Application of Elliptic Curves to Fermat's Last Theorem

Elise Alvarez-Salazar, Caroline Baldan, and Kelly Stump

2022 Mathematics Directed Reading Program, University of California - Santa Barbara

Abstract

For this year's Directed Reading Program, we studied elliptic curves and methods for finding all their rational solutions. The three theorems about to be mentioned all tell us that the abelian group over $E(\mathbb{Q})$ has a rich group structure. Using this knowledge, we tackle the specific case of $n = 4$ of Fermat's Last Theorem.

Preliminary Information

Definition: An elliptic curve over $\mathbb{Q}$ is a smooth cubic projective curve E defined over $\mathbb{Q}$, with at least one rational point $\mathscr{O} \in E(\mathbb{Q})$ that we call the *origin*.

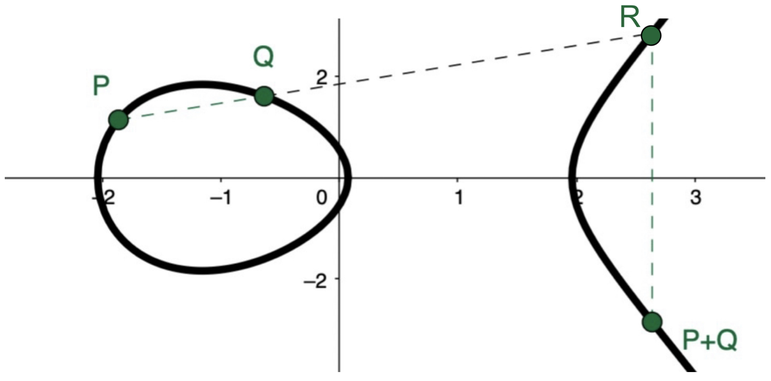
We will focus on elliptic curves of Weierstrass Form:

$$y^2 = x^3 + Ax + B \text{ where } A, B \in \mathbb{Z}$$

Defining P + Q

The operator for $E(\mathbb{Q})$ shall be defined as follows:

For $P, Q \in E(\mathbb{Q})$, where $P \neq Q$, we find the secant line which intersects both P and Q, $Y : y = ax + b$. Solving for the third point of intersection of Y with our curve E, labelled R, we see that $P + Q$ is the reflection of R over the x-axis.



For the case where $P = Q$, we consider the tangent line rather than the secant and a similar procedure follows to find $2P$. Note that every point has an inverse and our identity is the point at infinity, $\mathscr{O}$. Thus, we see that for this defined $+$ operator, we generate an abelian group on $E(\mathbb{Q})$.

Important Theorems

Mordell-Weil draws further conclusions about the previously created abelian group structure, stated below:

$E(\mathbb{Q})$ is a finitely generated abelian group. In other words, there are points $P_1, ..., P_n$ such that any other point $Q \in E(\mathbb{Q})$ can be expressed as a linear combination

$$Q = a_1P_1 + ... + a_nP_n$$

for some $a_i \in \mathbb{Z}$

From this theorem, and facts we know concerning finitely generated abelian groups, we find that:

$$E(\mathbb{Q}) \cong E(\mathbb{Q})_{\text{torsion}} \oplus \mathbb{Z}^{R_E}$$

Continuing on, we will refer to R_E as the *rank*. We can reach further conclusions about the group structure created over $E(\mathbb{Q})_{\text{torsion}}$ with Mazur's theorem stated in [**1**] as Thm 2.4.2.

Finding Rational Solutions

The natural continuation of the process of finding rational solutions for E is to next explore methods to calculate $E(\mathbb{Q})_{\text{torsion}}$ and $\mathbb{Z}^{R_E}$.

Specifically for calculating $E(\mathbb{Q})_{\text{torsion}}$, we have a theorem from Nagell-Lutz:

Let $E/\mathbb{Q}$ be an elliptic curve with Weierstrass equation $y^2 = x^3 + Ax + B$ where $A, B \in \mathbb{Z}$ Then, every torsion point $P \neq \mathcal{O}$ of E satisfies:

(1) The coordinates of P are integers, i.e. $x(P), y(P) \in \mathbb{Z}$.
(2) If P is a point of order $n \geq 3$ then $4A^3 + 27B^2$ is divisible by $y(P)^2$.
(3) If P is of order 2 then $y(P) = 0$ and $x(P)^3 + Ax(P) + B = 0$.

We have come up with two methods from our readings for trying to calculate the rank. The first uses Theorem 2.6.4 in [**1**]. And the other possible solution is found in section 2.9 of [**1**].

Scan the following QR code to be taken to our algorithm that will find the torsion points of assorted elliptic curves:

Example of Finding Rational Solutions

Let us consider the elliptic curve $E : y^2 = x^3 - x$. Applying our code to E, we see that $(0, 0)$, $(1, 0)$, $(-1, 0)$ and the point at infinity make up $E(\mathbb{Q})_{\text{torsion}}$, where each non-identity element has order 2. We find that the discriminant $\Delta_E = 64$. Thus, the only prime of bad reduction to consider is $p = 2$. We determine that 2 is of multiplicative bad reduction. Thus, by Thm. 2.6.4 in [**1**], we see that

$$R_E \leq m + 2a - 1 = 0$$

Thus, $E(\mathbb{Q}) \cong \mathbb{Z}_2 \times \mathbb{Z}_2$.

Fermat's Last Theorem (n = 4)

Problem Statement: Let $n = 4$. Are there any solutions to $a^n + b^n = c^n$ where $a, b, c \in \mathbb{Z}$ with $abc \neq 0$?

Solution: We claim that there are no non-trivial solutions. We are given the equation $a^4 + b^4 = c^4$, when $x = \frac{2(b^2+c^2)}{a^2}$ and $y = \frac{4b(b^2+c^2)}{a^3}$ are substituted in, we get the elliptic curve $E : y^2 = x^3 - 4x$.

Applying our given algorithm to this elliptic curve, we find that $E(\mathbb{Q})_{\text{torsion}} = \{(0, 0), (2, 0), (-2, 0), \mathscr{O}\}$. Note that these torsion points correspond to trivial solutions of $a^4 + b^4 = c^4$.

For the free part, an attempt to bound the rank proves insufficient as the prime of bad reduction is additive. Thus, we move onto use of the algorithm in 2.9 of [**1**] which tells us that the rank is 0. Thus, $E(\mathbb{Q}) \cong E(\mathbb{Q})_{\text{torsion}}$.

