_0 L_∞

The **Kauffman Bracket Skein Module** (KBSM) of M is

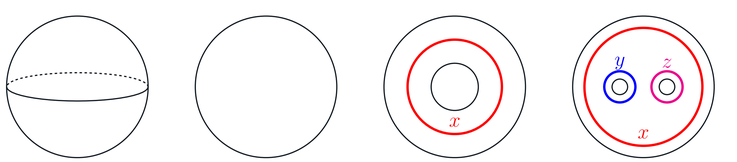
$$\mathcal{S}_{2,\infty}(M; R, A) = R\mathcal{L}^{fr}/S^{sub}_{2,\infty}.$$

When $M = \mathbb{R}^3$ or $M = S^3$ and $R = \mathbb{Z}[A^{\pm 1}]$, $\mathcal{S}_{2,\infty}(M) = R\{\varnothing\}$, which descends canonically to the Kauffman bracket polynomial by looking at the coefficient on $\varnothing$. Consider the special case in which $M = \Sigma \times I$, for some surface Σ. It turns out that in this case, we can endow the KBSM with a multiplication to get a **Kauffman Bracket Skein Algebra** (KBSA); given framed links $L_1, L_2 \subseteq \Sigma \times I$, define $L_1 \cdot L_2$ by placing L_1 over L_2, where $L_1 \subseteq \Sigma \times (1/2, 1), L_2 \subseteq \Sigma \times (0, 1/2)$. We write the KBSA of a thickened surface as $\mathcal{S}^{alg}(\Sigma; R, A)$. More generally, when $A = \pm 1$, any KBSM admits an algebra structure with multiplication as the disjoint sum.

Manifolds of the form $\Sigma \times I$ provide a large class of well behaved yet nontrivial KBSMs, which we can compute using the following theorem.

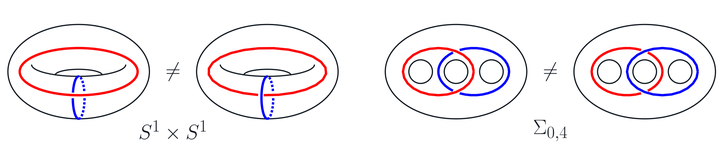
Theorem (Przytycki): Let M be an oriented 3-manifold which is either equal to $\Sigma \times I$, where Σ is an oriented surface, or equal to a twisted I bundle over Σ, $\Sigma \tilde{\times} I$, where Σ is an unoriented surface. Then the KBSM $\mathcal{S}_{2,\infty}(M; \mathbb{Z}[A^{\pm 1}]) = \mathbb{Z}[A^{\pm 1}]B(\Sigma)$, where $B(\Sigma)$ consists of $\varnothing$ and links in Σ without contractible components.

Example:



$\mathcal{S}^{alg}(S^2) \cong \mathbb{Z}[A^{\pm 1}]$ $\mathcal{S}^{alg}(D^2) \cong \mathbb{Z}[A^{\pm 1}]$ $\mathcal{S}^{alg}(S^1 \times I) \cong \mathbb{Z}[A^{\pm 1}][x]$ $\mathcal{S}^{alg}(\Sigma_{0,3}) \cong \mathbb{Z}[A^{\pm 1}][x, y, z]$

Figure 1: The four commutative KBSAs of $\Sigma \times I$ over $\mathbb{Z}[A^{\pm 1}]$ with their generators.



$S^1 \times S^1$ $\Sigma_{0,4}$

Figure 2: Noncommutative multiplications in the torus and the four-punctured sphere.

Zariski Topology

Let R be a ring. Let S be the set of all prime ideals of R. For each subset (equivalently, we can consider ideals) $I \subseteq R$, let $V(I)$ denote the set of all prime ideals of R which contain I. We have the following properties:

$$V(0) = X, \quad V(R) = \emptyset, \quad V\left(\sum_{j \in J} I_j\right) = \bigcap_{j \in J} V(I_j), \quad V(I_1 \cap I_2) = V(I_1) \cup V(I_2).$$

Thus, the sets $V(I)$ satisfy the axioms for the closed sets of a topological space, whose points are the elements of S. Such a space is called $\text{Spec}(R)$, and the topology is called the **Zariski topology**.

A set $X \subseteq \mathbb{C}[x_1, \ldots, x_m]$ which is closed under the Zariski topology is called a **closed algebraic set**. Given an algebraic set X, let $I(X)$ be the ideal of polynomials which vanish on X. Then the **coordinate ring** of X is $\mathbb{C}[x_1, \ldots, x_m]/I(X)$.

Let M be a compact orientable manifold. An $SL_2(\mathbb{C})$ **representation** of the fundamental group $\pi_1(M)$ is a group homomorphism $\rho : \pi_1(M) \to SL_2(\mathbb{C})$. Then the **character** of the representation is $\chi_\rho = \text{tr} \circ \rho$.

Culler-Shalen Theory

Denote the set of all characters by $X(M)$. To each homotopy class of curves $\gamma \in \pi_1(M)$, we can associate a function $t_\gamma : X(M) \to \mathbb{C}$, where $t_\gamma(\chi_\rho) = \chi_\rho(\gamma)$. Together, Marc Culler (a UCSB alumni!) and Peter Shalen proved the following two results:

Theorem (Culler-Shalen):

(i) There exists a finite set of elements $\{\gamma_1, \ldots, \gamma_m\}$ in $\pi_1(M)$ such that every t_γ is an element of the polynomial ring $\mathbb{C}[t_{\gamma_1}, \ldots, t_{\gamma_m}]$.

(ii) If every t_γ is an element of $\mathbb{C}[t_{\gamma_1}, \ldots, t_{\gamma_m}]$, then $X(M)$ is a closed algebraic subset of $\mathbb{C}[t_{\gamma_1}, \ldots, t_{\gamma_m}]$.

Let $\mathcal{R}(M)$ be the coordinate ring of $X(M)$. While we could make different choices of coordinates, all choices are isomorphic by polynomial maps. Moreover, $\mathcal{R}(M)$ makes sense as a subset of the algebra $\mathbb{C}^{X(M)}$ since choice of representative produces the same function, as we are quotienting out by functions which are identically zero.

Main Theorem

Now, we wish to associate to each knot K an element of $\pi_1(K)$. A priori, this may depend on the orientation of K. Fix an orientation, and pick a curve γ freely homotopic to the now oriented $\vec{K}$. Trace is invariant under conjugation, so "'freely" did not matter, and for any character χ_ρ,

$$\chi_\rho(\gamma^{-1}) = \text{tr}(\rho(\gamma^{-1})) = \text{tr}(\rho(\gamma)^{-1}) = \frac{\text{tr}(\rho(\gamma))}{\det(\rho(\gamma))} = \text{tr}(\rho(\gamma)) = \chi_\rho(\gamma).$$

Thus, our choice of orientation also did not matter, so we can speak of *the* map t_γ determined by K.

Theorem (Bullock): Let $\tilde{\Phi} : \mathbb{C}\mathcal{L}^{fr} \to \mathcal{C}^{X(M)}$ be the linear map mapping each knot K to a function $\tilde{\Phi}(K) : X(M) \to \mathbb{C}$, $\tilde{\Phi}(K)(\chi_\rho) = -\chi_\rho(K)$, links to the product of the images of their components, and $\varnothing$ to 1. Then $\tilde{\Phi}$ descends to a well defined surjective algebra homomorphism $\Phi : \mathcal{S}^{alg}(M; \mathbb{C}, -1) \to \mathcal{R}(M)$, with $\ker(\Phi) = \text{nil}(\mathcal{S}^{alg}(M; \mathbb{C}, -1))$.

Since M is compact and orientable, there exist finitely many generators $K_1, \ldots, K_m$. Then $\text{im}(\Phi) \subseteq \mathbb{C}[-\Phi(K_1), \ldots, -\Phi(K_m)]$, so such $-\Phi(K_i)$ are coordinates on $X(M)$, handling surjectivity. Now, to show that Φ is well defined on the KBSA, we must show $\tilde{\Phi}(S^{sub}_{2,\infty}) = 0$. By direct computation, we have that the image of the trivial component relation is

$$\tilde{\Phi}(L \sqcup \bigcirc + ((-1)^2 + (-1)^{-2})L) = \tilde{\Phi}(L)\tilde{\Phi}(\bigcirc + 2\varnothing) = \tilde{\Phi}(L)(-\chi_\rho(\bigcirc) + 2) = \tilde{\Phi}(L)(-\text{tr}(\text{Id}) + 2) = 0.$$

Note that $\rho(\bigcirc) = \text{Id}$ since $\bigcirc$ is nullhomotopic. To compute the image of the skein relation, we need slightly more machinery, which we will now introduce.

Skein Relation to Trace Identity

For the skein relation, it suffices to show resolutions of skein triples (L, L_0, L_∞) in which L (and by choosing the right orientation, L_0) are knots are mapped to 0. In this case, L_∞ will have two components K_1, K_2. Choose $*$ to be a basepoint in the neighborhood where the skein triple differs. We can then find $a, b \in \pi_1(M, *)$ where ab is freely homotopic to $\vec{L}$ (recall, our choice of orientation is inconsequential). Choosing the appropriate orientations $\vec{K_1}, \vec{K_2}$, we have $a\bar{b} \simeq \vec{L_0}, a \simeq \vec{K_1}, b \simeq \vec{K_2}$.

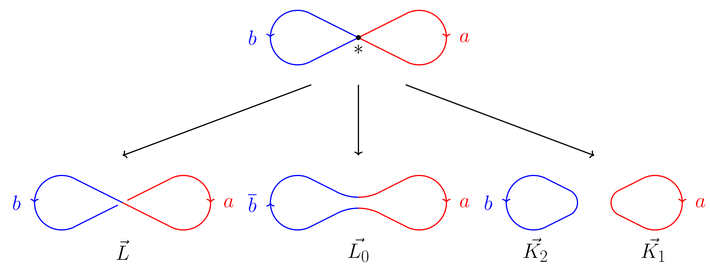


Figure 3: Free homotopies between the skein triple and words in $\pi_1(M, *)$.

For any χ_ρ, let $A = \rho(a), B = \rho(b)$, and we have

$$\begin{aligned}\tilde{\Phi}(L + L_0 + L_\infty)(\chi_\rho) &= -\chi_\rho(L) - \chi_\rho(L_0) + \chi_\rho(K_1)\chi_\rho(K_2) \\ &= -\text{tr}(AB) - \text{tr}(AB^{-1}) + \text{tr}(A)\text{tr}(B).\end{aligned}$$

Our skein relation maps to the **fundamental** $SL_2(\mathbb{C})$ **trace identity**: Let $A, B \in SL_2(\mathbb{C})$. The Cayley-Hamilton theorem gives

$$A^2 - \text{tr}(A)A + \det(A)\text{Id} = 0 \implies A - \text{tr}(A) + A^{-1} = 0.$$

Multiplying both sides by B and taking the trace,

$$AB + A^{-1}B = \text{tr}(A)B \implies \text{tr}(AB) + \text{tr}(A^{-1}B) = \text{tr}(A)\text{tr}(B).$$

Thus, $\tilde{\Phi}(L + L_0 + L_\infty)(\chi_\rho) = 0$. Therefore, $\tilde{\Phi}(S^{sub}_{2,\infty}) = 0$, so $\tilde{\Phi}$ descends to Φ. A powerful result arose three years later, refining Bullock's work:

Theorem (Przytycki-Sikora): Let Σ be an oriented surface, and $M = \Sigma \times I$. If R has no divisors, $\mathcal{S}^{alg}(M; R, A)$ has no divisors.

Here, $\text{nil}(\mathbb{C}) = 0 \implies \text{nil}(\mathcal{S}^{alg}(M; \mathbb{C}, -1) = 0 \implies \mathcal{S}^{alg}(M; \mathbb{C}, -1) \cong \mathcal{R}(M)$.

Acknowledgements

Thank you to my mentor, Rhea Palak Bakshi, for sharing her expertise and enthusiasm in knot theory, and to the DRP for the opportunity to participate.

References

[1] M. F. Atiyah, I. G. Macdonald, Introduction to Commutative Algebra, *Addison-Wesley Publishing Company* (1969).

[2] D. Bullock, Rings of $SL_2(\mathbb{C})$-characters and the Kauffman bracket skein module, *Comment. Math. Helv.* 72 (1997), no. 4, 521-542.

[3] J. H. Przytycki, Fundamentals of Kauffman bracket skein modules, *Kobe Math. J.*, 16(1), (1999), 45-66. arXiv:math/9809113 [math.GT].

[4] J. H. Przytycki, R. P. Bakshi, D. Ibarra, G. Montoya-Vega, D. Weeks, Lectures in Knot Theory, An Exploration of Contemporary Topics, *Springer Universitext* (2024).

The Word Problem in Coxeter Groups

Benjamin Schoeb

University of California - Santa Barbara

Cayley graphs

Let G be a group with a set of generators S. The **Cayley graph** of G, denoted $\mathrm{Cay}(G, S)$ is a directed graph with the elements of G as the vertex and edge set $\{(g, gs) : g \in G, s \in S\}$. We call $s \in S$ an **involution** if it has order 2 and the edge (g, gs) is not directed.

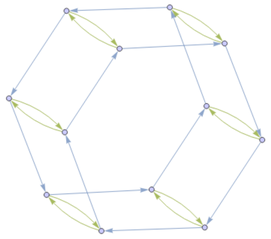


Figure 1: *$\mathrm{Cay}(D_{12}, S)$ with generating set $S = \{r, s\}$, where s is an involution*

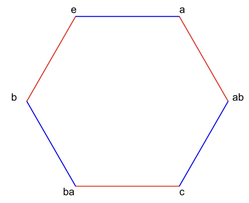


Figure 2: *$\mathrm{Cay}(D_6, S)$ with generating set $S = \{a, b\}$, where a and b are involutions*

Pre-Reflection and Reflection Systems

A **pre-reflection system** of a group G consists of a subset R of G, an action of G on a connected simple graph Ω, and a basepoint $v_0 \in \mathrm{Vert}(\Omega)$ where

1. Every element in R is an involution
2. For $g \in G$ and $r \in R$, $grg^{-1} \in R$
3. For each edge in Ω there exists a unique $r \in R$ that exchanges its endpoints, and each r corresponds to an edge in such a way
4. R generates G

A wall H_r is the set of midpoints of edges which are flipped by r. We can use walls to define a group that intuitively resembles the geometric notion of "reflection".
A **reflection system** is a pre-reflection system with the additional condition:

1. For each $r \in R$, $\Omega \setminus H_r$ has exactly two components.

Example: D_6 is a reflection system

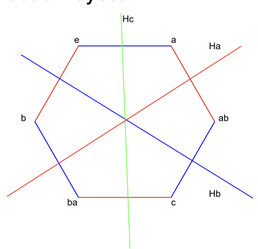


Figure 3:
$\mathrm{Cay}(D_6, S)$,$R = \{a, b, c\}$ and walls H_a, H_b, and H_c

Coxeter Systems

A **Coxeter System**, (W, S), is a group with presentation

$$W = \langle S | (s_i s_j)^{m_{ij}} \ \forall i, j \in I \rangle$$

were $s_i^2 = 1 \, \forall i \in I$ and $m_{ij} \in \mathbb{N} \cup \infty$.
Example: A right-angled Coxeter group has a presentation

$$W = \langle S | s_i^2 \ \forall i \in I \rangle$$

Example: (D_6, S) with generating set $S = \{a, b\}$ with presentation

$$W = \langle a, b | a^2, b^2, (ab)^3 \ \forall i \in I \rangle$$

We can connect Coxeter groups to reflection groups and its Cayley graph by the following lemma:

Lemma 0.1. *For Coxeter system (W, S), let $\Omega = \mathrm{Cay}(W, S)$ and $R = \{wsw^{-1} : w \in W, s \in S\}$. Then (Ω, R) is a pre-reflection system.*

The Word Problem

Given a group G with generating set S, a **word** $\mathbf{s}$ is a sequence consisting of the generators in S. Words can be visualized as a path on the Cayley graph of G beginning at the identity and ending at the product of the generators that make up the word. The **word length** of $g \in G$ with respect to S is

$$l_S(g) = \min\{n \in \mathbb{N} : \exists s_1, ..., s_n \in S \text{ where } g = s_1 ... s_n\}.$$

A word $\mathbf{s} = (s_1, ..., s_n)$ is called **reduced** if $l_S(g) = n$ and $g = s_1 ... s_n$.
Example: In D_6, as defined above, an example of a word would be

$$\mathbf{s} = (a, b, a, b, a, b, a, b, a, b, b, a)$$

where

$$l_S(abababababba) = 1$$

The word problem: Consider a group G with generating set S. Given any words $\mathbf{s}_1, \mathbf{s}_2$, we want to be able to determine if they represent the same element of G. In many groups, there is no algorithm that can make this determination. However, for Coxeter groups, their structure and relationship to their Cayley graphs provide strategies for minimizing words in these groups.