Acknowledgements

We would like to thank the DRP team for organizing this year's program. We would also like to thank our DRP Mentor, Marcos Reyes, for guiding us through our project. He was a wonderful resource while reading through our elliptic curve texts and taught us well.

References

[1] Á. Lozano-Robledo, *Elliptic Curves, Modular Forms and their L-functions*, American Mathematical Soc., 2011

[2] J. Silverman, J. Tate, *Rational Points on Elliptic Curves*, Springer-Verlag, New York, 2015

Alexander Polynomial the Great

Alycia Doucette and Elizabeth Benda - Mentored by Melody Molander

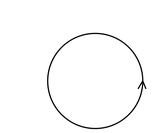
2022 Mathematics Directed Reading Program. University of California - Santa Barbara

Introduction

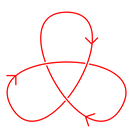
What is a **knot**? Simply speaking, a knot is a closed curve in space that does not intersect itself in any way. Knots have many applications to other fields of science and are fun for mathematicians to study. One of the main questions posed when studying knots is how to tell whether or not two different projections are the same knot. A tool that has developed as a way to distinguish two knots from each other is representing knots as polynomials. In this poster we will focus on one of the three major polynomial representations of knots, the Alexander polynomial.

Definitions

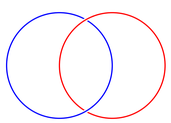
- **Projection:** A two-dimensional picture representation of a knot.
- **Orientation:** A direction in which you travel around the knot.
- **Crossing number:** The least number of crossings that occur in any projection of a particular knot.
- **Link:** A set of knotted loops tangled up together.
- **Unknot:** The unknot is also known as the trivial knot, and it looks as follows:



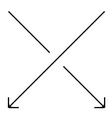
(a) Hi, I'm an oriented unknot!



(b) Hi, I'm an oriented trefoil!



(c) Hi, I'm a link!



(d) Hi, I'm a crossing!

The Alexander Polynomial

The Alexander polynomial was a method invented in 1928 as a way to represent knots and links as polynomial equations. It is an invariant for all representations of knots and links up to the same orientation. The Alexander polynomial is dependent on the orientation of the knot or link being assessed. The formula to compute the Alexander polynomial was refined by John Conway in 1969, and is now based on the following two rules:

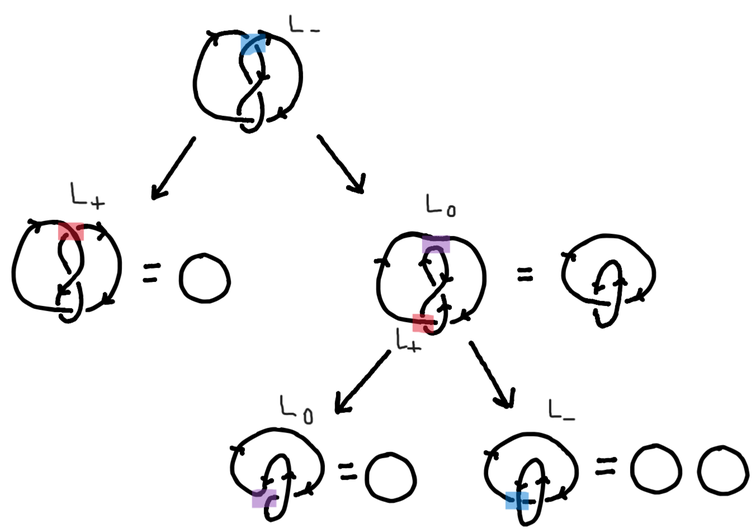
L_+ L_- L_0

$$\Delta(\bigcirc) = 1 \quad (1)$$

$$\Delta(L_+) - \Delta(L_-) + (t^{\frac{1}{2}} - t^{-\frac{1}{2}})\Delta(L_0) = 0 \quad (2)$$

The main tool used to compute the Alexander polynomial is called the **resolving tree**. The resolving tree is an easy way to break a knot down into a series of unknots and trivial links. In order to create the resolving tree, you choose one crossing of the knot, and determine whether it is an L_+, L_-, or L_0 crossing. From there, the chosen crossing is broken down into two new knots. These new knots are dependent on what type of crossing the original one is.

Resolving Tree of the Figure-Eight Knot



Alexander Polynomial of the Figure-Eight Knot

$$\Delta(L_+) - \Delta(L_-) + (t^{\frac{1}{2}} - t^{-\frac{1}{2}})\Delta(L_0) = 0$$

$$\Delta(L_+) = \Delta(\bigcirc) = 1$$

$$\Delta(L_0) = \Delta(\bigcirc \cup \bigcirc) - (t^{\frac{1}{2}} - t^{-\frac{1}{2}})\Delta(\bigcirc) = -(t^{\frac{1}{2}} - t^{-\frac{1}{2}})$$

$$\Rightarrow \Delta(L_-) = \Delta(L_+) + (t^{\frac{1}{2}} - t^{-\frac{1}{2}})\Delta(L_0) = 1 + (t^{\frac{1}{2}} - t^{-\frac{1}{2}})(-t^{\frac{1}{2}} + t^{-\frac{1}{2}})$$

$$= \boxed{3 - t - t^{-1}}$$

Since $3 - t - t^{-1} \neq 1$ we know that the figure-eight knot is not a projection of the unknot.

Other Polynomial Representations

The other polynomial representations we looked at were the Jones polynomial and the HOMFLY polynomial. The Jones polynomial, $V(t)$, is derived using three rules, and the base variable $t^{\frac{1}{2}}$. All prime knots with 9 or fewer crossings have a distinct Jones polynomial. The HOMFLY polynomial, unlike the other two, is multivariable. However, it does maintain a similar structure to that of the Alexander polynomial, using L_+, L_-, and L_0. Knots under both the HOMFLY and Jones polynomials are not affected by orientation, however, when computing the HOMFLY of a link, orientation between the two links does affect the result.

Consider the rules of the HOMFLY polynomial:

$$P(\bigcirc) = 1 \quad (1)$$

$$\alpha P(L_+) - \alpha^{-1} P(L_-) = z P(L_0) \quad (2)$$

The Alexander and Jones polynomials can be derived from the HOMFLY rules as follows:

$$\Delta(t) = P(\alpha = 1, z = t^{-\frac{1}{2}} - t^{\frac{1}{2}})$$

$$V(t) = P(\alpha = t^{-1}, z = t^{\frac{1}{2}} - t^{-\frac{1}{2}})$$

Conclusion

Each polynomial representation of knots has its own benefits and drawbacks. While the HOMFLY polynomial comes the closest to distinguishing between all knots and links, there is not currently any polynomial representation of knots that can completely distinguish all knots and links. Knots are the best!

References and Acknowledgements

It was fascinating to read and learn about how knots, simple strings in space, can be transformed into different polynomials. We would like to thank our graduate mentor Melody Molander, and the DRP, for creating this space for us to explore and grow our interests in mathematics.

Adams, Colin. The Knot Book. American Mathematical Society, 2004.

Introduction to Algebraic Number Theory

Robin Lee, Mr. Mulun Yin

University of California, Santa Barbara

UC SANTA BARBARA

Introduction

In this poster, we will overview the fundamentals of Algebraic Number Theory, focusing on the basic definitions of rings and fields, algebraic numbers, and algebraic integers.

Rings and Fields

As the most fundamental concept of Algebraic Number Theory, rings and fields are algebraic structures that contain two binary operations (addition and multiplication) with properties similar to those for integers $\mathbb{Z}$. In [1], we can define a **ring** as a non-empty set R with addition and multiplication. Assuming R is a ring, we mean it has the following characteristics:

- a set closed under addition $a+b \in R$ and multiplication $ab \in R$
- commutative under addition $a+b=b+a$
- associative under addition $a+(b+c)=(a+b)+c$ and multiplication $a(bc)=(ab)c$
- contains the additive identity $a+0=a, \forall a \in R$,for some $0 \in R$
- contains additive inverses: $\forall a \in R$, $\exists s \in R$ such that $a+s=0$
- contains the multiplicative identity $1*a=a*1=a, \forall a \in R$, for some $1 \in R$

Example of Rings: $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, $\mathbb{Z}[\sqrt{2}]$, $\mathbb{Z}[i]$

An element of the ring $\mathbb{Z}[\sqrt{2}]$ is $a+b\sqrt{2}$ where $a,b \in \mathbb{Z}$.
An element of the ring $\mathbb{Z}[i]$ is $a+bi$ where $a,b \in \mathbb{Z}$.
An example of the ring's addition and multiplication properties is:
$(1+\sqrt{2})+(2+\sqrt{2})=3+2\sqrt{2}$, and
$(1+\sqrt{2})*(2+\sqrt{2})=2+2\sqrt{2}+\sqrt{2}+2=4+3\sqrt{2}$

Similar to rings, **fields** not only contain the same properties as a ring, but also contain multiplicative inverses (in addition to additive inverses) and is commutative under multiplication. In other words, a field F is a unique configuration of a commutative ring that contains at least two elements such that every non-zero element in F is both commutative under addition and multiplication. Furthermore, a field contains a multiplicative inverse.

Example of Fields: $\mathbb{Z}_n$ where n is a prime and positive integer, $\mathbb{Q}$, $\mathbb{Q}(\sqrt{2})$, $\mathbb{Q}(i)$

An element of the field $\mathbb{Q}(\sqrt{2})$ is $a+b\sqrt{2}$ where $a,b \in \mathbb{Q}$.
An element of the field $\mathbb{Q}(i)$ is $a+bi$ where $a,b \in \mathbb{Q}$.

Algebraic Numbers and Minimal Polynomials

Diving deeper into our understanding of fields and rings, it is imperative we first overview an essential element to utilizing Algebraic Number Theory: Algebraic Numbers. According to [2], we can say that a complex number α is **algebraic** if it is the root of a polynomial with specifically integer coefficients, and **transcendental** if it is not. Furthermore, in the following proof, we can conclusively prove that there are only a countably large amount of Algebraic Numbers.
Given any polynomial with integer coefficients:

$$p(X)=C_0X^d+C_1X^{d-1}+\cdots+Cd=0.$$

with $C_i \in \mathbb{Z}$ and $C_0 \neq 0$, we can define the "height" $H(p)$ as:

$$H(p)=d+|C_0|+\cdots+|C_d| \in \mathbb{Z}$$

Such that given any $n \in \mathbb{Z}$, there are only finitely many such polynomials whose heights are $\leq n$. So, every polynomial with integer coefficients (which corresponds to an algebraic number) can thus be controlled by an integer, but $\mathbb{Z}$ is countably infinite—proving that Transcendental Numbers not only exist, but are also more prevalent than their Algebraic counterparts as $\mathbb{C}$ is uncountable.
Note: From the aforementioned properties, we can conclude that every rational $\frac{m}{n}$, where $m,n \in \mathbb{Z}$, is algebraic, since it is always a root of $nX-m=0$

With our definition of Algebraic Numbers established, we are able to quickly perceive the definition of the **Minimal Polynomial** of an algebraic number α. The minimal polynomial of α is a (unique) polynomial that consist of the following attributes: (1) coefficients are in $\mathbb{Q}$, (2) leading coefficient is 1 (monic), (3) smallest possible degree, and (4) α is a root.

Example of Minimal Polynomials: If $\alpha=\sqrt{2}$, then $f(x)=x^2-2$ is the minimal polynomial of $\sqrt{2}$, because all the coefficients in $f(x)$ are $\in \mathbb{Q}$, it is monic as the leading coefficient is 1, of the smallest degree (2), and α is a root. Similarly, the minimal polynomial of i is x^2+1.

Field of Algebraic Numbers

Utilizing our newfound knowledge of Algebraic Numbers and Minimal Polynomials, we can finally discuss the **Field of Algebraic Numbers**.

Let us define the set A of algebraic numbers. We actually know that set A is a field, but this will be proven later using field extension. Because it is a field, we can infer that it has the same properties as the ones we have mentioned in the "Rings and Fields" section. As such, if α and β are algebraic numbers, then so are the following:
$\alpha+\beta$, $\alpha-\beta$, $\alpha\beta$, $\frac{\alpha}{\beta}$ where $\beta \neq 0$.

This is important, because for example, assume we want to find the minimal polynomial of $\sqrt{2}+\sqrt{3}$ to check using definition whether it's algebraic. This may be difficult to compute at first glance, but using our knowledge of the field of algebraic numbers, we already know $\sqrt{2}+\sqrt{3}$ is an algebraic number. Even though we did not find the minimal polynomial, we know this is algebraic, as both $\sqrt{2}$ and $\sqrt{3}$ are algebraic.