Theorem 0.2. *Assume W is generated by a set of distinct involutions S. Then the following are equivalent*

1. *(W,S) is a Coxeter system*
2. *if $\Omega = \mathrm{Cay}(W, S)$ and $R = \{wsw^{-1} : w \in W, s \in S\}$ then (Ω, R) is a reflection system*
3. *(W,S) satisfies the exchange condition*
4. *(W,S) satisfies the deletion condition*

The Exchange-Deletion Condition

The exchange condition: If $s_1 ... s_k = w \in W$ then for any $s \in S$, $l_S(sw) = k + 1$ or

$$w = (s s_1 ... \hat{s_i} ... s_k)$$

where $\hat{s_i}$ is deleted.
This is a consequence of the deletion condition.
The deletion condition: Assume $w = s_1 ... s_k$. If $\mathbf{s} = (s_1, ..., s_k)$ is a word in S with $K > l_S(s_1 ... s_k)$ then there exist indices $i < j$ such that

$$\mathbf{s} = (s_1, ..., \hat{s_i}, ..., \hat{s_j}, ... s_k)$$

The deletion condition is derived from the action of $r_i = r_j$ over its wall. This allows the vertices w_i and w_{j-1} to be omitted from the path while preserving the types of edges, deleting two edges and therefore two elements from the sequence.

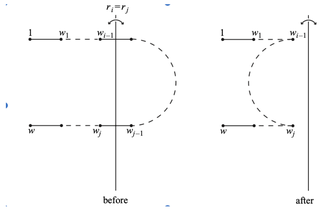


Figure 4: *The deletion condition*

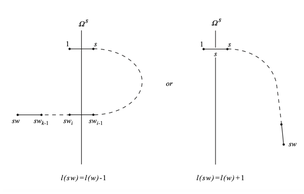


Figure 5: *The exchange condition*

Tits' solution

A word is **M-reduced** if it can't be shortened by the following operations

1. Delete a subword of the form (s, s)
2. Replace an alternating subword of the form $(s, t, ...)$ of length m_{st} by the alternating subword of the form $(t, s, ...)$ of length m_{st}

These operations follow from the structure of Coxeter groups. Following from **Theorem 0.2** we have

Theorem 0.3. *Assume W is generated by a set of distinct involutions S and Theorem (0.2) applies. Then a word in S is reduced if and only if it is M-reduced.*

Acknowledgements

Reference Material:
"The Geometry and Topology of Coxeter Groups" – Michael Davis
"Geometric and Topological Aspects of Coxeter Groups and Buildings" – Anne Thomas
"Office Hours with a Geometric Group Theorist" – Matt Clay, Dan Margalit

Thank you to the UCSB Directed Reading Program and to my mentor Alfredo Ramirez for all the guidance.

Fusion Categories and Jellyfish Relations

Tom Lindquist with mentor Quinn Kolt

Department of Mathematics, University of California, Santa Barbara

Monoidal Categories

A *monoidal category* $\mathcal{C}$ is a category equipped with a multiplication which is unital and associative, up to specified isomorphisms. This means we have a unit object $\mathbf{1}\colon \mathcal{C}$, and a multiplication functor $\otimes\colon \mathcal{C}\times\mathcal{C}\to\mathcal{C}$. Since the multiplication is not "strictly" unital and associative, we have to specify isomorphisms:

- The *left unitor* $\lambda_A\colon \mathbf{1}\otimes A\to A$.
- The *right unitor* $\rho_A\colon A\otimes\mathbf{1}\to A$.
- The *associator* $\alpha_{A,B,C}\colon (A\otimes B)\otimes C\to A\otimes(B\otimes C)$.

We don't want to be able to derive a map $(A\otimes B)\otimes C\to A\otimes(B\otimes C)$ other than the associator by composing associators and unitors. It suffices for the following diagrams to commute (indices on the associators and unitors omitted for space):

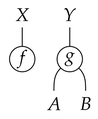
$$\begin{array}{ccc}
((A\otimes B)\otimes C)\otimes D & \xrightarrow{\alpha\otimes D} & (A\otimes(B\otimes C))\otimes D \\
\downarrow{\scriptstyle\alpha} & & \downarrow{\scriptstyle\alpha} \\
(A\otimes B)\otimes(C\otimes D) & & \\
\downarrow{\scriptstyle\alpha} & & \\
A\otimes(B\otimes(C\otimes D)) & \xleftarrow{A\otimes\alpha} & A\otimes((B\otimes C)\otimes D)
\end{array}
\qquad
\begin{array}{ccc}
(A\otimes\mathbf{1})\otimes B & \xrightarrow{\alpha} & A\otimes(\mathbf{1}\otimes B) \\
& {\scriptstyle\rho\otimes B}\searrow \quad \swarrow{\scriptstyle A\otimes\lambda} & \\
& A\otimes B &
\end{array}$$

Some monoidal categories which are particularly relevant to us:

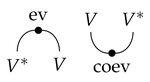
- **Vec**, the category of finite-dimensional complex vector spaces, with $\mathbf{1}=\mathbb{C}$ the one dimensional space, and $\otimes$ the tensor product.
- **Rep**(G), the category of finite-dimensional complex representations of the finite group G, again with $\mathbf{1}=\mathbb{C}$, the trivial representation, and $\otimes$ the tensor product of representations.
- **Vec**(G) the category of finite-dimensional G-graded vector spaces. The unit has one dimension, graded in the group identity, and the tensor product is given by $(V\otimes W)_g=\bigoplus_h V_h\otimes W_{h^{-1}g}$.

String Diagrams

We can visualize morphisms in a monoidal category in a nice way using *string diagrams*. An example string diagram is drawn on the left. Each object in the category is represented by a string, and horizontal juxtaposition of strings represents the monoidal product. The monoidal unit is sometimes written with a dotted string, but more often the strings representing unit objects are simply omitted. Similarly, we don't draw the associator explicitly. We represent morphisms with little labeled circles or boxes. So, the example diagram represents $f\otimes g\colon A\otimes B\to X\otimes Y$, where $f\colon \mathbf{1}\to X$ and $g\colon A\otimes B\to Y$.



Rigidity



In the category of finite-dimensional vector spaces, every vector space V has a *dual space*, written V^*, consisting of the linear functionals on V. Furthermore, we have a linear map $\mathrm{ev}\colon V^*\otimes V\to\mathbf{1}$ which evaluates a linear functional at a vector, and another map $\mathrm{coev}\colon \mathbf{1}\to V\otimes V^*$, which is dual to evaluation. These morphisms are drawn as string diagrams on the left. The evaluation and coevaluation maps satisfy the relations below. This makes **Vec** an example of a *rigid monoidal category*.

Simple Objects

Simple groups are common objects of study in group theory. Similarly, we can talk about simple objects in an arbitrary category, which are objects with no nontrivial proper quotients. The categories we care about on this poster are abelian categories enriched over **Vec**, which essentially means we have a direct sum operation on our objects, and the Hom sets form complex vector spaces with bilinear composition. In this context, an important theorem about simple objects is *Schur's lemma*:

Theorem 1. *If A and B are simple objects in an abelian category enriched over* **Vec**, *then either A and B are not isomorphic and* $\mathrm{Hom}(A,B)=0$, *or A and B are isomorphic, and* $\mathrm{Hom}(A,B)$ *is one-dimensional. In particular,* $\mathrm{Hom}(A,A)=\mathbb{C}$.

An important concept related to simple objects is *semisimplicity*. A category is semisimple if every object can be decomposed into a finite direct sum of simple objects.

Fusion Categories

A fusion category is a rigid linear monoidal category which:

- is semisimple,
- has finitely many simple objects, up to isomorphism,
- and has a simple monoidal unit $\mathbf{1}$.

The *rank* of a fusion category is the number of simple objects it has, up to isomorphism. All the examples of monoidal categories to the left are in fact fusion categories:

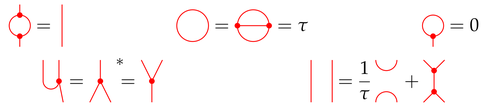
- In **Vec**, the only simple object is the one-dimensional vector space; every vector space can be decomposed into a direct sum of one-dimensional spaces. This is the only fusion category with exactly one simple object.
- In **Rep**(G), the simple objects are the irreducible representations.
- In **Vec**(G), there is a simple object for each element g of G, which is a one-dimensional vector space in the g-graded component and zero dimensional in all the other components.

Examples of small fusion categories

The simplest nontrivial examples of fusion categories we can come up with are ones with only two simple objects, one of which is the monoidal unit.

- One such category is **Vec**$(\mathbb{Z}/2)$, the category of (finite-dimensional) $\mathbb{Z}/2$-graded vector spaces. The two simple objects are $\mathbf{1}$, which is one-dimensional in the 0-graded component, and T, which is one-dimensional in the 1-graded component. We have that $T\otimes T$ must be a direct sum of simple objects, in this case $T\otimes T=\mathbf{1}$. Representing T by a blue strand, we can derive the following string diagram relations:

- Another small fusion category is **Fib**, which has simple objects $\mathbf{1}$ and A. In **Fib**, the decomposition of $A\otimes A$ is $\mathbf{1}\oplus A$. Because the tensor product distributes over direct sums, we can use this to show that the number of summands isomorphic to A in $A^{\otimes n}$ will be the n-th fibonacci number, hence the name **Fib**. Having an isomorphism $A\otimes A\to\mathbf{1}\oplus A$ means that we have maps $A\otimes A\to A$ and $A\to A\otimes A$. We represent A as a red strand, and we represent these maps as trivalent vertices, giving the following relations, where τ is the golden ratio:



We can take linear combinations of morphisms like this because our categories are linear, thus the Hom-sets are vector spaces.

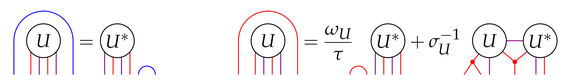
Free Product of Fusion Categories

We are interested in using our simple examples of fusion categories to build more complicated ones. A useful construction is the *free product* of fusion categories. The objects in the free product are formal direct sums of tensor products of objects from the factors. This means that the simple objects in a free product will be products of simple objects of the factors. For example, in the case of **Vec**$(\mathbb{Z}/2)*$**Fib**, the simple objects will be of the form $A, T, A\otimes T, T\otimes A, A\otimes T\otimes A,\cdots$. A string diagram in this category would look like the following:

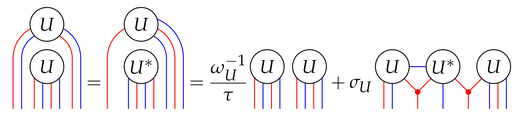
We have a blue diagram from **Vec**$(\mathbb{Z}/2)$ overlaid on a red diagram from **Fib**, such that the strands from one don't overlap the strands from the other. We can apply the relations we derived for the red and blue strings, as long as the strands don't overlap. This free product is still a semisimple rigid monoidal category, but it's no longer a fusion category: it has infinitely many nonisomorphic simple objects. We want to introduce some kind of relation which makes some of these simple objects isomorphic. In fact, in the case of **Vec**$(\mathbb{Z}/2)$ and **Fib**, it must be of the form $U\colon (A\otimes T)^{\otimes n}\to(T\otimes A)^{\otimes n}$. On the right, we see the case $n=2$.

Jellyfish Relations

Once we introduce our isomorphism U, we want the resulting category to be fusion, and so we can use the properties of fusion categories in order to derive relations on the isomorphism U. However, we want to check that these relations are "consistent." We might find that by introducing the isomorphism U, we may have introduced more relations on the simple objects than we intended to.
In order to check this, we derive "jellyfish relations" for our category, pictured below. These relations apply for any n, so we represent $T\otimes(A\otimes T)^{\otimes n-1}$ with a purple strand. The number ω_U is some $2n$-th root of unity, and σ_U is a square root of ω_U.



Then, we use the jellyfish relations in order to "normalize" a particular diagram in two different ways, drawn below in the case $n=1$.



We move the inner U out leftwards until we get a linear combination of diagrams to which no jellyfish relations apply, as shown above. We also move it rightwards to get another such linear combination. Then, we can set these two expressions equal to each other, and see if there are values for ω_U and σ_U which solve the equation. If we can solve the equation without introducing new linear dependencies, then we will know the isomorphism U we added hasn't introduced any more unwanted relations between simple objects.

Sources and Acknowledgements

- M. Izumi, S. Morrison, D. Penneys. Fusion categories between $\mathcal{C}\boxtimes\mathcal{D}$ and $\mathcal{C}*\mathcal{D}$. *https://arxiv.org/abs/1308.5723*

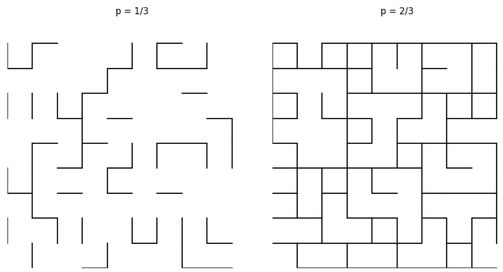
I'd also like to thank Quinn Kolt for giving me the opportunity to work on this project and teaching me a lot of math, and the Directed Reading Program for organizing all this.

Percolation Theory

Shihao Chen, Yikai He, Siyuan (Jeff) Fan
Mentor: Kyle Hansen
University of California, Santa Barbara

What is Percolation?

Percolation is a core mathematical model for studying phase-transition phenomena involving connectivity in random media. In a regular lattice, such as the 2-dimensional $\mathbb{Z}^2$, each edge (**bond percolation**) or node (**site percolation**) is independently set to be"open" or "closed" with probability p. Here, the values of p do matter. The figure below shows samples of a 9×9 square lattice generated by Python by setting $p = \frac{1}{3}$ and $\frac{2}{3}$. Observe that, when $p = \frac{2}{3}$, one cluster extends from top to bottom and from left to right of the sample. That is, there exists **vertical crossing** when $p = \frac{2}{3}$, but no such crossing exists when $p = \frac{1}{3}$.



Why we care: Percolation underpins real-world phenomena like the spread of wildfire. Wildfires that occur frequently cause immense suffering and losses. The devastating wildfire in Los Angeles is a vivid case of percolation in ecological systems. When vegetation is sparse, the spread of fire tends to remain localized. However, as the density of flammable material increases, the chain reaction will trigger a large-scale systematic spread. So, we can model this through percolation on a 2D lattice.

The central question: One fundamental question in percolation theory asks at what probabilities p an infinite connected path exists in the graph. The lower bound of such probability, denoted by p_c, is called the **critical threshold**. For bond percolation on the planar lattice $\mathbb{Z}^2$, the proof of $p_c = \frac{1}{2}$ is one of the field's landmark results.

Harris-Kesten Theorem

The **Harris-Kesten Theorem** is a foundational result in percolation theory for the $\mathbb{Z}^2$ lattice. It explains why the **critical threshold** $p_c = \frac{1}{2}$:

- **Harris (1960)** proved that for $p \leq \frac{1}{2}$, all clusters are almost surely finite, including at $p_c = \frac{1}{2}$.
- **Kesten (1980)** showed that for $p > \frac{1}{2}$, a unique infinite cluster emerges almost surely.

The proof relies on the lattice's **self-duality** (linking open paths in $\mathbb{Z}^2$ to closed paths in its dual), and the construction of **square annulus** in **Harris's Theorem**.

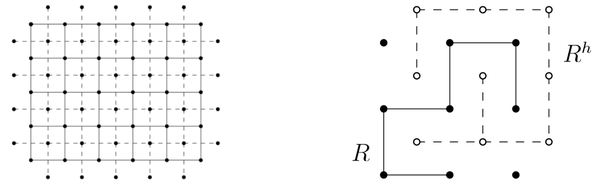
Acknowledgements

We are deeply grateful to our mentor, Kyle Hansen, for his meticulous guidance and we thank UCSB's Directed Reading program for providing us with the opportunity to deeply explore percolation theory and collaborate on this project.

Dual Lattice

In planar bond percolation, we first need to study the probability of **open crossings** in a bounded subset R of the Cayley graph of $\mathbb{Z}^2$, a cornerstone of the **Harris-Kesten Theorem**.

- The **planar dual** of R is R^h, shown with dashed lines in the lefthand figure below.
- **Edges in Dual:** there always exist a vertex $v \in R^h$ corresponding to each face of R, so we define an edge of R^h is **open** if and only if the corresponding edge of R is **closed**, as the righthand figure below shows.



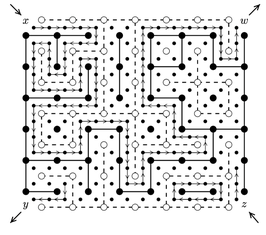
- Define $H(R)$ and $V(R)$ respectively as the events that there exists **horizontal crossing** or **vertical crossing** in R.

Lemma: Self-Duality

Let R be an n by $n+1$ rectangle. Then, exactly one of the events $H(R)$ and $V^*(R^h)$ holds, but not both. If S is an n by n square and $p = \frac{1}{2}$, then

$$\mathbb{P}_{\frac{1}{2}}(H(S)) = \mathbb{P}_{\frac{1}{2}}(V(S)) \geq \frac{1}{2}.$$

We construct a maze-like graph as the figure below shows. Imagine the arrowed edges are paths a player can traverse, and edges of R and R^h are walls. If you enter at x, you will always keep walls of R on the right, and walls of R^h on the left. You either exit at w or at y. If you exit at w, there was always a wall of R on your right, which makes a horizontal crossing of R. Similarly, if you exit at y you get a vertical crossing of R^h. With this heuristic, we can see that one of the events $H(R)$ and $V^*(R^h)$ will hold. Moreover, one can apply the **Jordan Curve Theorem** to show that these events are mutually exclusive.