Field Extension

In our case, a field extension of $\mathbb{Q}$ can be defined as $\mathbb{Q}(\alpha)$, denoted by $\mathbb{Q}(\alpha)/\mathbb{Q}$, where $\mathbb{Q}(\alpha)$ is the smallest field containing $\mathbb{Q}$ and α (an algebraic number); there are a few examples in the Rings and Fields section. An element of $\mathbb{Q}(\alpha)$ is a polynomial with "variable" α (though α is fixed), with coefficients in $\mathbb{Q}$.

Note that we are able to combine two elements in $\mathbb{Q}(\alpha)$ as they are both polynomials and follow the usual rules for scalar multiplication and addition for polynomials. As such, $\mathbb{Q}(\alpha)$ is a vector space over $\mathbb{Q}$. Furthermore, the **degree of the field extension** is defined to be the dimension of the $\mathbb{Q}$ vector space $\mathbb{Q}(\alpha)$. Referring to $\mathbb{Q}(\sqrt{2})$, the dimension is 2, because we have a basis $\{1,\sqrt{2}\}$ that consists of 2 elements.

There exists a lemma that states α is algebraic if and only if the field extension $\mathbb{Q}(\alpha)/\mathbb{Q}$ has a finite degree.

Using the aforementioned lemma, because α and β are algebraic, we know that $[\mathbb{Q}(\alpha):\mathbb{Q}]$ and $[\mathbb{Q}(\beta):\mathbb{Q}]$ are both finite. Thus, $[\mathbb{Q}(\alpha,\beta):\mathbb{Q}]$ must also be finite. As we can infer that $\alpha+\beta \in \mathbb{Q}(\alpha,\beta)$, this implies $\mathbb{Q}(\alpha+\beta) \subseteq \mathbb{Q}(\alpha,\beta)$ and $[\mathbb{Q}(\alpha+\beta):\mathbb{Q}]$ is finite. We know from the aforementioned lemma that α is algebraic if and only if the field extension $\mathbb{Q}(\alpha)/\mathbb{Q}$ has a finite degree. Therefore, $\alpha+\beta$ must be algebraic. Note that in order to show algebraic numbers make a field, we just need to show that they are closed under the operations, since those axioms(say, associativity) are all inherited from C.
We can utilize this proof with $\alpha-\beta$, $\alpha\beta$, $\frac{\alpha}{\beta}$(where $\beta \neq 0$), because they are all $\in \mathbb{Q}(\alpha,\beta)$. Hence, algebraic numbers form a field.

Example of Field Extension: The field extension $\mathbb{Q}(\sqrt{2})=\{a+b\sqrt{2}|a,b \in \mathbb{Q}\}$ and the degree is 2, and $\sqrt{2}$ is algebraic. We can test that $u+v$ and uv are still in the form $a+b\sqrt{2}$ where $a,b \in \mathbb{Q}$ when u,v are.

- $(a_1+_1 b\sqrt{2})+(a_2+b_2\sqrt{2})=(a_1+a_2)+(b_1+b_2)\sqrt{2}$
- $(a_1+b_1\sqrt{2})(a_2+b_2\sqrt{2})=(a_1a_2+2b_1b_2)+(a_1b_2+a_2b_1)\sqrt{2}$

We define a **number field** K as an extension of $\mathbb{Q}$ of finite degree.

Integrality and the Ring of All Algebraic Integers

The **ring of all algebraic integers** I can be defined as an algebraic number α where the minimal polynomial of α over $\mathbb{Q}$ has coefficients in $\mathbb{Z}$. Thus, it is a subset of algebraic numbers, and in the following sections, we will prove that it forms a ring.

First and foremost, suppose α is a root of $x^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0=0$ where $a_i \in \mathbb{Q}$. Then, we have d = common multiple of denominators of a_i, then $d^n(\alpha^n+a_{n-1}\alpha^{n-1}+\cdots+a_1\alpha+a_0)=0$. Thus:

$d^n(\alpha^n+a_{n-1}\alpha^{n-1}+\cdots+a_1\alpha+a_0)=0$
$=>(d\alpha)^n+d\alpha_{n-1}(d\alpha)^{n-1}+\cdots+d^{n-1}a_1(d\alpha)+d^na_0=0$

Because $d\alpha$ is a root of the new equation, $x^n+da_{n-1}x^{n-1}+\cdots+d^{n-1}a_1x+d^na_0=0$. This means that $da_i \in \mathbb{Z}$, because we multiply a_i by its common denominator multiple and $a_i \in Q$. Thus, all the coefficients are integers and $d\alpha$ is an algebraic integer.

Therefore, $\forall \alpha \in A, \exists d \in \mathbb{Z}$ st $d\alpha \in I$,i.e., $d\alpha$ is an algebraic integer.

This means that every algebraic number α is an algebraic integer divided by an integer, which is analogous to a rational. Namely,

$$algebraic\ numbers=\frac{algebraic\ integers}{integers}$$

We define $\mathbb{Z}[\alpha]$ as the smallest ring containing $\mathbb{Z}$ and α, which is analogous to $\mathbb{Q}(\alpha)$. Similar to the lemma in the field extension, α is an algebraic integer if and only if $\mathbb{Z}[\alpha]$ is a finitely generated $\mathbb{Z}$ module. Now we are trying to prove that this a ring using this lemma.

Here R is a finitely generated $\mathbb{Z}$ module means every element in R can be written uniquely as a linear combination of fixed n elements. For example, every element in $\mathbb{Z}[i]$ is in the form $a+bi$ where $a,b \in \mathbb{Z}$. We will be able to show every element in $\mathbb{Z}[\alpha]$ is the root of a monic polynomial with coefficients $\in \mathbb{Z}$, i.e., they are all algebraic integers.

Analogous to the lemma in Field Extension, because α and β are algebraic integers, we know that $\mathbb{Z}[\alpha]$ is finitely generated and $\mathbb{Z}[\beta]$ is finitely generated. Thus, $\mathbb{Z}[\alpha,\beta]$ is finitely generated. Our previous proofs suggest that $\alpha+\beta \in \mathbb{Z}[\alpha,\beta]$, which means $\mathbb{Z}[\alpha+\beta] \subseteq \mathbb{Z}[\alpha,\beta]$ and $\mathbb{Z}[\alpha+\beta]$ is finitely generated. Therefore, $\alpha+\beta$ must be algebraic.
We can utilize this proof with $\alpha-\beta$, $\alpha\beta$, because they are all in $\mathbb{Z}[\alpha,\beta]$

Therefore, the set of all algebraic integers forms a ring.