Now, if R' is a copy of R rotated by $90°$ with probability p that an edge is open in R, we know that $\mathbb{P}_p(H(R)) = \mathbb{P}_p(V(R'))$ and for n by $n+1$ rectangle R,

$$\mathbb{P}_p(H(R)) + \mathbb{P}_{1-p}(V(R')) = 1.$$

If $p = \frac{1}{2}$, then $\mathbb{P}_{\frac{1}{2}}(H(R)) = \frac{1}{2}$ for R. This also implies that, for the smaller n by n square S,

$$\mathbb{P}_{\frac{1}{2}}(H(S)) = \mathbb{P}_{\frac{1}{2}}(V(S)) \geq \frac{1}{2}.$$

References

[1] Stauffer, D., & Aharony, A. (2018). *Introduction to percolation theory*. Taylor & Francis.
[2] Bollobás, B., & Riordan, O. (2006). A short proof of the Harris–Kesten theorem. *Bulletin of the London Mathematical Society*, 38(3), 470–484.

Harris's Theorem: Percolation at $\mathrm{p} = \frac{1}{2}$

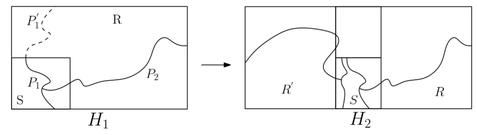
Main Idea: By combining crossing probabilities with geometric decomposition, the proof shows that no infinite open cluster exists at $p_c = \frac{1}{2}$ in $\mathbb{Z}^2$.

Lemma: Exclusive Percolation in Dual Lattice

Let $R = [0, m] \times [0, 2n]$, $m \geq n$, by an m by $2n$ rectangle. Let $X(R)$ be the event that there are paths P_1 and P_2 of open edges, where P_1 crosses $S = [0, n] \times [0, n]$ from top to bottom, and P_2 lies inside R and joins some vertex on P_1 to some vertex on the right-hand side of R. Then

$$\mathbb{P}_p(X(R)) \geq \mathbb{P}_p(H(R))\mathbb{P}_p(V(S))/2.$$

In H_1, the Lemma tells, by conditioning on the existence of the **leftmost vertical crossing** P_1, the probability of finding a connecting path P_2 is **at least half the probability** of a horizontal crossing of R.



In H_2, two overlapping rectangles R and its reflection R' share a square S. Then, we will use $X(R)$ same as in the H_1, $X(R')$, the reflection of $X(R)$, and $H(S)$. Now, based on increasing event lemma and previous lemma, we have:

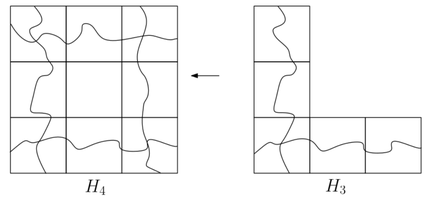
$$\mathbb{P}_{\frac{1}{2}}(H(R \cup R')) \geq \mathbb{P}_{\frac{1}{2}}(X(R))^2 \mathbb{P}_{\frac{1}{2}}(H(S)) = \mathbb{P}_{\frac{1}{2}}(H(R))^2/2^5.$$

Repeatedly applying this to scaled rectangles gives a uniform lower bound for horizontal crossing probabilities in long thin rectangles, we prove:

Corollary

Let $\rho > 1$ be a fixed integer. There is a constant $c(\rho) > 0$ depending only on ρ such that for any $2\rho n$ by $2n$ rectangle R we have

$$\mathbb{P}_{\frac{1}{2}}(H(R)) \geq c(\rho).$$



In bond percolation, $\theta(p)$ is the probability that the origin belongs to an infinite open cluster of connected edges. By the above corollary, long thin rectangles have a probability of being crossed. This idea allows us to build the configuration of two thin rectangles sharing a square to create H_3, and extending this yields H_4, a **square annulus** fully enclosing the center. Each rectangle crossing is independent and occurs with probability $\geq c$, so the full cycle forms with probability $\geq c^4$. Repeating this around the origin with disjoint annuli, the probability of no surrounding cycle in the dual lattice tends to zero. Thus, we have:

Theorem: Harris's Theorem

For bond percolation in $\mathbb{Z}^2$, there is almost certainly no infinite open cluster, showing that

$$\theta\left(\frac{1}{2}\right) = 0.$$

Solving the Schrödinger Equation: The Hydrogen Atom

UC **SANTA BARBARA**
Department of Mathematics

Aileen Arreola and Rebekah Pustelnik

Partial Differential Equations

Partial differential equations (PDEs) are equations that involve partial derivatives of a function with multiple variables. They are often used in physics, chemistry, and engineering to describe how a quantity (temperature, energy, pressure, etc.) changes with respect to space and time.

Ex. Heat Equation

$$\frac{\partial u}{\partial t} = \alpha \frac{\partial^2 u}{\partial x^2}$$

Separation of Variables

Separation of variables is a technique used to solve linear homogeneous PDEs by assuming the solution is a product of functions, each depending on only one variable.

Heat Equation

$$\frac{\partial u}{\partial t} = \alpha \frac{\partial^2 u}{\partial x^2}, \quad u(0,t) = 0, \quad u(L,t) = 0, \quad u(x,0) = f(x)$$

$$u(x,t) = X(x)T(t)$$

$$X(x)T'(t) = \alpha X''(x)T(t)$$

$$\frac{T'(t)}{\alpha T(t)} = \frac{X''(x)}{X(x)} = -\lambda$$

So now we have two equations with respect to one variable:

$$X'' + \lambda X = 0 \quad \text{and} \quad T' + \alpha \lambda T = 0$$

which can be solved as simpler ODEs.

Schrödinger Equation

The **Schrödinger equation** gives rise to the wave function, $\Psi(x,t)$, which explains the state of a quantum system over time. Specifically, it describes the probability of finding a particle in a certain place or state.

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial x^2} + V\Psi \tag{1}$$

$\Psi(x,t)$: wave function with respect to position and time, $\hbar$: reduced Planck constant, m: mass, V: potential energy function

Probability Density Functions

Probability can be modeled using discrete or continuous variables. For a discrete variable X, let $P(j)$ be the probability that X is equal to j. The function $P(j)$ is called the probability mass function of X and satisfies

$$0 \le P(j) \le 1 \quad \text{and} \quad \sum_{j=0}^{\infty} P(j) = 1.$$

Similarly, for a continuous probability distribution, let $P_{a,b}$ be the probability that X lies in $[a,b]$. We define $\rho(x)$ such that

$$P_{a,b} = \int_a^b \rho(x)dx \quad \text{and} \quad \int_{-\infty}^{\infty} \rho(x)dx = 1.$$

We call $\rho(x)$ the **probability density function** (PDF) of X. The solutions to the Schrödinger equation (1) are functions Ψ which describe the quantum state of the system; the wave functions are such that $|\Psi|^2 = \rho$ is the PDF for the position of the particle.

Acknowledgments

We would like to thank Evan Tufte for his guidance and encouragement and the UCSB Directed Reading Program for the opportunity to work on this project.

References

[1] David J. Griffiths and Darrell F. Schroeter. *Introduction to quantum mechanics*. Cambridge University Press, Cambridge; New York, NY, third edition edition, 2018. ISBN 978-1-107-18963-8.

Time Independent Schrödinger Equation

If we assume V is independent of time, we can apply the method of separation of variables to the 1-D Schrödinger equation (1):

$$\Psi(x,t) = \psi(x)\varphi(t)$$

$$i\hbar \frac{1}{\varphi}\frac{d\varphi}{dt} = \frac{\hbar^2}{2m}\frac{1}{\psi}\frac{d^2\psi}{dx^2} + V.$$

Hence, both sides must be equal to a constant, E

$$i\hbar \frac{1}{\varphi}\frac{d\varphi}{dt} = \frac{\hbar^2}{2m}\frac{1}{\psi}\frac{d^2\psi}{dx^2} + V = E \implies \varphi(t) = e^{-iEt/\hbar}, \quad \frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + V\psi = E\psi. \tag{2}$$

We obtain solutions ψ_n with a corresponding allowed energy E_n, where any linear combination of separable solutions is a solution.

Particle in a Box

Consider a particle in a well such that the sides prevents the particle from being outside the well. That is, the potential can be described as

$$V(x) = \begin{cases} 0, & 0 \le x \le a \\ \infty, & \text{otherwise.} \end{cases}$$

The requirement $V = \infty$ outside of $[0,a]$ tells us $\psi(x) = 0$ outside of $[0,a]$. As $V = 0$ in $0 \le x \le a$, (2) can be written as

$$\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} = E\psi, \quad 0 \le x \le a.$$

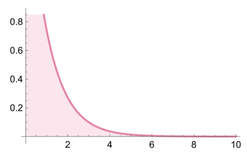
We then write

$$\frac{d^2\psi}{dx^2} = -k^2\psi, \quad k \equiv \frac{\sqrt{2mE}}{\hbar}, \quad \psi(0) = \psi(a) = 0.$$

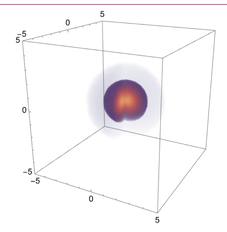
We can then solve as ODEs to obtain distinct solutions

$$E_n = \frac{n^2\pi^2\hbar^2}{2ma^2}, \quad \psi_n(x) = \sqrt{\frac{2}{a}}\sin\left(\frac{n\pi}{a}x\right) \quad 0 \le x \le a.$$

Hydrogen Orbitals

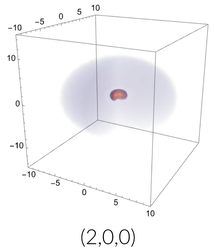


$R_{10}(r)$

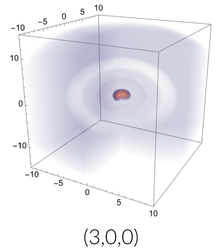


(1,0,0)

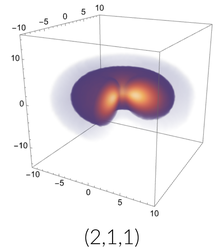
Given the solutions ψ_{nlm} from (4), we can plot the PDFs to obtain the orbital shapes of hydrogen. For the example of ψ_{100}, we observe how the radial equation, R_{10}, describes the shape of the 1s orbital of hydrogen.



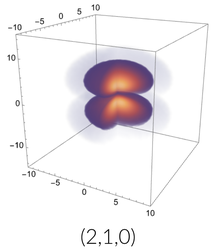
(2,0,0)



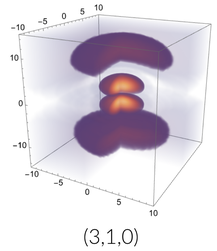
(3,0,0)



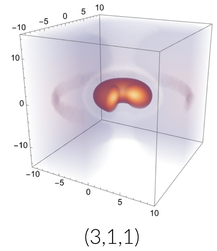
(2,1,1)



(2,1,0)



(3,1,0)



(3,1,1)

The Solution for the Hydrogen Atom

1. Time independent 3-D Schrödinger equation:
$$\frac{-\hbar^2}{2m}\nabla^2\psi + V\psi = E\psi$$
2. Convert to spherical coordinates and apply separation of variables to the radial and angular equations
$$\psi(r,\theta,\phi) = R(r)Y(\theta,\phi)$$
$$\frac{1}{R}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right) - \frac{2mr^2}{\hbar^2}[V(r) - E] = l(l+1)$$
$$\frac{1}{Y}\left\{\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2 Y}{\partial\phi^2}\right\} = -l(l+1)$$
Note $l(l+1)$ is the separation constant.
3. Separation of Variables for the angular equation
$$Y(\theta,\phi) = \Theta(\theta)\Phi(\phi)$$
$$\Phi(\phi) = e^{im\theta} \quad \text{and} \quad \Theta(\theta) = AP_l^m(\cos\theta)$$
where P_l^m is the associated Legendre polynomial.
4. Change of variables of the radial equation
$$u(r) = rR(r)$$
$$\frac{-\hbar^2}{2m}\frac{d^2u}{dr^2} + \left[V + \frac{\hbar^2}{2m}\frac{l(l+1)}{r^2}\right] = Eu \tag{3}$$
5. Solving the 3-D equation with the hydrogen potential energy function
$$V(r) = \frac{-e^2}{4\pi\varepsilon_0}\frac{1}{r}$$

Define
$$\kappa = \frac{\sqrt{-2m_eE}}{\hbar}, \quad \rho = \kappa r, \quad \rho_0 = \frac{m_e e^2}{2\pi\varepsilon_0\hbar^2\kappa}.$$

Using (3), we obtain
$$\frac{d^2u}{d\rho^2} = \left[1 - \frac{\rho_0}{\rho} + \frac{l(l+1)}{\rho^2}\right]u.$$

Define a new function
$$u(\rho) = \rho^{1+1}e^{-\rho}v(\rho).$$

Expressing $v(\rho)$ as a power series, we can develop a recurrence relation for the coefficients. The solution is $v(\rho) = L_{n-l-1}^{2l+1}(2\rho)$ where $L_q^p(x)$ is the associated Laguerre polynomial. Hence, the allowed energies are

$$E_n = -\left[\frac{m_e}{2\hbar^2}\left(\frac{e^2}{4\pi\varepsilon_0}\right)^2\right]\frac{1}{n^2} = \frac{E_1}{n^2}.$$

We then can write the spatial wave equations by calling their quantum numbers

$$\psi_{nlm}(r,\theta,\phi) = R_{nl}(r)Y_l^m(\theta,\phi)$$

$$\psi_{nlm} = \sqrt{\left(\frac{2}{na}\right)\frac{(n-l-1)!}{2n(n+l)!}}e^{-r/na}\left(\frac{2r}{na}\right)^l\left[L_{n-l-1}^{2l+1}(2r/na)\right]Y_l^m(\theta,\phi) \tag{4}$$

Connections to Linear Algebra

We make note of two important properties of the wave functions obtained from the Schrödinger equation:

1. ψ_{nlm} are orthogonal.
$$\int \psi_{nlm}^*\psi_{n'l'm'}\, r^2\sin\theta\, drd\theta d\phi = \delta_{nn'}\delta_{ll'}\delta_{mm'}$$
2. ψ_{nlm} are eigenfunctions with eigenvalue E_n of allowed energy.
$$\left(\frac{-\hbar^2}{2m}\nabla^2 + V\right)\psi_{nlm} = E_n\psi_{nlm}$$

The Toric Code: Homology and Qubits

Parsa Pourghasem Mentor: Nayeong Kim
University of California - Santa Barbara
Department of Mathematics - Directed Reading Program 2025

Introduction

The toric code is an important model in physics which is a simplest-case, analytically solvable model of many different phenomena. It is the simplest case of a $\mathbb{Z}_2$ lattice gauge theory and exhibits topological order. Most importantly, in this case, it is a basic form of a **quantum error correcting code**. That means that it can correct errors due to quantum "bit flips" and noise in its constituent **qubits**. A **qubit** is a quantum two-state system which can be represented as the normalized column vector:

$$\psi = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$$

Where $\alpha, \beta \in \mathbb{C}$ and $|\alpha|^2 + |\beta|^2 = 1$. Unlike a traditional bit, when the state of the qubit is measured, it will return a result of 0 with probability $|\alpha|^2$ and a result of 1 with a probability of $|\beta|^2$. In Dirac notation, we can write these vectors as follows:

$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$$

In this case, this notation is just shorthand but it can be much more versatile.

Pauli Operators and Multi-qubit Systems

The Pauli operators are defined as:

$$\sigma^x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma^y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma^z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

These have vital meanings in the measurement of qubits and as generators of the Lie algebra $\mathfrak{sl}(2, \mathbb{C})$. All three have eigenvalues of 1 and -1 and obey the following properties:

- Traceless: $Tr(\sigma_i) = 0 \quad \forall i \in \{x, y, z\}$
- Anti-commutative: $\sigma_i \sigma_j = -\sigma_j \sigma_i \quad \forall i \neq j$
- Involutary: $\sigma_i^2 = I \quad \forall i \in \{x, y, z\}$
- Hermitian: Their transpose conjugates are equal to themselves

These properties, along with the fact that:

$$\sigma^y = i\sigma^x\sigma^z$$

gives all the products of these matrices. We also sometimes include the identity matrix I as one of the operators to form a basis.
When working with systems with more than one qubit, we tensor product them together. In the case of two qubits we have

$$|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle$$

We can also tensor product Pauli operators and the identity in different ways to act on different parts of a multi-qubit system. We denote a Pauli operator acting on a qubit j (and identity on all others) as σ_j.

References

[1] D Browne.
Lectures on Topological Codes and Quantum Computation.
2014.

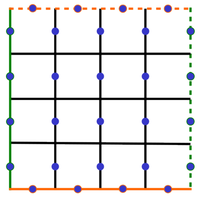
[2] A.Yu. Kitaev.
Fault-tolerant quantum computation by anyons.
Annals of Physics, 303(1):2–30, January 2003.

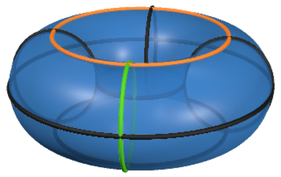
[3] M. Nakahara.
Geometry, Topology and Physics.
CRC Press, 2018.

[4] A Wang.
The toric code.
2024.

$\mathbb{Z}_2$ Homology of the Torus

Consider wrapping a $k \times k$ square lattice around a torus, joining opposite edges:

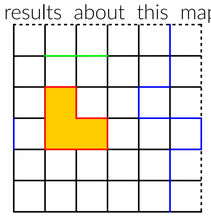




We define **0-cells** to be points/vertices on this lattice (marked in green), **1-cells** to be edges (blue), and **2-cells** to be the squares (red). We also define an **n-chain** to be sum of an element of $\mathbb{Z}_2$ (0 or 1) multiplied by a set of n-cells (the line in yellow below is a 1-chain). The space of n-chains is denoted C_n and is a $\mathbb{Z}_2$ vector space.

We now define the **boundary map** d_n which takes an n-chain and returns its boundary as an (n-1)-chain. The following are very important results about this map:

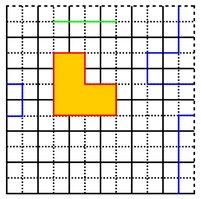
- n-chains with no boundaries are **n-cycles** and the subspace of n-cycles (the kernel of d_n) is denoted Z_n.
- n-chains which are the boundary of an (n+1)-chain are called **n-boundaries** and the subspace of these (image of d_{n+1} is denoted B_n
- All n-boundaries are n-cycles ($d_n d_{n+1} = 0$)



We define the homology group $H_n = Z_n/B_n$ and the homology groups of a torus are $\mathbb{Z}_2, \mathbb{Z}_2 \oplus \mathbb{Z}_2$ and $\mathbb{Z}_2$ respectively (up to isomorphism).

We then make another lattice called the **dual lattice** as follows:

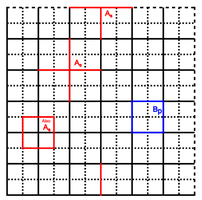
- An **n-cocell** is a (2-n)-cell on the dual lattice (ex: a 0-cocell is a square plaquette on the dual lattice)
- The chain group on the dual lattice is called the cochain group C^n
- We define coboundary (d^n), cocycles (Z^n), coboundaries (B^n), and cohomologies (H^n) analogously to that of the regular (primal) lattice.



The Toric Code

Take a $k \times k$ toric lattice like that seen before. Now place a qubit on each edge (tensor them) to create a vector space (specifically a Hilbert space) $\mathcal{N}$ and define the following operators:

$$A_s = \prod_{j \in +_s} \sigma_j^x,$$
$$B_p = \prod_{j \in \square_p} \sigma_j^z$$



This is the **toric code**. We define the **protected subspace** on the vector space of this code to be:

$$\mathcal{L} = \{|\psi\rangle \in \mathcal{N} \mid \forall s, p: \; A_s\,|\psi\rangle = |\psi\rangle\,, B_p\,|\psi\rangle = |\psi\rangle\}$$

For the set of linear operators on this space $\boldsymbol{L}(\mathcal{L})$, this is the same as setting $A_s = 1$ and $B_p = 1$:

$$\boldsymbol{L}(\mathcal{L}) \cong \boldsymbol{L}(\mathcal{N})/I$$

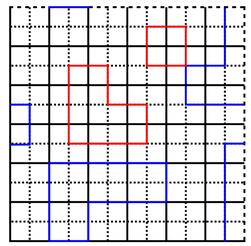
Where I is the left sided ideal generated by all $A_s - 1$ and $B_p - 1$.

Dimension of the Protected Subspace

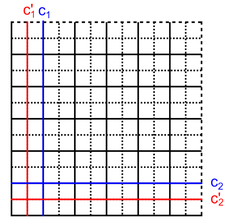
No matter the extension to $\boldsymbol{L}(\mathcal{N})$, elements of $\boldsymbol{L}(\mathcal{L})$ will always commute with A_s and B_p. Thus we loosely write that $\boldsymbol{L}(\mathcal{L}) \subseteq \mathcal{G}$ where $\mathcal{G} \subset \boldsymbol{L}(\mathcal{N})$ is the space of operators that commute with A_s and B_p which can be shown to be generated by:

$$\prod_{i \in C} \sigma_i^z, \quad \prod_{i \in C'} \sigma_i^x$$

Where C and C' are in the 1-cycles and cocycles respectively. In other words, this is all closed loops on the lattice and dual lattice.



We can see then that $\boldsymbol{L}(\mathcal{L})$ is generated by all cycles and cocycles modulo all boundaries (I being an ideal is not important in this case). In other words the generators of $\boldsymbol{L}(\mathcal{L})$ is isomorphic to $H_1 \oplus H^1$ which for the torus has rank 4, corresponding to the following cuts and operators:



$$\mathcal{Z}_1 = \prod_{i \in c_1} \sigma_i^z, \quad \mathcal{Z}_2 = \prod_{i \in c_2} \sigma_i^z,$$
$$\mathcal{X}_1 = \prod_{i \in c_1'} \sigma_i^x, \quad \mathcal{X}_2 = \prod_{i \in c_2'} \sigma_i^x$$

This generates 16 unique operators and thus the dimension of the protected subspace is 4.

Uses of the Protected Subspace

Each of the topologically defined states in the protected subspace can be considered a "codeword". These codewords can have errors on them that cause changes. We consider an error of the form:

$$E = \prod_i (\sigma_i^x)^{\alpha_i} \prod_i (\sigma_i^x)^{\beta_i} \qquad \alpha_i, \beta_i \in 0, 1$$

Our tools to catch these errors are **syndrome measurements**, measurements of the eigenvalues of A_s and B_p. This means that an error will not be caught only if it commutes with A_s and B_p ($E \in \mathcal{G}$). However, if the error is a product of only A_s and B_p then it will not change the codeword of the state so E would not be an error. Thus, to cause an error, E must contain a non-contractible chain or cochain which can only happen if the action of E is non-trivial for at least k qubits. We can also define a Hamiltonian (an energy) such that the protected subspace is the ground state:

$$H = -\sum_s A_s - \sum_p B_p$$

Then we can imagine that for small errors (which are excitations to this energy), the system will "cool" down to its ground state and automatically error-correct.

Fermat's Last Theorem

Eitan Boaz, Sohom Dutta, Talia Martin - Mentor: Waqar Ali Shah

University of California Santa Barbara

Background

In the margins of his copy of Diophantus' *Arithmetica*, Pierre de Fermat scribbled the following intriguing conundrum, circa 1637.

> *"It is impossible to separate a cube into two cubes, or a fourth power into two fourth powers, or in general, any power higher than the second, into two like powers. I have discovered a truly marvelous proof of this, which this margin is too narrow to contain."*
>
> —Pierre de Fermat

This sparked one of the most famous mathematical problems in modern history that came to be known as **Fermat's Last Theorem** (FLT). It asserts that no three positive integers a, b, c satisfy

$$a^n + b^n = c^n$$

for any integer $n > 2$. Despite its seemingly simple appearance, the theorem eluded a complete proof for over 350 years. Although FLT applies to all integers $n > 2$, it can be reduced to two fundamental cases:

- $n = 4$, which Fermat himself proved using his method of *infinite descent*.
- $n = p$, where p is an odd prime.

Indeed, any $n \geq 3$ must have one of the aforementioned factors. The general case for odd primes remained unresolved for centuries. Our project focuses on a major breakthrough achieved in the 19th century by Ernst Eduard Kummer, who resolved FLT for a special class of prime exponents. His work laid the foundations for modern algebraic number theory and ultimately led to Andrew Wiles' complete proof in the 1990s.

Key Definitions and Concepts

Def 1. A *number field* K is a field containing $\mathbb{Q}$ such that K is a finite dimensional vector space over $\mathbb{Q}$. The *ring of integers* of K is the set of all elements in K that are a root of a monic polynomial with integer coefficients.

Def 2. The *ideal class group* C_K of a number field K is the quotient group J_K/P_K where J_K is the group of fractional ideals of the ring of integers of K, and P_K is the subgroup of principal ideals in J_K. It is a finite abelian group.

Def 3. The *class number* of a number field K is the order of its ideal class group C_K. It quantifies the failure of the ring of integers of K to be a unique factorization domain.

Def 4. A *Dedekind domain* is an integral domain in which all non-zero ideals factor uniquely into a product of prime ideals.

Def 5. A primitive p-th *root of unity* is a number $\zeta_p \in \mathbb{C}$ such that $(\zeta_p)^p = 1$ but $(\zeta_p)^k \neq 1$ for all $k \in \mathbb{N}$ with $k < p$. The p-th *cyclotomic field* is the number field generated by $\mathbb{Q}$ and any choice of a primitive ζ_p. We will henceforth denote such a ζ_p by ζ.

Key facts

Fact 1. The ring of integers of $\mathbb{Q}(\zeta)$ is $\mathbb{Z}[\zeta]$, which is a Dedekind domain.

Fact 2. The ideal $\langle 1 - \zeta \rangle$ of $\mathbb{Z}[\zeta]$ is a prime ideal and $\langle 1 - \zeta \rangle^{p-1} = \langle p \rangle$.

Fact 3. All units in $\mathbb{Z}[\zeta]$ are of the form $\epsilon\zeta^i$ where ϵ is real and i is an integer.

Fact 4. If I is an ideal of $\mathbb{Z}[\zeta]$ such that I^k is principal for some positive integer k that does not divide the class number of $\mathbb{Q}(\zeta)$, then I is itself principal.

Kummer's Special Case

In the 19th century, Ernst Eduard Kummer introduced groundbreaking ideas that laid the foundation of modern algebraic number theory and used them to prove FLT for a conjecturally infinite class of prime exponents p known as *regular* primes—those for which p does not divide the class number of $\mathbb{Q}(\zeta_p)$.

- Use the factorization $a^p + b^p = (a + b)(a + \zeta b) \cdots (a + \zeta^{p-1} b)$.
- Work in $\mathbb{Z}[\zeta]$, which is analogous to $\mathbb{Z}$ but lacks the unique factorization property.
- Introduce *ideal numbers* (early version of ideals) to handle the failure of unique factorization in $\mathbb{Z}[\zeta]$.
- Establish that if p is regular, then one can bypass the failure of unique factorization in $\mathbb{Z}[\zeta]$ and prove FLT for the exponent p.

This resolved FLT for many small primes. Indeed, the only *irregular* primes less than 100 are 37, 59 and 67. It is known that there are infinitely many irregular primes and it conjectured that the number of regular primes is also infinite.

Proof Structure

We outline a proof of the following special case of Kummer's breakthrough:

Theorem. *If p is an odd regular prime, the Diophantine equation $x^p + y^p = z^p$ has no solutions in positive integers x, y, z where $p \nmid xyz$.*

Suppose on the contrary that there is such a triplet. Then there exists another triplet whose members are pairwise coprime and which also satisfies FLT. Using properties of roots of unity and passing to ideals, we can factor $x^p + y^p$, yielding:

$$\prod_{i=0}^{p-1} \langle x - \zeta^i y \rangle = \langle z \rangle^p \qquad (\dagger)$$

Each factor on the left hand side must be pairwise coprime. Assuming otherwise implies there is some prime ideal $\mathfrak{p}$ that contains z and $y(1 - \zeta)\zeta^k$ for some integer k. Since ζ^k is a unit and $\mathfrak{p}$ is a prime, either $\mathfrak{p}$ contain y or $1 - \zeta$. Suppose

- $y \in \mathfrak{p}$. Then since y and z are coprime, we have $1 \in \mathfrak{p}$, contradiction.
- $1 - \zeta \in \mathfrak{p}$. As $\langle 1 - \zeta \rangle$ is prime by the first part of **Fact 2**, we have $\mathfrak{p} = \langle 1 - \zeta \rangle$. The second part of **Fact 2** then forces $p \mid z$ through a series of congruences, a contradiction.

Since the right hand side of (†) is a p-th power, and since, by **Fact 1**, a Dedekind domain like $\mathbb{Z}[\zeta]$ guarantees unique factorization into prime ideals, each factor on the left hand side must also be a p-th power of an ideal. Therefore, there exists some ideal I such that

$$\langle x - \zeta y \rangle = I^p.$$

Since I^p is a principal ideal and the prime p is regular, I itself is a principal ideal by **Fact 4**. Therefore, there exists some $\delta \in \mathbb{Z}[\zeta]$ such that $I = \langle \delta \rangle$. This implies that $x - \zeta y = u\delta^p$ where $u \in \mathbb{Z}[\zeta]$ is a unit. By **Fact 3**, all units in $\mathbb{Z}[\zeta]$ can be expressed in the form $\epsilon\zeta^i$ for some real element $\epsilon \in \mathbb{Z}[\zeta]$ and some integer i. We therefore have

$$x - \zeta y = \epsilon\zeta^i \delta$$

By considering this equation modulo $\langle p \rangle$, rearranging, and substituting, we can deduce that

$$\frac{x}{p}\zeta^{-i} + \frac{y}{p}\zeta^{1-i} - \frac{x}{p}\zeta^p - \frac{y}{p}\zeta^{i-1} \in \mathbb{Z}[\zeta].$$

Since p does not divide x or y, some of the exponents must be congruent modulo p in order for their corresponding terms to combine. The possible congruences produce contradictions as follows:

- $i \equiv 0 \pmod{p}$ and $i \equiv 1 \pmod{p}$ imply the contradiction $p \mid xyz$.
- $2i \equiv 1 \pmod{p}$ implies that $p = 3$, which can be solved by working mod 9.

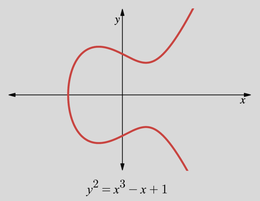
Remark. The case in which $p \mid xyz$ was also shown by Kummer using similar methods.

From Fermat To Modularity

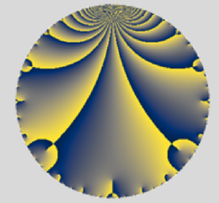
Modularity Theorem. *Every elliptic curve over $\mathbb{Q}$ is modular.*

That is, there exists a correspondence between two seemingly distinct objects: elliptic curves and modular forms.

- *Elliptic curves* are smooth projective algebraic curves of the form $y^2 = x^3 + ax + b$, whose rational points form a finitely generated abelian group.
- *Modular forms* are complex analytic functions on the upper half-plane that exhibit a rich symmetry under the action of modular transformations.



An elliptic curve



A disk model plot of the associated modular form

Andrew Wiles And The Proof

In 1985, Gerhard Frey observed that a hypothetical counterexample to Fermat's Last Theorem would give rise to a special elliptic curve—now known as the *Frey curve*—which, if the modularity conjecture were true, could not be modular. Building on this insight, Ken Ribet proved that this was indeed the case a year later.

Andrew Wiles's crucial contribution was his proof of the modularity theorem for *semistable* elliptic curves. Using the theory of *Galois representations*, he uncovered a profound link between the arithmetic of elliptic curves and modular forms. However, the initial proof contained a serious gap. With the help of Richard Taylor, Wiles later resolved the issue. Combined with Ribet's theorem, this breakthrough eliminated any possible counterexample to Fermat's Last Theorem. The final proof appeared in 1995.

Andrew Wiles at the 1993 Cambridge lecture.

Acknowledgements

We are deeply grateful to our mentor, Waqar Ali Shah, for his insightful guidance and support throughout this project. We also thank the UCSB Directed Reading Program for making this opportunity possible.

References

[1] David S. Dummit and Richard M. Foote. *Abstract Algebra*. 3rd. John Wiley and Sons, 2004.

[2] Ian Stewart and David Tall. *Algebraic Number Theory and Fermat's Last Theorem*. 3rd. 2002.

[3] Lawrence C. Washington. *Introduction to cyclotomic fields*. 2nd. Springer-Verlag, 1997.

UC **SANTA BARBARA**

Permutation Wordle and List Strategies

Mia Mastrocola and Roselyn Soto, Mentored by Joel Pion

University of California, Santa Barbara — Directed Reading Program

Introduction

Permutation Wordle is a Wordle-inspired game in which both the secret code and each guess are permutations of $\{1, \dots, n\}$, and after every guess you learn which entries lie in the correct position (analogous to Wordle's green tiles). It challenges a guesser to recover a hidden permutation $\sigma \in S_n$ by making successive permutation guesses with this "correct-position" feedback.

Permutation Wordle

Hidden:	3 2 1	
Guess 1:	1 2 3	1 correct (position 2)
Guess 2:	3 2 1	All correct!

We study the **Circular Shift** strategy, proposed by Kutin and Smithline [1]. We prove that it constitutes a list strategy and we show that any valid game-tree induces such a list. We conjecture that Circular Shift is optimal among all list strategies.

Definitions

For any strategy A and target $\sigma \in S_n$, let $G_A(\sigma)$ denote the number of guesses A uses to correctly identify σ. For a fixed σ we write γ_r for the rth guess made by a strategy. We say a permutation π is *consistent with* the first $r-1$ guesses $\gamma_1, \dots, \gamma_{r-1}$ if, under feedback, σ would not have been eliminated by those guesses.