Again, is $\sqrt{2}+\sqrt{3}$ an algebraic integer? We know the answer is yes, despite the fact that we didn't even compute it's minimal polynomial!

Integers in Number Fields

As a consequence, let us look at the **integers in a number field** K, which is by definition, $K \cap I$(namely, algebraic integers that are in K):

If we take $\alpha,\beta \in K \cap I$ (integers in K, as we just defined), then we can prove that their sum $\alpha+\beta \in K \cap I$.

Because they are in the intersections of K and I, $\alpha,\beta \in K$ and $\alpha,\beta \in I$. Furthermore, since I is a ring as proven above, $\alpha+\beta \in I$. Similarly, K is a field (closed under addition), so $\alpha+\beta \in K$.

Therefore, $\alpha+\beta \in K \cap I$. This is similar to $\alpha-\beta$ and $\alpha*\beta$, as both are in $K \cap I$.
In conclusion, $K \cap I$ forms a ring.

Example:
The integers in $\mathbb{Q}(\sqrt{2})$ is $\mathbb{Z}[\sqrt{2}]$
The integers in $\mathbb{Q}(i)$ is $\mathbb{Z}[i]$

Further Applications

With all these definitions, we could study number theory, say the theory of prime numbers, in a much broader context. Some familiar results about $\mathbb{Z}$ are still true in this new setting, but some are not (as an example, unique factorization of an integer into primes fail in general). These discoveries lead us to the modern algebraic number theory...
And yes, we are also able to show that the integers in a number field are always finitely generated—just as all the existing examples suggest.

References

[1] David R. Finston and Patrick J. Morandi. "Abstract Algebra: Structure and Application". In: (2010).

[2] Frazer Jarvis. "Algebraic Number Theory". In: (2010).

Elliptic Curve Cryptography

Rocky Beaty
Mentor: **Mychelle Parker**
University of California, Santa Barbara | 2022 Directed Reading Program
UC SANTA BARBARA
Department of Mathematics

What is an Elliptic Curve?

Definition 1 An *elliptic curve over a field K* is defined by an equation

$$E : y^2 = x^3 + ax + b \tag{1}$$

where $a, b \in K$ and $\Delta \neq 0$ where Δ is the *discriminant* of E and is defined as

$$\Delta = -16(4a^3 + 27b^2).$$

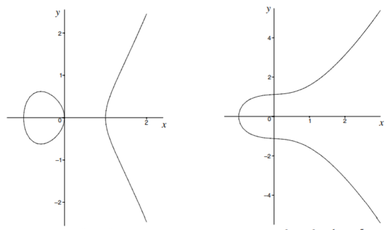
Note: There is a more general form of the equation:

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6.$$

However, if the characteristic of $K \neq 2$ or 3, then the equation can be expressed as in (1). This assumption applies to all elliptic curves used in cryptography, and thus equation (1) is sufficient for us.

Definition 2 Let K be a field over which an elliptic curve is defined. Then the *K-rational* points, denoted $E(K)$, are all points on E with coordinates in K, along with the point at infinity denoted ∞. The *order* of the curve, $\#E(K)$, is the total number of points on the curve.

Elliptic curves can be defined over infinite fields such as $\mathbb{R}$ or $\mathbb{Q}$, or they can be defined over finite fields such as $\mathbb{Z}/p\mathbb{Z}$ or $\mathbb{F}_q$. Consider the following graphs of various elliptic curves:



(a) $E_1 : y^2 = x^3 - x$

(b) $E_2 : y^2 = x^3 + \frac{1}{4}x + \frac{5}{4}$

Fig. 1: Elliptic Curves over $\mathbb{R}$. [1]

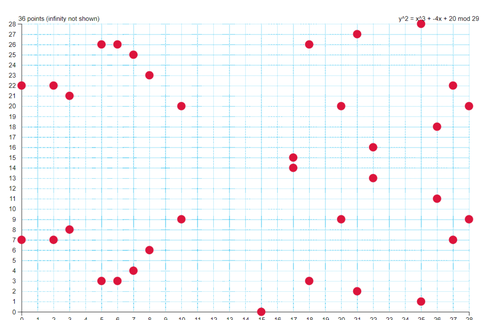
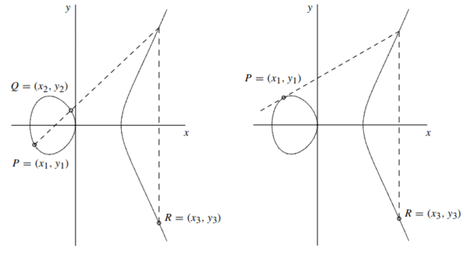


Fig. 2: Elliptic Curve over finite field $\mathbb{F}_{29}$. [2]

Group Law

There is a convenient way of defining an addition operation for two points in $E(K)$ to give a third point in $E(K)$. With this operation, the set of points in $E(K)$ forms an abelian group, where ∞ serves as the identity. The addition operation has a clear geometric interpretation. First, notice that any line will intersect an elliptic curve E at most 3 times. Given any two distinct points, $P = (x_1, y_1)$ and $Q = (x_2, y_2)$, on E, then $P + Q = R = (x_3, y_3)$ is found by drawing a line through P and Q, find the third point this line intersects E. Then to obtain R reflect this point about the x-axis. Doubling a point P is the same, though the tangent line at the point P is used. Note: $P - Q$ is performed by taking $-Q = (x_2, -y_2) \in E(K)$.



(a) Addition: $P + Q = R$. (b) Doubling: $P + P = R$.

Fig. 3: Point Addition and Point Doubling. [1]

From this abelian group comes the basis for the scheme of elliptic curve cryptography.

What is Elliptic Curve Cryptography?

Elliptic Curve Cryptography (ECC) is a modern public-key cryptography technique based on the mathematics of elliptic curves over finite fields. ECC relies on the difficulty of solving the Elliptic Curve Discrete Logarithm Problem.