Weak domination: Strategy A *weakly dominates* B if its average guess-count

$$\overline{G}(A) \coloneqq \frac{1}{n!} \sum_{\sigma \in S_n} G_A(\sigma)$$

satisfies $\overline{G}(A) \leq \overline{G}(B)$

Strong domination: For each $r \geq 1$, let

$$N_A(r) \coloneqq |\{\sigma \in S_n \colon G_A(\sigma) \leq r\}|$$

Then A *strongly dominates* B if

$$\forall r \colon \; N_A(r) \geq N_B(r)$$

List Strategy: A *list strategy* L is defined by a fixed total ordering

$$L \coloneqq \left(\sigma^{(1)}, \sigma^{(2)}, \dots, \sigma^{(n!)}\right) \subset S_n$$

In which $\gamma_1 = \sigma^{(1)}$, and for each $r > 1$,

$$\gamma_r \coloneqq \sigma^{(j)} \text{ such that j = } \min_L \{i \colon \sigma^{(i)} \text{ is consistent with } \gamma_1, \dots, \gamma_{r-1}\}$$

Excedances and Eulerian Numbers

Permutation Wordle's performance under Circular Shift is governed by *excedances*. An *excedance* of $\sigma \in S_n$ is a position i with $\sigma(i) > i$; we write $exc(\sigma)$ to denote how many exceedances σ has. For example:

3 1 **6** 2 **7** 4 5

a permutation of S_7, with bold boxes showing excedances $\sigma(i) > i$.

The *Eulerian number*

$$A(n,k) \coloneqq \left|\{\sigma \in S_n \colon exc(\sigma) = k\}\right|$$

counts permutations in S_n with exactly k excedances. Under Circular Shift, each targetσ is solved in precisely $1 + exc(\sigma)$ guesses which yields an average guess count of $(n+1)/2$ across all permutations, see [1] for details. In practice, this means that almost all permutations are solved in very close to the average number of guesses, underscoring both the predictability and efficiency of Circular Shift.

List Strategies and Game Trees

A game tree encodes every possible play of a list strategy as a rooted branching diagram: each node is a guess, and each outgoing edge corresponds to the feedback received. The depth of each node gives the number of guesses needed for that permutation.

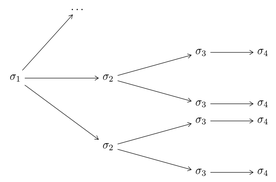


Figure 1: A partial tree diagram showing the sequence of guesses σ_i; ellipses indicate further continuation.

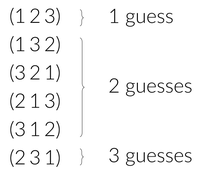


Figure 2: A predetermined ordering of permutations of S_3 by guess count that defines a list strategy.

By recording the nodes of any valid game tree in breadth-first order, you obtain exactly the ordering that reproduces that tree's guess behavior. Conversely, given any list, you can rebuild its game tree by following each path of guesses

Circular Shift in Action

The **Circular Shift** strategy: begin with the identity, then at each step cyclically rotate all misplaced entries one position to the right, leaving fixed entries untouched.

CircularShift Strategy (n=4)

Hidden:	2 3 1 4	
Guess 1:	1 2 3 4	Feedback: Only position 4 correct
Guess 2:	3 1 2 4	Feedback: Only position 4 correct
Guess 3:	2 3 1 4	Feedback: All correct!

Proposition: Circular Shift is a List Strategy.

Let $S(\sigma) = i$ denote the number of guesses Circular Shift needs to reach σ, for all i with $1 \leq i \leq n!$. We proceed by induction.

Base Case ($k = 1$). Only the identity permutation id = $(1, 2, ..., n)$ is solved in one guess, so $\sigma^{(1)} = \text{id}$ and $S(\text{id}) = 1$.

Inductive Hypothesis. Assume that, for a target σ, Circular Shift's first k guesses are

$$(\gamma_1, \gamma_2, \dots, \gamma_k) \;=\; (\sigma^{(1)}, \sigma^{(2)}, \dots, \sigma^{(k)}) \quad \text{where } S(\sigma^{(i)}) = i \quad \text{for all } i \text{ with } 1 \leq i \leq k.$$

Inductive Step. Consider $\sigma^{(k+1)}$. Note $\sigma^{(k+1)} = \sigma$. That is,

$$(\sigma^{(k+1)})^{(1)} = \sigma^{(1)}, \; (\sigma^{(k+1)})^{(2)} = \sigma^{(2)}, \; \dots, \; (\sigma^{(k+1)})^{(k)} = \sigma^{(k)}.$$

Notice that if for some $1 \leq j \leq k$ we had

$$\left(\sigma^{(k+1)}\right)^{(j)} \neq \sigma^{(j)},$$

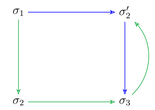
then $\sigma^{(j)}$ would conflict with the feedback from the first $j-1$ guesses. But by our hypothesis, those first $j-1$ guesses already agree, and Circular Shift's deterministic nature forces the jth guess to coincide as well. Thus all k initial guesses agree, so

$$\gamma_r = \sigma^{(r)}, \quad r = 1, \dots, k,$$

and $\gamma_{k+1} = \sigma^{(k+1)}$ with $S(\sigma^{(k+1)}) = k+1$. Moreover, this next guess is unique: if two targets ρ, ω are both consistent with guesses $\gamma_1, \dots, \gamma_k$ and both reach guess γ_{k+1} in the strategy, then we must have $\rho^{(k+1)} = \omega^{(k+1)}$, hence $\rho = \omega$. This completes the induction. Therefore, by listing all permutations in order of increasing $S(\sigma)$, we obtain a list L, similar to that in Figure 2. Circular Shift follows such a list exactly, so it is indeed a list strategy.

Our Conjectures

- **DAGs Dominate General Stratgies**
  What We Know: In a general strategy, modeled by an arbitrary directed graph, tree structures have the potential for cycles, allowing for the repetition of previous guesses.



  The diagram above illustrates a general strategy whose directed graph contains such a cycle. From the root guess σ_1 one can follow the green feedback path:

$$\sigma_1 \to \sigma_2 \to \sigma_3 \to \sigma_2'$$

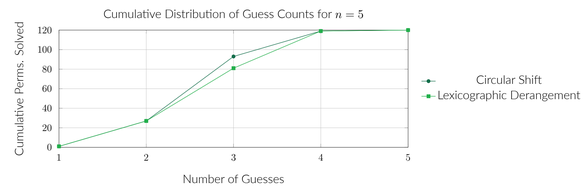
  Creating a loop in which guess σ_3 can be repeated. By contrast, a *directed acyclic graph* (DAG) strategy forbids any such cycles: its acyclic nature guarantees no closed loops, meaning no possibility of repeated guesses. We conjecture that for any cyclic general strategy, there is a DAG that solves every target in no more guesses.
  Concept of Proof: We hypothesize that successful induction on the removal of the number of loops within a directed cyclic graph would result in a dominating DAG structure.

- **List Strategies Dominate DAGs**
  What We Know: In a DAG strategy, a single guess can be reached by multiple feedback paths. As shown in the figure below, the two nodes σ_2 and σ_2' both point to σ_3, as well as σ_4 and σ_4' with σ_5:

  Conversely, within a list strategy, each node has exactly one incoming edge. Namely, there is a unique parent for each guess, so no two different feedback histories can ever converge on the same next guess. We conjecture that for any DAG-based strategy, there is a list strategy that solves every target in no more guesses than the DAG.

- **Circular Shift is a List Strategy** ✓

- **Circular Shift Dominates List Strategies**
  What We Know: Not all list strategies perform equally. In computational experiments, we compared Circular Shift against multiple other valid list strategies, such as Lexicographic Derangements, shown in the graph below. In every case, Circular Shift strongly dominated. We conjecture that Circular Shift is *optimal* among *all* list strategies.



  Concept of Proof: We hypothesize that for any strategy A, we can bound $N_A(r)$ above by $\sum_{i=1}^{r} A(n,i)$, the sum of the *Eulerian Number* for each value of r. Whereas for Circular Shift, $N(r)$ is exactly equal to $\sum_{i=1}^{r} A(n,i)$ [1].

Acknowledgements

We would like to thank our mentor Joel Pion for his enthusiasm and support throughout this process and for introducing us to Permutation Wordle! We also thank the DRP program for the opportunity to work on this project.

References

[1] S. Kutin and L. Smithline, *Permutation Wordle*, arXiv:2408.00903v2 [math.CO], 2024.

Lie Algebra of a Lie Group

Zih-Yu Hsieh Mentor: Arthur Jiang
University of California Santa Barbara

Tangent Vectors as Derivations

When embedding smooth manifolds into $\mathbb{R}^n$, tangent vectors are associated with directional derivatives. To generalize tangent vectors into abstract smooth manifold, we need an analogy:

Definition

Any point $u \in M$, a **Derivation at** u, is a linear map $v_u : C^\infty(M) \to \mathbb{R}$, that satisfies the product rule:

$$\forall f, g \in C^\infty(M), \quad v_u(fg) = f(u)(v_u g) + g(u)(v_u f)$$

The vector space of all derivations at u, or $T_u(M)$, is the **Tangent Space** of M at u, and each derivation $v_u \in T_u(M)$ is a **Tangent Vector** of u.

Vector Fields & Smooth Condition

Definition

a vector field is a map $X : M \to TM$ (TM denotes the **Tangent Bundle**), with $X(u) = X_u \in T_u(M)$.
Which, X is a **Smooth Vector Field**, if $X : M \to TM$ is a smooth map.
A collection of smooth vector fields on M is $\mathfrak{X}(M)$, which is an $\mathbb{R}$-vector space.

Another equivalent condition of saying X is smooth, is through smooth functions $f \in C^\infty(M)$: For all $u \in M$, $X(u) = X_u \in T_u(M)$ is a derivation at u, define $Xf : M \to \mathbb{R}$ by $Xf(u) = X_u(f)$. Which, the **Derivation** is an equivalent condition for smooth vector field:

Theorem

Given vector field X, $X \in \mathfrak{X}(M)$ iff it satisfies product rule. i.e. For all $u \in M$, and all $f, g \in C^\infty(M)$:

$$X(fg)(u) = X_u(fg) = f(u)(X_u g) + g(u)(X_u f) = f(u)Xg(u) + g(u)Xf(u)$$
$$\implies X(fg) = f(Xg) + g(Xf)$$

Vector Fields of Different Manifolds

Given M, N two smooth manifolds, and smooth map $F : M \to N$. Let $X \in \mathfrak{X}(M)$, an ideal situation is mapping X to a smooth vector field of N through F. Yet, this requires F to be bijective:

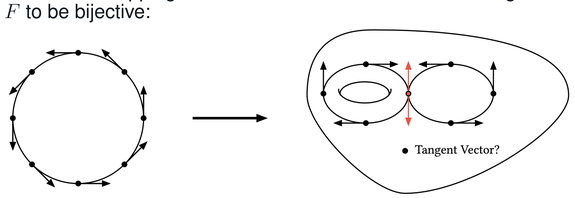


Figure 1: Example of Conflicting Tangent Vectors

So, we'll consider a weaker notion:

Definition

Given $X \in \mathfrak{X}(M)$ and $Y \in \mathfrak{X}(N)$, the two are ***F*-related**, if for all $u \in M$, the following is true:

$$dF_u(X_u) = Y_{F(u)}$$

Simply speaking, F maps the tangent vectors collected by X, to be compatible with tangent vectors collected by Y.

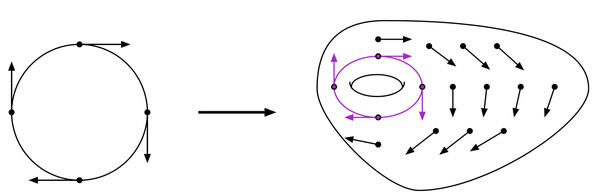


Figure 2: A demonstration of F-Relation

Lie Bracket of Vector Fields

The initial motivation is to combine two vector fields $X, Y \in \mathfrak{X}(M)$ to be another vector field. For all $f \in C^\infty(M)$, since $Yf \in C^\infty(M)$, then $XYf := X(Yf) \in C^\infty(M)$. But, in general XY is not a derivation, hence not a vector field:

Example

Define vector fields $X = \frac{\partial}{\partial x}$, $Y = x\frac{\partial}{\partial y}$ on $\mathbb{R}^2$. Take smooth functions $f(x, y) = x$ and $g(x, y) = y$, then we get the following:

$$XY(fg) = X\left(x\frac{\partial}{\partial y}(xy)\right) = \frac{\partial}{\partial x}(x^2) = 2x$$

But, product rule doesn't hold for this example:

$$f(XYg) + g(XYf) = x\left(X\left(x\frac{\partial}{\partial y}(y)\right)\right) + y\left(X\left(x\frac{\partial}{\partial y}(x)\right)\right) = x$$

So, we need to define a new operation on vector fields:

Definition

The **Lie Bracket** $[\cdot, \cdot] : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$, is defined as:

$$\forall X, Y \in \mathfrak{X}(M), \quad [X, Y] = XY - YX$$

Which, the output $[X, Y] \in \mathfrak{X}(M)$, since it satisfies product rule:

$$[X,Y](fg) = X(Y(fg)) - Y(X(fg)) = X(f(Yg) + g(Yf)) - Y(f(Xg) + g(Xf))$$
$$= f(XYg)+(Yg)(Xf)+g(XYf)+(Yf)(Xg)-f(YXg)-(Xg)(Yf)-g(YXf)-(Xf)(Yg)$$
$$= f(XYg - YXg) + g(XYf - YXf) = f[X,Y](g) + g[X,Y](f)$$

Lie Bracket also satisfies these properties:

- **Bilinearity:** $[aX + bY, Z] = a[X, Z] + b[Y, Z]$
- **Antisymmetry:** $[X, Y] = -[Y, X]$
- **Jacobi's Identity:** $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$

Moreover, Lie Bracket inherits relation of smooth maps:

Theorem

Given smooth map $F : M \to N$, if $X_1, X_2 \in \mathfrak{X}(M)$ and $Y_1, Y_2 \in \mathfrak{X}(N)$ are F-related respectively, then $[X_1, X_2] \in \mathfrak{X}(M)$ and $[Y_1, Y_2] \in \mathfrak{X}(N)$ are also F-related.

Lie Groups & Left-Invariant Vector Fields

The initial motivation is to study group structures in some smooth manifolds.

Definition

A **Lie Group** G, is a smooth manifold along with group structure, such that the group operation $P : G \times G \to G$ by $P(g, h) = gh$, and the inversion map $i : G \to G$ by $i(g) = g^{-1}$ are both smooth maps between manifolds.

For all $g \in G$, denote the left multiplication $L_g : G \to G$ by $L_g(h) = gh$, since $L_g = P\,|_{\{g\}\times G}$, it is a smooth map. Hence, there's a notion of X being L_g-related to itself:

Definition

Given any $X \in \mathfrak{X}(G)$ and all $g \in G$, X is a **Left-Invariant Vector Field**, if for all $g \in G$, X is L_g-related to itself. Which, for all $g \in G$:

$$d(L_g)_e(X_e) = X_{L_g(e)} = X_g$$

So, X is uniquely determined by its tangent vector at identity, $X_e \in T_e(G)$. In fact, each $v_e \in T_e(G)$ also corresponds to a unique Left-Invariant vector field.
The collection of Left-Invariant vector fields $\mathfrak{g} \subseteq \mathfrak{X}(G)$, is itself a linear subspace, and $\mathfrak{g} \cong T_e(G)$ as vector spaces, based on the above relation.

Recall that Lie Bracket of vector field preserves F-relation between manifolds, so:

Theorem

For all $X, Y \in \mathfrak{g}$, since for all $g \in G$, X and Y are L_g-related to themselves, then the Lie Bracket $[X, Y]$ is also L_g-related to $[X, Y]$. Hence, $[X, Y]$ is also left-invariant, or $[X, Y] \in \mathfrak{g}$. So, $\mathfrak{g}$ is closed under Lie Bracket's operation.

Lie Algebra on a Lie Group

Definition

Given a vector space $\mathfrak{g}$ over $\mathbb{R}$ or $\mathbb{C}$, with a binary operation $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, such that the following holds:

- **Bilinearity:** $[aX + bY, Z] = a[X, Z] + b[Y, Z]$
- **Antisymmetry:** $[X, Y] = -[Y, X]$
- **Jacobi's Identity:** $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$

Then, the pair $(\mathfrak{g}, [\cdot, \cdot])$ is a **Lie Algebra**.

In general, Lie Algebra is non-associative, so Jacobi's Identity is an alternative condition. Finally, we can define **Lie Algebra of a Lie Group:**

Definition

Given a lie group G, since the subset of left-invariant vector fields $\mathfrak{g} \subseteq \mathfrak{X}(G)$ forms a linear subspace, while closed under Lie Bracket's operation, then the pair $(\mathfrak{g}, [\cdot, \cdot])$ forms a **Lie Algebra** of G, denoted as $Lie(G)$.

Here's an example of Lie Algebra on a Lie Group:

Example

General Linear Group & its Lie Algebra:
Given $M_n(\mathbb{R}) \cong \mathbb{R}^{n^2}$, and $GL_n(\mathbb{R}) \subset M_n(\mathbb{R})$ as an open subset, it's a natural smooth manifold with dimension n^2. The product of matrices and the inversion are smooth maps, so $GL_n(\mathbb{R})$ is a Lie Group.
Now, consider $\mathfrak{g} = Lie(GL_n(\mathbb{R}))$: Each $X \in \mathfrak{g}$ is uniquely characterized by $X_{I_n} \in T_{I_n}(GL_n(\mathbb{R}))$. And, as vector spaces, $\mathfrak{g} \cong T_{I_n}(GL_n(\mathbb{R}))$.

Lie Algebra on $M_n(\mathbb{R})$:
Given $M_n(\mathbb{R})$ as $\mathbb{R}$-vector space and the commutator $[A, B] = AB - BA$, the pair $(M_n(\mathbb{R}), [\cdot, \cdot])$ in fact forms a Lie Algebra, denoted as $\mathfrak{gl}_n(\mathbb{R})$.

Lie Algebra Isomorphism between $\mathfrak{g}$ and $\mathfrak{gl}_n(\mathbb{R})$:
$GL_n(\mathbb{R})$ has a global coordinate provided by $M_n(\mathbb{R})$, denote as $(X^i_j)_{1\le i,j\le n}$.
For each $A \in \mathfrak{gl}_n(\mathbb{R})$, it corresponds to a tangent vector in $T_{I_n}(GL_n(\mathbb{R}))$:

$$A = (A^i_j) \mapsto A^i_j \left.\frac{\partial}{\partial X^i_j}\right|_{I_n}$$

The above tangent vector defines a Left-Invariant vector field $A^L \in \mathfrak{g}$. For all $X \in \mathfrak{g}$, the left multiplication L_X is in fact a linear operator on $M_n(\mathbb{R})$, so its differential is identical to itself. Which, it provides the following relation:

$$A^L_X = d(L_X)_{I_n}\left(A^i_j \left.\frac{\partial}{\partial X^i_j}\right|_{I_n}\right) = X^i_j A^j_k \left.\frac{\partial}{\partial X^i_k}\right|_X, \quad A^L = X^i_j A^j_k \frac{\partial}{\partial X^i_k}$$

Then, for arbitrary $A, B \in \mathfrak{gl}_n(\mathbb{R})$, Lie Bracket of $A^L,\ B^L \in \mathfrak{g}$ generates:

$$\left[A^L, B^L\right] = X^i_j A^j_k \frac{\partial}{\partial X^i_k}(X^p_q B^q_r)\frac{\partial}{\partial X^p_r} - X^p_q B^q_r \frac{\partial}{\partial X^p_r}(X^i_j A^j_k)\frac{\partial}{\partial X^i_k}$$

Because each $A^j_k,\ B^q_r$ are constants, while $\frac{\partial}{\partial X^i_k} X^p_q = 1$ iff $(i, k) = (p, q)$ and is 0 otherwise. Then, match up related indices, we get:

$$\left[A^L, B^L\right] = X^i_j (A^j_k B^k_r - B^j_k A^k_r)\frac{\partial}{\partial X^i_r} = (AB - BA)^L = [A, B]^L$$

Hence, the map $\mathfrak{gl}_n(\mathbb{R}) \to \mathfrak{g}$ by $A \mapsto A^L$ is a Lie Algebra Isomorphism.

Acknowledgements & Reference

I really thank my mentor Arthur Jiang for the effort, guidance, and insights on the materials and this project, and also UCSB Math DRP for this opportunity. Finally, check out my peer **Siyu Chen's** cool poster about Lie Group and Lie Algebra's applications in physics!

Reference: Lee, J.M. *Introduction to Smooth Manifolds*; 2nd ed.; Springer: New York, 2012; 9781441999825

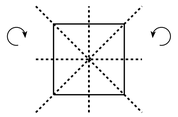
Geometric Realization of Coxeter Groups

Nate Annau and Jesse Cobb — Mentor: Benedict Lee

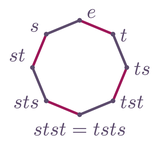
University of California, Santa Barbara

Coxeter Systems

Dihedral groups, well known from elementary group theory, encode the symmetries of regular polygons. For example, the square admits eight fundamental symmetries: reflections through its horizontal and vertical axes and two diagonals, and rotations through 90, 180, and 270 degrees.



$$D_8 = \langle r, s \mid r^4 = s^2 = (sr)^2 = 1 \rangle$$



Through the transformation $t := sr$, we can describe this group solely in terms of reflections, which gives rise to the notion of a **geometric reflection group**—a more generic group generated by reflections across a particular set of hyperplanes H_i acting on Euclidean, hyperbolic, or spherical spaces.

Abstracting this further, beyond purely geometric considerations, produces the **Coxeter group**, defined by the presentation

$$W = \langle s_1, s_2, ..., s_n \mid (s_i s_j)^{m_{ij}} = 1 \rangle$$

where $m_{ii} = 1$ and $m_{ij} = m_{ji} \in \{2, 3, ...\} \cup \{\infty\}$ for all distinct i, j. Observe this definition implies each generator s_i is an involution and thus corresponds to a reflection. Denoting the indexing set $\{s_i\}$ by S, we define a **Coxeter System** as the pair (W, S).

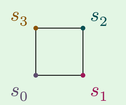
Importantly, for some $T \subseteq S$, we define the **parabolic subgroup** W_T of W by $W_T = \langle T \rangle$; we can show that (W_T, T) is a Coxeter System as well.

Chambers and Nerves

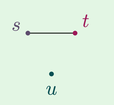
For the following definitions let (W, S) be a Coxeter system. Then, for the diagrams shown below, let $S^{(0)} = \{s_0, s_1, s_2, s_3\}$, $S^{(1)} = \{s, t, u\}$, and

$$W^{(0)} = \langle s_0, s_1, s_2, s_3 \mid s_i^2 = (s_i s_{i+1})^2 = 1, \forall i \in \mathbb{Z}_4 \rangle$$
$$W^{(1)} = \langle s, t, u \mid s^2 = t^2 = u^2 = 1; (st)^3 = 1 \rangle \cong D_6 * C_2.$$

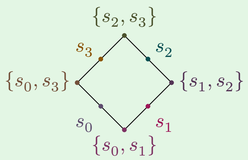
- An **abstract simplicial complex** is a set V, called the vertex set, and a collection X of finite subsets of V such that:
  - $\{v\} \in X$ for all $v \in V$
  - if $\Delta \in X$ with $\Delta' \subseteq \Delta$ then $\Delta' \in X$
- The **nerve** of (W, S), denoted by $L = L(W, S)$, is the abstract simplicial complex with a simplex σ_T for each $T \subseteq S$, where $T \neq \emptyset$ and W_T is finite
- Let L' be the **barycentric subdivision** of L (adding additional simplices at the barycenters of existing simplices representing the parabolic subgroups W_T)
- The **chamber** K is the cone on L'. For each $s \in S$, we can define the closed star in L' of the vertex s to be $K_s \subseteq K$
- The point added by the cone is the empty set $\emptyset$ in the simplicial complex, which represents the subgroup $W_\emptyset$



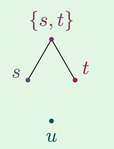
$L(W^{(0)}, S^{(0)})$



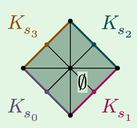
$L(W^{(1)}, S^{(1)})$



$L'(W^{(0)}, S^{(0)})$



$L'(W^{(1)}, S^{(1)})$



$K(W^{(0)}, S^{(0)})$

$K(W^{(1)}, S^{(1)})$

The Davis Complex as a Basic Construction

We wish to realize a Coxeter group, returning it to its geometric origins. This is the idea behind the **basic construction** $\mathcal{U}(W, X)$.

We begin with some additional definitions. If (W, S) is a Coxeter system and X is a connected and Hausdorff topological space, define a **mirror structure** on X over S by a collection $(X_s)_{s \in S}$, where each X_s is a nonempty, closed subset of X. Call each X_s the s-mirror of X. The idea is to "glue" copies of X along the mirrors.

For each point $x \in X$, define $S(x) \subseteq S$ by $S(x) := \{s \in S : x \in X_s\}$. Define a relation $\sim$ on $W \times X$ by

$$(w, x) \sim (w', x') \text{ if and only if } x = x' \text{ and } w^{-1}w' \in W_{S(x)}.$$

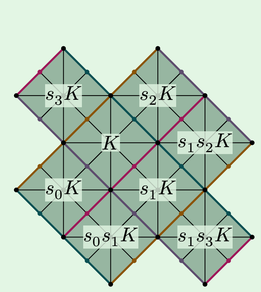
Now the basic construction is the quotient

$$\mathcal{U}(W, X) = W \times X / \sim$$

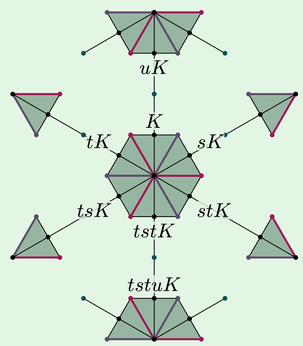
equipped with the quotient topology. We then define the **Davis complex** $\Sigma(W, S)$ as

$$\Sigma(W, S) = \mathcal{U}(W, K)$$

where K is a chamber with the mirror structures $(K_s)_{s \in S}$ as defined previously.



$\Sigma(W^{(0)}, S^{(0)})$



$\Sigma(W^{(1)}, S^{(1)})$

The Davis Complex is *CAT(0)*

We say a **geodesic space** X is $\mathbf{CAT(0)}$ if the triangles in X appear no "fatter" than triangles in a Euclidean space $\mathbb{E}^n$ of same dimension. Similarly, we can define $\mathbf{CAT(-1)}$ and $\mathbf{CAT(1)}$ for triangles in X compared to triangles in hyperbolic space $\mathbb{H}^n$ and spherical space $\mathbb{S}^n$ respectively. In order to realize this condition for the Davis complex, we must construct a metric for it.

We first choose a collection $\underline{d} = (d_s)_{s \in S}$ for which $d_s > 0$ for any $s \in S$. For finite W_T, let C_T be a chamber in $\mathbb{R}^n$ generated by the intersection of the half spaces produced by the hyperplanes H_t for $t \in T$ as in the **Tits Representation**. Then we can define the unique point x_T in the interior of C_T such that $d(x_T, H_t) = d_t$ for all $t \in T$. We then metrize each cell of $\Sigma(W, S)$, wW_T, as a copy of the polytope generated by the W_T-orbit of x_T (a standard choice for $\underline{d}$ is $d_s = \frac{1}{2}$ for all s constructing the path metric for the 1-skeleton of $\Sigma(W, S)$).

Using this metric, it was shown by Moussong and Gromov that the Davis complex for a Coxeter group W, $\Sigma(W, S)$, is a complete $\mathrm{CAT}(0)$ space. This result implies the **contractability** of the Davis complex, and shows that the **word problem** (whether two words represent the same element) and **conjugacy problem** (whether two words represent conjugate elements) are solvable for W.

Tits Representation

A key result due to Jacques Tits gives a faithful linear representation for (W, S),

$$\rho : W \to \mathrm{GL}_n(\mathbb{R}),$$

with $n = |S|$, such that for each $s_i \in S$, $\rho(s_i) = \sigma_i$ is a linear involution whose fixed set is a hyperplane and for all $i \neq j$, the product $\sigma_i \sigma_j$ has order m_{ij}. Consider the real vector space V with basis $\{e_1, ..., e_n\}$, and define a symmetric bilinear form B on V by

$$B(e_i, e_j) = \begin{cases} -\cos\left(\frac{\pi}{m_{ij}}\right) & \text{if } m_{ij} \text{ is finite} \\ -1 & \text{if } m_{ij} = \infty. \end{cases}$$

Then define the hyperplanes by $H_i = \{v \in V : B(e_i, v) = 0\}$ and the linear maps by

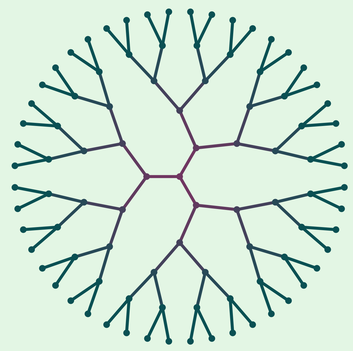
$$\sigma_i(v) = v - 2B(e_i, v)e_i,$$

which we note is the usual form of reflections in Euclidean geometry.

Buildings

A **building of type** (W, S) is a simplicial complex Δ, which is a union of subcomplexes called **apartments**, where each apartment is a copy of the **Coxeter complex** (or alternatively the Davis complex) for (W, S). With *chambers* defined to be the maximal simplices in Δ, the following hold:

1. Any two chambers are contained in a common apartment
2. If A and A' are arbitrary apartments, then there is an isomorphism $A \to A'$ which fixes $A \cap A'$ pointwise



For example, take the infinite dihedral group

$$W = \langle s, t \mid s^2 = t^2 = 1 \rangle \cong D_\infty,$$

whose Coxeter complex (and similarly its Davis complex) is the tessellation of the real line $\mathbb{E}^1$ under the action of W.

The 3-regular tree T_3 (shown to the left) is a building of type (W, S) when we take the system of apartments to be the collection of all bi-infinite lines in T_3. Each line segment corresponds to a chamber in the building and each path through the tree corresponds to a single Davis complex $\Sigma(W, S)$.

Observe that the first condition is satisfied since any two edges in the tree are contained in a common line. The second condition follows because we can trivially find a map between any pair of bi-infinite lines. Thus, T_3 is a building of type (W, S).

Acknowledgements

We want to thank our mentor, Benedict Lee, for his guidance and support in our readings and the Directed Reading Program at UCSB for this opportunity. Finally, we thank the contributors to the Typst typsetting language for making this poster a joy to create.

References

[1] A. Thomas, *Geometric and Topological Aspects of Coxeter Groups and Buildings*. European Mathematical Society, 2018.

[2] J. E. Humphreys, *Reflection groups and Coxeter groups*. Cambridge University Press, 1990.

Learning to Act: Methods of Reinforcement Learning

Alexandra Boog, Soros Qin
University of California, Santa Barbara

History

Reinforcement Learning (RL) emerged in the late 1970s from psychology, neuroscience, and control theory, focusing on hedonistic systems that maximize reward through trial and error. Today, RL powers real-world applications in gaming, web services, finance, and healthcare, with the goal of learning through interaction for better decision-making.

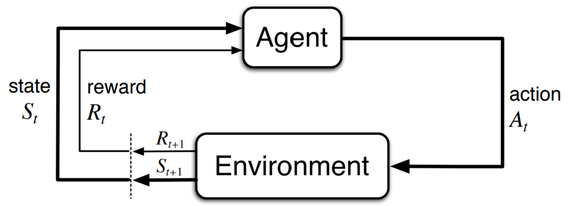


Figure 1: the action-agent interaction in a Markov Decision Process (MDP).

Background

Fundamentally, Reinforcement Learning uses mathematical methods to find the **optimal policy**, π^*, that an agent can follow to achieve a specific goal in an environment. This is done by having the agent and environment interact in a specific **action**, $a \in A$ from a **state**, $s \in S$, placing the agent in a new *state*, $s' \in S$. A policy π tells us which action to take in a given state, also known as a **state-action pair**. Each *state-action pair* gives us a different **return**, which is the total accumulated reward, r, that the agent seeks to maximize.

We assign a **value function** to each *state*, v_π, and *state-action pair*, q_π, that's dependent on a specific policy and we define one policy to be better than the other if it's expected return is greater. Using this logic, we can find the **optimal policy** by finding out which policy maximizes these *value functions*. As time progresses, the influence of future rewards diminishes. This is shown using a **discount factor**, $\gamma \in [0, 1]$, which reduces the contribution of rewards received further in the future.

Key Equations (Bellman optimality):

$$v^*(s) = \max_a \sum_{s',r} p(s', r \mid s, a)\big[r + \gamma\, v^*(s')\big],$$

$$q^*(s, a) = \sum_{s',r} p(s', r \mid s, a)\big[r + \gamma \max_{a'} q^*(s', a')\big].$$

We can represent this by a backup diagrams, which show how we gather information backwards form future states to update our *optimal policy*.

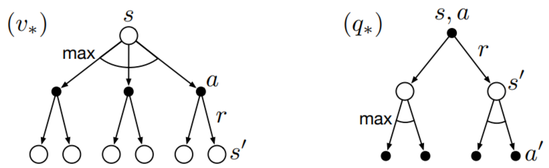


Figure 2: Backup Diagrams for v_* and q_*

Acknowledgements

We thank Jack Pfaffinger for his guidance as well as the UCSB Directed Reading Program for the opportunity to work on this project.

References

[1] Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction*. Cambridge, Massachusetts; London, England: The MIT press, 2018.

Dynamic Programming (DP)

Definition. A family of algorithms for computing optimal value functions and policies when the full MDP model is known. It uses the known transition/reward model $p(s', r \mid s, a)$ to perform *expected updates* (policy evaluation, iteration, value iteration). DP performs **expected updates** using the Bellman equations to refine value estimates. We begin with **iterative policy evaluation**, where an arbitrary v_0 is updated under a given policy π by applying the Bellman equation as the **Key Update:**

$$v_{k+1}(s) = \sum_a \pi(a \mid s) \sum_{s',r} p(s', r \mid s, a)\left[r + \gamma v_k(s')\right].$$

As this process repeats $v_k \longrightarrow v_\pi$. **Policy improvement** chooses actions that maximize expected return, aka what actions maximize $q_\pi(s, a)$. This gives us a new, *greedy* policy.

$$\pi'(s) = \arg\max_a \sum_{s',r} p(s', r \mid s, a)\left[r + \gamma v_\pi(s')\right].$$

Alternating these processes is **policy iteration** and it yields increasingly better policies with repetition until we eventually converge to the optimal policy, π^*.

$$\pi_0 \xrightarrow{\mathsf{E}} v_{\pi_0} \xrightarrow{\mathsf{I}} \pi_1 \xrightarrow{\mathsf{E}} v_{\pi_1} \xrightarrow{\mathsf{I}} \pi_2 \xrightarrow{\mathsf{E}} \cdots \xrightarrow{\mathsf{I}} \pi_* \xrightarrow{\mathsf{E}} v_*$$

Monte Carlo (MC)

Definition. Methods for estimating value functions by averaging sample returns from complete episodes, without requiring any model. The return G_t from time step t is defined as:

$$G_t = R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \cdots = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1}.$$

The value function is updated using the **Key Update (constant-α first-visit MC):**

$$V(s) \leftarrow V(s) + \alpha\big(G_t - V(s)\big),$$

where $\alpha > 0$ is the learning rate. There are two main types of MC Methods: **first-visit** and **every-visit** which estimate $v_\pi(s)$ following the first visit or all visits to s respectively.