Definition 3 The *elliptic curve discrete logarithm problem* (ECDLP) is: given an elliptic curve E defined over a finite field $\mathbb{F}_q$, a point $P \in E(\mathbb{F}_q)$ of order n, and a point $Q \in \langle P \rangle$, find the integer $l \in [0, n-1]$ such that $Q = lP$. The integer l is called the discrete logarithm of Q to the base P, denoted $l = \log_P Q$.

Simply put, the ECDLP is the problem of finding an integer n such that $Q = nP$. It exploits the fact that, as shown above, it is rather easy to double a point $P \in E(K)$ together, but it is thought to be very difficult to figure out how many times the point was doubled. The essential pieces of a secure ECC scheme are:

1. Elliptic Curve $E(\mathbb{F}_p)$ over finite field $\mathbb{F}_p$, p prime
2. P : generator - $P \in E(\mathbb{F}_p)$ *is a generator*
3. d : private key - $d \in \mathbb{Z}$ *is selected uniformly at random from the interval [1,n-1]*
4. Q : public key - *a point* $Q = dP \in E(\mathbb{F}_p)$
5. k : random integer - *used to increase security of encryption scheme*

In the ECC scheme, a sender's message is represented as a point M, and encrypted by adding it to kQ, where $Q = dP$ is the intended recipient's public key. The sender transmits the points $C_1 = kP, C_2 = M + kQ$ to the recipient who uses their private key d to compute

$$dC_1 = d(kP) = k(dP) = kQ$$

and can then easily recover $M = C_2 - kQ$. An attacker would have to find kQ, which is computationally infeasible using the public information.

Example

Note: It is possible to turn the geometric interpretation of point addition and point doubling into algebraic formulas by solving the cubic equations.
Let $K = \mathbb{F}_{97}$ and take

$$E : y^2 = x^3 + 2x + 3.$$

Consider $P = (3, 6)$, one can calculate the multiples of P using the mentioned algebraic formulas to obtain:

$$0P = \infty \;\; 1P = (3,6) \;\; 2P = (80,10) \;\; 3P = (80,87) \;\; 4P = (3,91)$$
$$5P = \infty \;\; 6P = (3,6) \;\; 7P = (80,10) \;\; 8P = (80,87) \;\; 9P = (3,91)$$

This pattern continues, so we see that $5P = \infty \implies P$ is a generator of order $n = 5$, and forms the *cyclic subgroup*

$$\langle P \rangle = \{\infty, P, 2P, 3P, 4P\}.$$

Now, consider the following problem. Let

$$P = (3,6),\ d = 3,\ Q = dP = 3P = (80,87),\ k = 9$$

and suppose the encoded message is $M = (24, 2)$. Using the algebraic formulas, one can calculate C_1 and C_2,

$$C_1 = kP = 9(3,6) = (3,91)$$

$$C_2 = M + kQ = (24,2) + 9(80,87) = (24,2) + (80,10) = (92,16).$$

The recipient receives C_1 and C_2, and then computes

$$dC_1 = d(kP) = k(dP) = kQ = (80,10)$$

so

$$M = C_2 - kQ = (92,16) - (80,10).$$

Notice that $-kQ = -(80,10) = (80,-10)$ where $-10 \equiv 87 \pmod{97}$ hence $-kQ = (80,87)$. So we get

$$M = (92,16) + (80,87) = (24,2)$$

as desired. An attacker wishing to recover M would likely know $E(\mathbb{F}_{97}), P, n, Q, C_1$ and C_2. However, even with this information it is computationally infeasible to compute kQ due to the cyclic nature of $\langle P \rangle$ and assuming k is sufficiently random.

Why ECC?

ECC is often preferred over RSA schemes because of the security and performance it offers using smaller key sizes. A common ECC key size of 256-bits is equivalent to a 3072-bit RSA key.

References

[1] *Elliptic Curves over Finite Fields*. `https://graui.de/code/elliptic2/`. Accessed: 2022-05-02.

[2] Darrel Hankerson, Alfred Menezes, and Scott Vanstone. *Guide to Elliptic Curve Cryptography*. Spring, 2004.

Algebraic Number Theory and Applications

Yanbo Cheng, Mychelle Parker
UC Santa Barbara

Motivations

It is known that a prime p can be written in the form $p = x^2+y^2$ with $x, y \in \mathbb{Z}$ if and only if $p \equiv 1 \mod 4$. Since we can factorize such p in $\mathbb{Z}[i]$ as $p = (x+iy)(x-iy)$, it is natural to think of the prime elements in $\mathbb{Z}[i]$. We then want to relate the field $\mathbb{Q}(i)$ to $\mathbb{Z}[i]$, and a proposition was found that illustrates such relationship.
Proposition 1.

$$\mathbb{Z}[i] = \{x \in \mathbb{Q}(i) : x^2 + ax + b = 0 \text{ for some } a, b \in \mathbb{Z}\}$$

This proposition can be seen as a motivation to study the properties of algebraic integers of an algebraic number field.

Introduction

We first establish some basic principles of algebraic number theory.

Definition 2. An **algebraic number field** K is a finite extension of $\mathbb{Q}$. A element $\alpha \in K$ is called an **algebraic integer** if $f(\alpha) = 0$ for some monic polynomial $f(x) \in \mathbb{Z}[x]$.

Definition 3. Let $A \subseteq B$ be a ring extension. Then, $b \in B$ is **integral** over A if $f(b) = 0$ for some monic polynomial $f(x) \in A[x]$. We then define the **integral closure** to be the set $\overline{A} = \{b \in B : b \text{ integral over } A\}$. A is then called **integrally closed** if $A = \overline{A}$.

As in linear algebra, traces and norms play an important role in algebraic number theory. We thus give their definition.

Definition 4. For a finite field extension $L|K$. The **trace** of an element $\alpha \in L$ is the trace of the endomorphism $\psi : L \to L$, $\psi(x) = \alpha x$ where L is seen as a K-vector space. The **norm** pf α is then the determinant of ψ, that is:

$$Tr_{L|K}(\alpha) = Tr(\psi), \qquad N_{L|K}(\alpha) = \det(\psi)$$

There is an extra property in of traces and norms in a separable extension $L|K$ that uses field embeddings from L into an algebraic closure $\overline{K}$ of K.