Off-Policy

MC methods estimate value functions based on complete episodes of experience. This allows us to evaluate a target policy π, while generating episodes using a different behavior policy b. To correct for the difference, we use **importance sampling**, which re-weights returns based on how likely we are to choose an action-state pair under π versus b.

$$\text{importance sampling ratio: } \rho_t = \prod_{k=t}^{T-1} \frac{\pi(A_k \mid S_k)}{b(A_k \mid S_k)}$$

The basic form, *ordinary importance sampling*, scales each return by ρ_t and averages the results. Although unbiased, this method can have high variance if b differs greatly from π. Although biased, *weighted importance sampling* solves this and reduces variance by normalizing the weights.

$$\text{ordinary: } V(s) = \frac{\sum_{t \in \mathcal{T}(s)} \rho_t\, G_t}{|\mathcal{T}(s)|} \qquad \text{weighted: } V(s) = \frac{\sum_{t \in \mathcal{T}(s)} \rho_t\, G_t}{\sum_{t \in \mathcal{T}(s)} \rho_t}$$

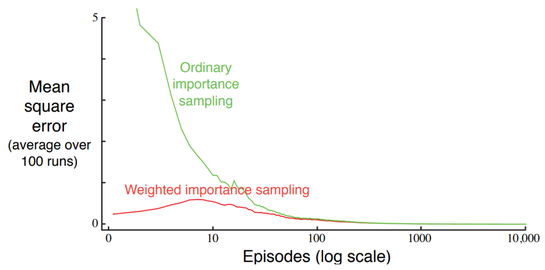
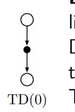


Figure 3: Weighted importance sampling produces lower error estimates of the value of a single blackjack state (normal rules) from off-policy episodes.

Temporal Difference (TD)



Definition. A hybrid of MC and DP that learns directly from experience like MC and **bootstraps**—updating estimates using other estimates— like DP without requiring a model. While MC waits until the end of an episode to update, TD updates after a specified number of steps. The special case, TD(0), updates after each step, but the more general form, TD(n) updates after n steps.

Key Update (TD(n)):

$$V_{t+n}(S_t) = V_{t+n-1}(S_t) + \alpha\left[G_{t:t+n} - V_{t+n-1}(S_t)\right], \quad 0 \le t < T$$

Note that the quantity in brackets is the **TD error**, δ_t, which is the difference between successive estimates. This method of updating is incremental which allows for quick updates.

n-step SARSA (on-policy)

We estimate the action-value function $q_\pi(s, a)$ under the current policy π, using the same n-step update structure we described above for state values.

$$Q_{t+n}(S_t, A_t) = Q_{t+n-1}(S_t, A_t) + \alpha\left[G_{t:t+n} - Q_{t+n-1}(S_t, A_t)\right], \quad 0 \le t < T$$

Q-Learning (off-policy)

We directly approximate q_*, independent of the policy being followed, by using the maximum estimated value at the next state (from TD(0) specifically).

$$Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha\left[R_{t+1} + \gamma \max_a Q(A_{t+1}, a) - Q(S_t, A_t)\right].$$

n-step Expected SARSA (on-policy)

We average over the expected value over the distribution of future *state-action pairs* instead of the max. Thus assuming $t + n < T$ and G_t is the return function,

$$G_{t:t+n} = R_{t+1} + \gamma R_{t+2} + \cdots + \gamma^{n-1} R_{t+n} + \gamma^n \sum_a \pi(a \mid S_{t+n}) Q(S_{t+n}, a)$$

$$Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha\left[G_{t:t+n} - Q(S_t, A_t)\right]$$

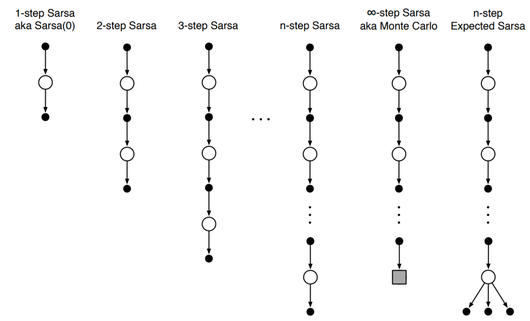


Figure 4: The backup diagrams for the spectrum of n-step methods for state–action values.

Trade-Offs

Dynamic Programming (DP)
Pros: Converges to exact solution, uses full model.
Cons: Requires known dynamics, computationally expensive (not scalable).

Monte Carlo (MC)
Pros: Model-free, easy to implement.
Cons: High variance, needs complete episodes (can be slow), no bootstrapping.

Temporal Difference (TD)
Pros: Model-free, bootstraps, lower variance (and faster) than MC, online and incremental.
Cons: Biased updates, analysis more complex than MC.

Mathematical Gauge Theory

A Gauge Theoretic Approach to Electromagnetism

Faizan Hussaini[1] Jinwei Zou[1]

[1]Department of Mathematics, University of California, Santa Barbara

0. What is Gauge Theory

In the most general sense, Gauge Theory is the study of *gauge-theoretic equations*. That is, differential equations involving *connections* or *sections* on various types of *bundles*. These equations provide valuable information about the geometry and topology of the underlying *manifold*.

1. Mathematical Prerequisites

Here we shall define the terms in the above definition:

1. A smooth manifold is a manifold with a maximal smooth atlas. (A smooth surface)
2. Let M be a smooth manifold. Let $p \in M$. The tangent space of p in M, T_pM, is defined to be the set of all possible velocities of curves passing through p. The tangent bundle of M is defined as $\sqcup_{p\in M} T_pM$.
3. A Lie group is a smooth manifold with a group structure such that: each point g in the manifold corresponds to a translation diffeomorphism which sends the identity point e to g. These diffeomorphisms form a group.
4. If M is a smooth manifold, and G is a Lie group, then a smooth action of G on M is a smooth map $\cdot : G \times M \to M$ such that $e \cdot p = p$ and $g \cdot (h \cdot p) = (gh) \cdot p$
5. Suppose G acts on M and let X be a left-invariant vector field on G. Let $\phi_p \colon G \to M$ be given by $\phi_p(g) = g \cdot p$. Then a fundamental vector field on M is a vector field of the form
$$\tilde{X}_p = D_e\phi_p(X_e)$$
6. Let U, E, M be smooth manifolds and let $\pi \colon E \to M$ be a surjective differentiable map. If for every $x \in M$, $\pi^{-1}(x) \cong U$, then U is called the general fiber of π ($\cong$ means diffeomorphic). If, for every $x \in M$, we can find a neighbourhood V of x such that $\pi^{-1}(V) \cong U \times V$, then (E, π, M, U) is called a fiber bundle.
7. $U \to E \xrightarrow{\pi} M$ is called a principal U-bundle if it satisfies the following: U is a Lie group, the action of U preserves each fiber on E, and there is a bundle atlas of "G-equivariant" charts $E \to M \times U$.
8. The vertical tangent space, V_p, of a principal bundle is the tangent space to the fiber. $V_p = \ker(D_p\pi)$.
9. The horizontal tangent space, H_p, is a subspace of T_pE such that $H_p \bigoplus V_p = T_pP$. Note that horizontal tangent spaces are not defined uniquely. Each horizontal tangent space determines a different connection on our bundle, and each connection correspond to a different gauge. This is the main tool used in our exploration of Gauge theory. We look at invariants under different gauges to find symmetries in physics.
10. A connection form is a $\mathfrak{u}(1)$-valued 1-form
$$\omega \in \Omega^1(P, \mathfrak{u}(1))$$
satisfying
$$(r_g)^*\omega = \mathrm{Ad}_{g^{-1}} \circ \omega = \omega \quad \text{and} \quad \omega\big(\widetilde{X}_q\big) = X\,,$$
for each fundamental vector field $\widetilde{X}_q$ generated by $X \in \mathfrak{u}(1)$. $\ker\omega_q = \mathrm{Hor}_qP$ defines the horizontal subspace at $q \in P$.

2. Examples

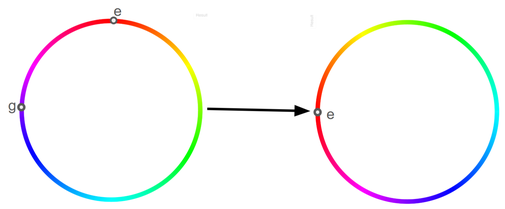


Figure 1: S^1 as a Lie group

The picture above shows a left translation diffeomorphism on the Lie group S^1. This is the Lie group in the principal bundle we use to study Electromagnetism.

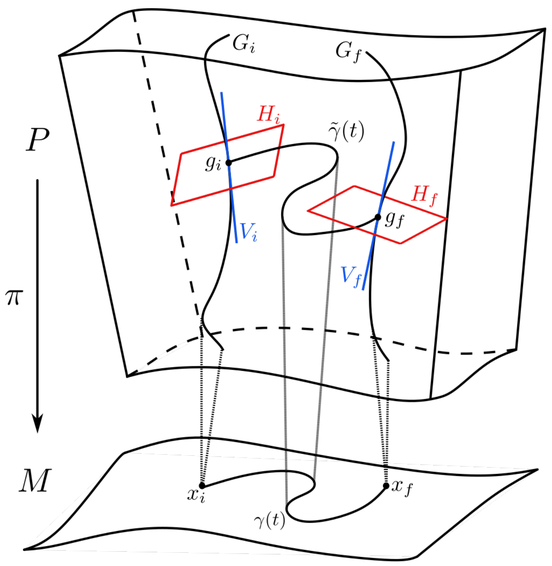


Figure 2: Principal Bundle

The figure above shows a principal bundle P over a manifold M. The structure group or general fiber of this bundle is a one-dimensional Lie group diffeomorphic to the strands G_i and G_f. H_i and H_f are two horizontal tangent spaces. Notice that they need not be orthogonal to V_i and V_f.

The Hopf Fibration

One of the most famous and useful fiber bundles in maths is the Hopf fibration: the bundle

$$S^1 \to S^3 \xrightarrow{\pi} S^2$$

with $\pi(z_0, z_1) = (2z_0z_1^*, z_0z_0^* - z_1z_1^*)$

3. Cartan's Theorem

Throughout our reading, this unbelievable theorem was used several times to prove critical results. The statement is the following:
Let G be a Lie group and suppose that $H \subseteq G$ is a subgroup in the algebraic sense. Then H is an embedded Lie subgroup if and only if H is a closed set in the topology of G.

4. E&M as a Gauge Theory

Classical electromagnetism, unified by Maxwell's equations, finds its mathematical expression through the framework of gauge theory. A formalism that encodes the electromagnetic field as a manifestation of local symmetry (specifically, invariance under the $U(1)$ gauge group) governing the phase freedom of charged fields.

5. The Electromagnetism Connection

Electromagnetism is formulated on a 4D Lorentzian manifold $(M, g_{\mu\nu})$, where $g_{\mu\nu}$ is the metric tensor with signature $(-,+,+,+)$. The gauge structure is a principal bundle $P \xrightarrow{\pi} M$ with structure group $G = \mathrm{U}(1) \cong S^1$, equipped with a connection form $\omega \in \Omega^1(P, \mathfrak{u}(1))$.

$$\omega_q : T_qP \longrightarrow \mathfrak{u}(1) \cong i\mathbb{R}\,.$$

In conclusion, this is $(P, \pi, M, \mathrm{U}(1))$, a principal U(1)-bundle over the Lorentzian manifold $(M, g_{\mu\nu})$, with right action

$$r_g : P \to P, \quad q \mapsto q \cdot g, \quad g \in \mathrm{U}(1).$$

The connection form ω splits $TP = \mathrm{Vert}(P) \oplus \mathrm{Hor}(P)$, where $\mathrm{Hor}(P)$ is annihilated by ω.

6. Electromagnetic Potential

Local sections $s : U \to P$ define the gauge potential A_μ, a $\mathfrak{u}(1)$-valued 1-form on U, via the pullback

$$s^*\omega = iA = iA_\mu dx^\mu \in \Omega^1(U, i\mathbb{R})$$

where $A_\mu \in C^\infty(U)$. For $\phi(x) \in C^\infty(U)$, changing the section to $s'(x) = s(x) \cdot g(x) = s(x)e^{-i\phi(x)}$ induces the gauge transformations:

$$\begin{aligned} s'^*(\omega) &= g(s^*\omega)g^{-1} + dg \cdot g^{-1} \\ &= iA + d(e^{i\phi}) \cdot e^{-i\phi} \\ &= i(A + d\phi)\,. \end{aligned}$$

In short, $A \mapsto A + d\phi$, i.e., $A_\mu(x) \mapsto A_\mu(x) + \partial_\mu\phi(x)$. We define covariant derivatives to compensate for local phase changes:

$$D_\mu = \partial_\mu - iA_\mu \quad \Longrightarrow \quad D_\mu\psi \mapsto e^{i\phi(x)}D_\mu\psi\,.$$

7. Electromagnetic Field Tensor

The curvature 2-form is $\Omega = d\omega + [\omega, \omega]$. Since U(1) is abelian, $[\omega, \omega] = 0$, so

$$\Omega = d\omega.$$

Given a local section $s : M \to P$, the pullback of ω gives the electromagnetic potential $A = s^*\omega$, a 1-form on M. The field strength is then

$$F = s^*\Omega = dA,$$

which in local coordinates is

$$F = \frac{1}{2}F_{\mu\nu}\,dx^\mu \wedge dx^\nu, \quad \text{with} \quad F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu.$$

Equivalently,

$$iF = i \cdot dA = d(s^*\omega) = s^*(d\omega) = s^*(\Omega)\,.$$

The electromagnetic field tensor $F_{\mu\nu}$ is considered geometrically as the curvature of a connection on a principal U(1)-bundle.

8. Visual Schematics

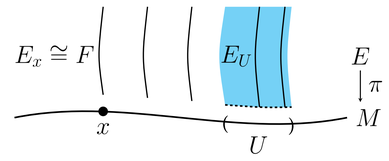


Figure 3: A schematic of the fiber bundle $\pi : E \to M$.

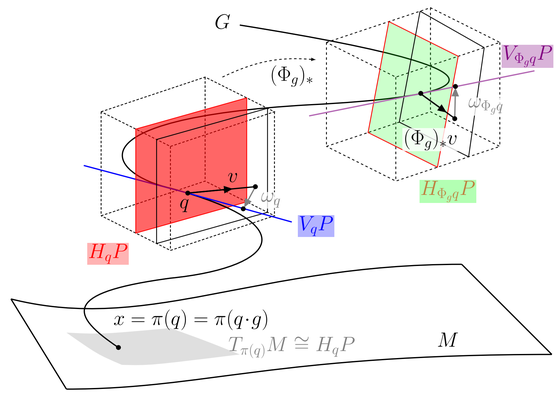


Figure 4: A schematic of the connection 1-form.

Figure 4 shows the visualization of

$$\omega_{\Phi_g q}\left((\Phi_g)_* v\right) = \mathrm{Ad}_g(\omega_q(v)) = \Phi_g^{-1}(\omega_q(v))\Phi_g\,, \forall\, v \in T_qP\,.$$

9. Yang-Mills Theory

The potential in electromagnetism takes the form of a connection on a principal U(1) bundle. We can generalize this picture to other interactions in the Standard Model. If we replace U(1) with an arbitrary compact Lie Group G, the resulting field theories are called Yang-Mills theories. The weak interaction corresponds to the choice $G = \mathrm{SU}(2)$ and the strong interaction corresponds to $G = \mathrm{SU}(3)$.

10. Acknowledgements

We thank the UCSB Directed Reading Program for this opportunity, and especially thank Edward Chen for directing our reading.

11. References

- Nicholas James Marks Ford. Electromagnetism as a gauge theory.
- Mark J.D. Hamilton. *Mathematical Gauge Theory.* Springer International Publishing, 2017.
- Rupert Way. Introduction to connections on principal fibre bundles.

Modular Forms and Elliptic Curves

Mingxin Du

University of California - Santa Barbara

Weierstrass $\wp$-Function

Let $\Lambda \subset \mathbb{C}$ be a lattice. The *Weierstrass $\wp$-function* associated to Λ is

$$\wp(z;\Lambda) = \frac{1}{z^2} + \sum_{\omega\in\Lambda\setminus\{0\}} \left(\frac{1}{(z-\omega)^2} - \frac{1}{\omega^2}\right).$$

It is an even, Λ-periodic meromorphic function on $\mathbb{C}$ with a double pole (zero residue) at each $\omega \in \Lambda$ and no other poles.

Parity and Zeros of $\wp'$

For any lattice Λ,

$$\wp(-z) = \wp(z), \qquad \wp'(-z) = -\wp'(z),$$

and on each fundamental parallelogram there are exactly three simple zeros of $\wp'$, which occur at the nonzero half–periods of Λ. Moreover $\wp'$ has a triple zero at the origin.

My Favorite Problem

Let $\Lambda = \Lambda_i$, the derivative $\wp'$ of the corresponding Weierstrass $\wp-$function has

- a triple pole at 0
- simple zeros at $\frac{1}{2}$, $\frac{i}{2}$, and $\frac{1+i}{2}$

Since $i\Lambda = \Lambda$ one shows

$$\wp(iz) = \wp_{i\Lambda}(iz) = i^{-2}\wp(z) = -\wp(z).$$

Show that

- $\wp(\frac{1+i}{2}) = 0$
- $\wp'(\frac{1+i}{2}) = 0$, then $\frac{1+i}{2}$ is a double zero, and the only zero of $\wp$ on $\mathbb{C}/\Lambda$
- after an appropriate scaling $\mathbb{C}/m\Lambda$ the Weierstrass equation becomes

$$y^2 = 4x(x-1)(x+1)$$

Solution

Since $i\Lambda = \Lambda$ shows

$$\wp(iz) = \wp_{i\Lambda}(iz) = i^{-2}\wp(z) = -\wp(z).$$

Plug in $z = \frac{1+i}{2}$, we have

$$\wp(i\frac{1+i}{2}) = \wp(\frac{i-1}{2}) = -\wp(\frac{1+i}{2}).$$

But $\frac{i-1}{2}$ differs from $\frac{1+i}{2}$ by a lattice vector, so $\wp(\frac{i-1}{2}) = \wp(\frac{1+i}{2})$. Hence

$$\wp(\frac{1+i}{2}) = -\wp(\frac{1+i}{2}) \Rightarrow \wp(\frac{1+i}{2}) = 0.$$

By hypothesis $\wp'$ has simple zeros at $\frac{1}{2}$,$\frac{i}{2}$, and $\frac{1+i}{2}$. In particular $\wp'(\frac{1+i}{2})$=0, so $\frac{1+i}{2}$ is a zero of $\wp$ of order at least two. One checks there are no other zeros mod Λ.

Complex-conjugation invariance $\Lambda = \overline{\Lambda}$ gives $\wp(\overline{z}) = \overline{\wp(z)}$, so

$$\wp(\frac{1}{2}) \in \mathbb{R}, \text{ and since } \wp(\frac{i}{2}) = -\wp(\frac{1}{2}),\ \wp(\frac{i}{2}) \in \mathbb{R}.$$

Computing some dominant terms for $\wp(\frac{1}{2})$ and $\wp(\frac{i}{2})$, one sees that $\wp(\frac{1}{2}) > 0$. Denote

$$e = \wp(\frac{1}{2}) > 0, \wp(\frac{i}{2}) = -e.$$

Then the usual form is

$$y^2 = 4(x-e_1)(x-e_2)(x-e_3)$$

where $e_1 = \wp(\frac{1+i}{2}) = 0, e_2 = \wp(\frac{1}{2}) = e, e_3 = \wp(\frac{i}{2}) = -e$. Hence

$$y^2 = 4x(x-e)(x+e) = 4x(x^2-e^2).$$

Since $\wp_{m\Lambda}(mz) = m^{-2}\wp_\Lambda(z)$ and $\wp'_{m\Lambda}(mz) = m^{-3}\wp'_\Lambda(z)$, we choose

$$X = \frac{x}{e},\ Y = \frac{y}{e^{3/2}}.$$

Then $y^2 = e^3 \cdot Y^2$ and $4x(x^2-e^2) = 4e^3X(X^2-1)$. By canceling e^3, we have

$$Y^2 = 4X(X^2-1) = 4X(X-1)(X+1).$$

To make the nonzero roots ± 1, we need

$$m^{-2}e = 1 \Rightarrow m = \sqrt{e} = \sqrt{\wp(\frac{1}{2})}.$$

Hence the torus $\mathbb{C}/\sqrt{\wp(\frac{1}{2})}\Lambda$ has Weierstrass equation

$$Y^2 = 4X(X^2-1) = 4X(X-1)(X+1).$$

Weierstrass Differential Equation

Define the invariants

$$g_2 = 60\sum_{\omega\in\Lambda\setminus\{0\}} \omega^{-4}, \qquad g_3 = 140\sum_{\omega\in\Lambda\setminus\{0\}} \omega^{-6}.$$

Then $\wp(z) = \wp(z;\Lambda)$ satisfies

$$\big(\wp'(z)\big)^2 = 4\,\wp(z)^3 - g_2\,\wp(z) - g_3.$$

The map $z \mapsto \big(\wp(z), \wp'(z)\big)$ identifies the torus $\mathbb{C}/\Lambda$ with the elliptic curve

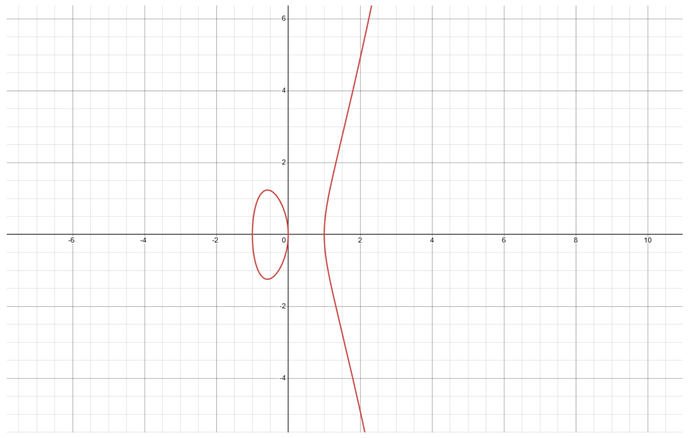
$$y^2 = 4x^3 - g_2x - g_3\,.$$

Reference

Fred Diamond and Jerry Shurman, *A First Course in Modular Forms*, Graduate Texts in Mathematics, Vol. 228, Springer, 2005.

Acknowledgments

Thank you to the UCSB Directed Reading Program and to my mentor Christine Alar for making this project possible.



$Y^2 = 4X(X-1)(X+1)$

WIGNER'S SEMI-CIRCLE LAW

Ray Qian

University of California Santa Barbara

Wigner Matrix and Semi-circle Law

Start with two independent families of independent and identically distributed (i.i.d.) zero mean, real-valued random variables $\{Z_{i,j}\}_{1\le i<j}$ and $\{Y_i\}_{1\le i}$, such that $EZ_{1,2}^2 = 1$ and, for all integers $k \ge 1$,

$$r_k := \max\left(E\left|Z_{1,2}\right|^k, E\left|Y_1\right|^k\right) < \infty$$

Consider the (symmetric) $N \times N$ matrix X_N with entries

$$X_N(j,i) = X_N(i,j) = \begin{cases} Z_{i,j}/\sqrt{N}, & \text{if } i < j \\ Y_i/\sqrt{N}, & \text{if } i = j \end{cases}$$

We call such a matrix a **Wigner matrix**, and if the random variables $Z_{i,j}$ and Y_i are Gaussian, we use the term **Gaussian Wigner matrix**.
Let λ_i^N denote the (real) eigenvalues of X_N, with $\lambda_1^N \le \lambda_2^N \le \cdots \le \lambda_N^N$, and define the empirical distribution of the eigenvalues as the (random) probability measure on $\mathbb{R}$ defined by

$$L_N = \frac{1}{N}\sum_{i=1}^{N} \delta_{\lambda_i^N}$$

Define the **semicircle distribution** (or law) as the probability distribution $\sigma(x)dx$ on $\mathbb{R}$ with density

$$\sigma(x) = \frac{1}{2\pi}\sqrt{4-x^2}\mathbf{1}_{|x|\le 2}$$

The following theorem can be considered the starting point of random matrix theory (RMT).

Theorem 1 (Wigner's Semi-circle Law) *For a Wigner matrix, the empirical measure L_N converges weakly, in probability, to the semicircle distribution.*
More precisely, Theorem 1 asserts that for any $f \in C_b(\mathbb{R})$, and any $\varepsilon > 0$,

$$\lim_{N\to\infty} P\left(\left|\langle L_N, f\rangle - \langle \sigma, f\rangle\right| > \varepsilon\right) = 0.$$

Define the moments $m_k := \left\langle \sigma, x^k \right\rangle$. Recall the **Catalan numbers**

$$C_k = \frac{\binom{2k}{k}}{k+1} = \frac{(2k)!}{(k+1)!k!}$$

One can easily check, for all integers $k \ge 1$,

$$m_{2k} = C_k, \quad m_{2k+1} = 0$$

Lemma 2 *$\beta_k = C_k < 4^k$. Further, the generating function $\hat{\beta}(z) := 1 + \sum_{k=1}^{\infty} z^k \beta_k$ satisfies, for $|z| < 1/4$,*

$$\hat{\beta}(z) = \frac{1-\sqrt{1-4z}}{2z}$$

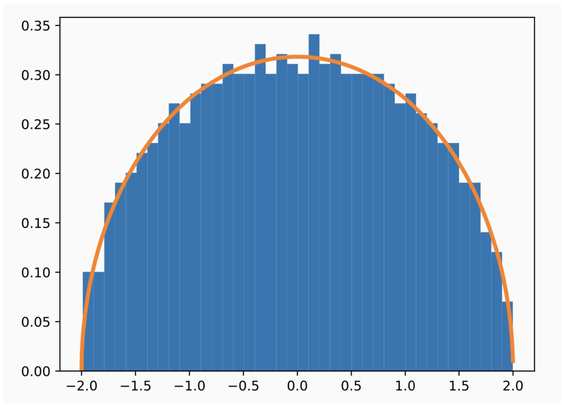


Figure 1: density of semicircle law

Stieltjes Transform

Definition 3 *Let μ be a positive, finite measure on the real line. The **Stieltjes transform** of μ is the function*

$$S_\mu(z) := \int_{\mathbb{R}} \frac{\mu(dx)}{x-z}, z \in \mathbb{C}\backslash\mathbb{R}$$

Note that for $z \in \mathbb{C}\backslash\mathbb{R}$, both the real and imaginary parts of $1/(x-z)$ are continuous bounded functions of $x \in \mathbb{R}$ and, further, $\left|S_\mu(z)\right| \le \mu(\mathbb{R})/|\Im z|$. These crucial observations are used repeatedly in what follows. Stieltjes transforms can be inverted. In particular, one has

Theorem 4 *For any open interval I with neither endpoint on an atom (element with nonzero mass) of μ,*

$$\begin{aligned} u(I) &= \lim_{\varepsilon\to 0}\frac{1}{\pi}\int_I \frac{S_\mu(\lambda+\mathrm{i}\varepsilon) - S_\mu(\lambda-\mathrm{i}\varepsilon)}{2\mathrm{i}}d\lambda \\ &= \lim_{\varepsilon\to 0}\frac{1}{\pi}\int_I \Im S_\mu(\lambda+\mathrm{i}\varepsilon)d\lambda \end{aligned}$$

Theorem 4 allows for the reconstruction of a measure from its Stieltjes transform. Further, one has the following.

Theorem 5 *Let $\mu_n \in M_1(\mathbb{R})$ be a sequence of probability measures.*

a) If μ_n converges weakly to a probability measure μ then $S_{\mu_n}(z)$ converges to $S_\mu(z)$ for each $z \in \mathbb{C}\backslash\mathbb{R}$.

b) If $S_{\mu_n}(z)$ converges for each $z \in \mathbb{C}\backslash\mathbb{R}$ to a limit $S(z)$, then $S(z)$ is the Stieltjes transform of a sub-probability measure μ, and μ_n converges vaguely to μ.

c) If the probability measures μ_n are random and, for each $z \in \mathbb{C}\backslash\mathbb{R}$, $S_{\mu_n}(z)$ converges in probability to a deterministic limit $S(z)$ that is the Stieltjes transform of a probability measure μ, then μ_n converges weakly in probability to μ.

(We recall that μ_n converges vaguely to μ if, for any continuous function f on $\mathbb{R}$ that decays to 0 at infinity, $\int f d\mu_n \to \int f d\mu$. Recall also that a positive measure μ on $\mathbb{R}$ is a sub-probability measure if it satisfies $\mu(\mathbb{R}) \le 1$.)

Theorem 6 *For $z \in \mathbb{C}$ such that $z \notin [-2,2]$, the Stieljes transform $S(z)$ of the semicircle law equals*

$$S(z) = \int \frac{1}{\lambda - z}\sigma(d\lambda) = \frac{-z+\sqrt{z^2-4}}{2} \tag{1}$$

Proof We use complex analysis to prove it. Notice that $\sigma(x) = \frac{1}{2\pi}\sqrt{4-x^2}\chi_{|x|\le 2}$. We have

$$S(z) = -\frac{1}{2\pi}\int_{-2}^{2}\frac{\sqrt{4-\lambda^2}}{z-\lambda}d\lambda.$$

Taking the branch such that $\Im\sqrt{z^2-4} \ge 0$. Let $\lambda = 2\cos(\theta)$ with $\theta \in [0,\pi]$. Then,

$$S(z) = \frac{1}{2\pi}\int_0^\pi \frac{4\sin^2\theta}{z-2\cos\theta}d\theta.$$

Let $u = e^{i\theta}, (0 \le \theta \le \pi)$. We have

$$S(z) = -\frac{1}{4\pi i}\oint_{|u|=1}\frac{(u-u^{-1})^2}{u^2-uz+1}du.$$

There are three poles, which are $u_\pm = \frac{z\pm\sqrt{z^2-4}}{2}$ and $u' = 0$. The choice of a branch makes $|u_-| < 1 < |u_+|$. Since

$$\text{Res}_{u=u_-} = -\sqrt{z^2-4} \quad \text{and} \quad \text{Res}_{u=0} = z,$$

by the residue theorem,

$$S(z) = -\frac{2\pi i}{4\pi i}\left(\text{Res}_{u=u_-} + \text{Res}_{u=0}\right) = \frac{-z+\sqrt{z^2-4}}{2} \qquad \square$$

Proof for Gaussian Wigner Matrices

We prove Theorem 1 when X_N is a Gaussian Wigner matrix. Recall Stein's identity.

Lemma 7 (Stein) *If ξ is a zero mean Gaussian random variable, then for f differentiable, with polynomial growth of f and f',*

$$E(\zeta f(\zeta)) = E\left(f'(\zeta)\right)E\left(\zeta^2\right)$$

Define next the matrix $\Delta_N^{i,k}$ as the symmetric $N \times N$ matrix satisfying

$$\Delta_N^{i,k}(j,l) = \begin{cases} 1, & (i,k) = (j,l) \text{ or } (i,k) = (l,j) \\ 0, & \text{otherwise} \end{cases}$$

Then, with X an $N \times N$ symmetric matrix,

$$\frac{\partial}{\partial X(i,k)}\mathbf{S}_X(z) = -\mathbf{S}_X(z)\Delta_N^{i,k}\mathbf{S}_X(z)$$

Therefore,

$$\begin{aligned} \frac{1}{N}E\,\mathrm{tr}\,\mathbf{S}_{X_N}(z) &= -\frac{1}{z} + \frac{1}{z}\frac{1}{N}E\left(\mathrm{tr}\,X_N\mathbf{S}_{X_N}(z)\right) \\ &= -\frac{1}{z} - \frac{1}{z}E\left[\left\langle L_N, (x-z)^{-1}\right\rangle^2\right] - \frac{1}{zN}\left\langle \bar{L}_N, (x-z)^{-2}\right\rangle \\ &\quad - \frac{1}{zN^2}\sum_i\left(\left(EY_i^2-2\right)E\mathbf{S}_{X_N}(z)(i,i)^2\right) \end{aligned}$$

It follows

$$\left|E\left[\left\langle L_N, (x-z)^{-1}\right\rangle^2\right] - \left\langle \bar{L}_N, (x-z)^{-1}\right\rangle^2\right| \to_{N\to\infty} 0.$$

This, and the boundedness of $1/(z-x)^2$ for a fixed z as above, imply the existence of a sequence $\varepsilon_N(z) \to_{N\to\infty} 0$ such that, letting $\bar{S}_N(z) := N^{-1}E\,\mathrm{tr}\,S_{X_N}(z)$, one has

$$\bar{S}_N(z) = -\frac{1}{z} - \frac{1}{z}\bar{S}_N(z)^2 + \varepsilon_N(z)$$

Thus any limit point $s(z)$ of $\bar{S}_N(z)$ satisfies

$$s(z)(z+s(z)) + 1 = 0 \tag{2}$$

For all $z \in \mathbb{C}$, with a suitable choice of the branch of the square-root,

$$s(z) = -\frac{1}{2}\left[z - \sqrt{z^2-4}\right].$$

Comparing with equation 1, one deduces that $s(z)$ is the Stieltjes transform of the semicircle law σ. It follows that $S_{L_N}(z)$ converges in probability to $s(z)$, solution of equation 2, for all $z \in \mathbb{C}\backslash\mathbb{R}$. The proof is completed by using part c) of Theorem 5.

Acknowledgements

We thank Hongjie Zeng for his guidance as well as the UCSB Directed Reading Program for the opportunity to work on this project.

References

[AGZ09] Greg W. Anderson, Alice Guionnet, and Ofer Zeitouni. *An Introduction to Random Matrices*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2009.