Proposition 5. *Let $L|K$ be a separable extension, and define the set $\Sigma = \{\sigma : L \to \overline{K}$ a field embedding $\}$. Then we have:*

$$Tr_{L|K}(\alpha) = \sum_{\sigma \in \Sigma} \sigma(\alpha)$$

$$N_{L|K}(\alpha) = \prod_{\sigma \in \Sigma} \sigma(\alpha)$$

We then give the definition of a Dedekind domain, which is the main object that algebraic number theory studies.

Definition 6. A **Dedekind domain** is a neotherian, integrally closed integral domain in which every nonzero prime ideal is maximal.

The product and sum of ideals defined such that

$$\mathfrak{a} + \mathfrak{b} = \{a + b : a \in \mathfrak{a}, b \in \mathfrak{b}\}$$

$$\mathfrak{ab} = \{\sum_{i \in I} a_i b_i : a_i \in \mathfrak{a}, b_i \in \mathfrak{b}, \forall i \in I\}$$

The importance of Dedekind domain is due to the fact that it gives unique prime factorization of prime ideals.

Dedekind domain

In this section, we denote $\mathcal{O}_K$ to be the ring of integers of an algebraic number field K. Such a ring has the following main properties:

Theorem 7. *$\mathcal{O}_K$ is a neotherian ring. It is integrally closed and every nontrivial prime ideal of $\mathcal{O}_K$ is a maximal ideal.*

Theorem 8. *Every ideal of $\mathfrak{a}$ of a Dedekind domain $\mathcal{O}$ that is nonzero and not the $\mathfrak{a} \neq \mathcal{O}$ admits a factorization in to nonzero prime ideal of $\mathcal{O}$:*

$$\mathfrak{a} = \mathfrak{p}_1 \cdots \mathfrak{p}_n$$

This factorization is unique up to reordering.

We see that this is similar to a unique factorization domain in which every element admits a factorization into a product of a unit and irreducible elements which is unique up to association and reordering.
Then we can thus look at the properties of the extensions of Dedekind domains.
Let o be a Dedekind domain with field of fraction K, let $L|K$ be a field extension with integral closure $\mathcal{O}$. Then we can decompose prime ideals of o in $\mathcal{O}$

Theorem 9. *Let $\mathfrak{p}$ be a prime ideal of o, then*

$$\mathfrak{p}\mathcal{O} = \mathfrak{P}_1^{e_1} \cdots \mathfrak{P}_n^{e_n}$$

with $f_i = [\mathcal{O}/\mathfrak{P}_i : o/p]$, we have the fundamental identity:

$$\sum_{i=1}^{n} f_i e_i = [L : K]$$

P-adic numbers

Now we introduce another topic, which are the p-adic numbers. We give two definitions of the p-adic integers $\mathbb{Z}_p$.

Definition 10. $\mathbb{Z}_p$ can be defined as the projective limit of if the rings $\mathbb{Z}/p^n\mathbb{Z}$, and thus

$$\mathbb{Z}_p = \lim_{n\leftarrow} \mathbb{Z}/p^n\mathbb{Z} = \{(x_n)_n \in \prod_{n=1}^{\infty} \mathbb{Z}/p^n\mathbb{Z} : x_{n+1} \equiv x_n \mod p^n\}$$

We could also define $\mathbb{Z}_p$ through Cauchy sequences.
Define the p-adic absolute value $||_p$ as follows:
Let $a = \frac{b}{c}$, $b, c \in \mathbb{Z}$, we can find some integer n such that $a = p^n\frac{b'}{c'}$ where $(b'c', p) = 1$. Then we have $|a|_p = \frac{1}{p^n}$. We can thus define a metric using $||_p$ just like what we did using the normal absolute value $||$. Thus, we can define the p-adic numbers using Cauchy sequence with respect to the metric $||_p$. The induced metric on $\mathbb{Z}_p$ is $d(x, y) = |x - y|_p$ for $x, y \in \mathbb{Z}_p$.

Definition 11. Let R be the ring of Cauchy Sequence with respect to $||_p$, and m be the ideal of nullsequence, that is, the Cauchy sequences that converges to zero. Then we define the p-adic numbers $\mathbb{Q}_p$ as

$$\mathbb{Q}_p = R/m$$

Then, define the p-adic integers as

$$\mathbb{Z}_p = \{x \in \mathbb{Q}_p : |x|_p \leq 1\}$$

The Unit theorem

From Minkowski Theory, we derive the Dirichlet's Unit Theorem by studying the exact sequence:

$$1 \to \mu(K) \to \mathcal{O}_K^* \to \Gamma \to 0$$

Where $\mathcal{O}_K^*$ is the group of units and $\mu(K)$ is the roots of unity that lie in K, and Γ is the image $\lambda(\mathcal{O}_K^*)$ defined by

$$\lambda(a) = (log|\tau(a)|)_\tau \in [\prod_\tau \mathbb{R}]^+$$

Where τ run over the complex embeddings $\tau : K \to \mathbb{C}$.

Theorem 12. *The group of units $\mathcal{O}_K^*$ of $\mathcal{O}_K$ is the direct product of the finite group $\mu(K)$, which is the group of roots of unity are in K, and a free abelian group of rank $r + s - 1$. Where r is the number of real embeddings $\sigma : K \to \mathbb{R}$ and s is the number of pairs of complex conjugate embeddings $\sigma, \overline{\sigma} : K \to \mathbb{C}$.*

This theorem give us a way to express any units u in $\mathcal{O}_K$ uniquely in the form

$$u = \xi u_1^{i_1} u_2^{i_2} \cdots u_{r+s-1}^{i_{r+s-1}}$$

where ξ is a root of unity and $u_1, u_2 \cdots$ are units of $\mathcal{O}_K$ that can be seen as a basis of the free abelian group mentioned above.

Applications

One of the applications of the Dirichlet's Unit Theorem is the solution of Pell's equations.

Corollary 13. *There exists infinily many pairs of solutions $x, y \in \mathbb{Z}$ to the equation*

$$x^2 + ny^2 = 1$$

with $n < 0$ not a perfect square and $n \in \mathbb{Z}$.

This is a direct application of the Dirichlet's Unit Theorem on the quadratic extension $K|\mathbb{Q}$, where $K = \mathbb{Q}(\sqrt{-n})$, and we use the fact that $r = 2$, $s = 0$, thus $r + s - 1 = 1$.
An application of the p-adic numbers is the following proposition:

Proposition 14. *Let $f(x_1, \cdots, x_n)$ be a polynomial with coefficients in integer. Then we have the equivalence:*

$$f(x_1, \cdots, x_n) \equiv 0 \mod p^n \text{ is solvable for all } n \geq 1$$
$$\iff f(x_1, \cdots, x_n) = 0 \text{ is solvable in p-adic integers}$$

Thus, the application of p-adic number also gives a way to solve problems in elementary number theory.

Acknowledgements

This is a poster of the Directed Reading Program in 2022. I would like to thank Mychelle Parker for being my mentor in this program.

References

Jürgen Neukirch, Algebraic Number Theory

Brouwer's Fixed Point Theorem with Application to Game Theory

Ruizhe Qian

Mentor: Pranav Arrepu

Introduction

We will prove Brouwer's Fixed Point Theorem by using fundamental groups. Then, we will show the application of Brouwer's Fixed Point Theorem to the game theory, namely the Nash Equilibrium.

Brouwer's Fixed Point Theorem in $\mathbb{R}$

Theorem (Brouwer's Fixed Point Theorem)**.** *Given that set $K \subset \mathbb{R}^n$ is compact and convex, and that function $f : K \to K$ is continuous, then there exists $c \in K$ such that $f(c) = c$.*

This is generalized statement of Brouwer's Fixed Point Theorem in $\mathbb{R}$. In this poster, we will explore the proof of a simple case, $D^2 \subset \mathbb{R}^2$. D^2 is homeomorphic to any closed and bounded compact subset of $\mathbb{R}^2$. But we will use the generalized version of this theorem to prove the existence of Nash equilibrium.

Algebraic Topology Preliminaries

We establish our theory from homotopy, an important equivalence relationship in topology,

Definition (Homotopy)**.** *[3] Two continuous maps $f_0, f_1 : X \to Y$ are said to be homotopic if there is a continuous map $F : X \times I \to Y$ such that $F(x, 0) = f_0(x)$ and $F(x, 1) = f_1(x)$. Then, we say $f_1(x) \simeq f_0(x)$*

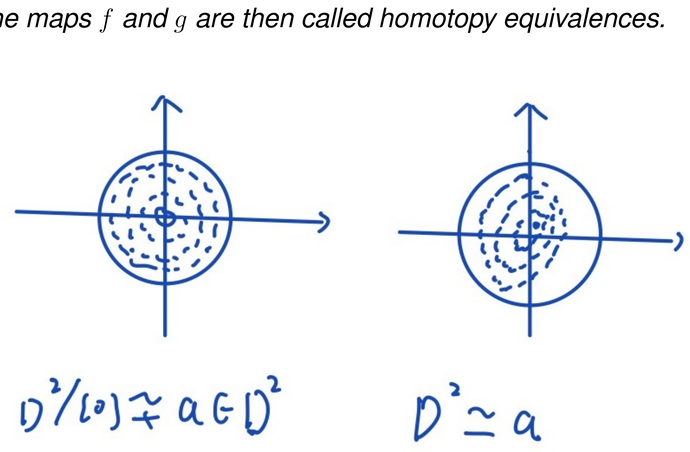
Definition (Homotopic Relative)**.** *Suppose that A is a subset of X and that f_0 and f_1 are two continuous functions from X to Y. We say f_0 and f_1 are homotopic relative to A if there is a homotopy $F : X \times I \to Y$ between f_0 and f_1 such that $F(a, t)$ does not depend on t for $a \in A$.*

Homotopy type, also known as homotopy equivalence, following from homotopy, is an important tool to classify topological space.

Definition (Homotopy Equivalence)**.** *Two spaces X and Y are homotopic equivalent if there exists continuous maps $f : X \to Y$ and $g : Y \to X$ such that*

$$g \circ f \simeq Id : X \to X$$
$$f \circ g \simeq Id : Y \to Y$$

The maps f and g are then called homotopy equivalences.



Spaces that are homotopy equivalent to a point are given a special name. The identity function of this space is homotopic to the constant map.

Definition (Contractible)**.** *A space X is said to be contractible if it is homotopy equivalent to a point.*

By using constant map and inclusion map, the following result can be easily derived.

Remark. *D^n is contractible and any convex subset of $\mathbb{R}^n$ is contractible.*

Consider that the cylinder, C and the circle S are a pair of homotopy equivalent spaces. Define $i : S \to C$ as the inclusion. This motivates the following definition.

Definition (Retraction)**.** *A subset A of a topological space X is called a retract of X if there is a continuous map $r : X \to A$ such that $r \circ i = Id : \quad A \to A$, where $i : A \to X$ is the inclusion map. The map r is called a retraction.*

Before we step into the definition of the fundamental group, we want to give the definition of some related concepts.

Definition (Path)**.** *A continuous mapping $f : [0, 1] \to X$ is called a path in X.*

Definition (Path Equivalent)**.** *Two path f, g in X are said to be equivalent if f and g are homotopic relative to $\{0, 1\}$. We write $f \sim g$*

Definition (Loop)**.** *A path is said to be closed if $f(0) = f(1)$. If $f(0) = f(1) = x$ then we say that f is based at x.*

Now, we have the definition of the fundamental group.

Definition (Fundamental Group)**.** *[1] The fundamental group of a space X will be defined so that its elements are loops in X starting and ending at a fixed basepoint $x \in X$ but two such loops are regarded as determining the same element of the fundamental group if one loop is homotopy equivalent to the other in space X. We denote this group as $\pi(X, x)$.*

We will explore the effect of continuous map between topological spaces $\psi : X \to Y$ has upon fundamental groups. Consider $\psi_* : \pi(X, x) \to \pi(Y, \psi(x))$ where $\psi_*[f] = [\psi f]$, f is a path in X.

Lemma. *ψ_* is a homomorphism of groups.*

Proof. $\psi_*([f][g]) = \psi_*([f * g] = [\psi(f * g)] = [\psi f * \psi g] = [\psi f][\psi g] = \psi_*[f]\psi_*[g]$. □

By proving this lemma, we can give ψ_* a name.

Definition (Induced Homomorphism)**.** *The homomorphism $\psi_* : \pi(X, x) \to \pi(Y, \psi(x))$ defined by $\psi_*[f] = [\psi f]$, where $\psi : X \to Y$ is a continuous map, is called the induced homomorphism.*

What if we have ψ as a homeomorphism?

Corollary. *If $\psi : X \to Y$ is a homeomorphism then $\psi_* : \pi(X, x) \to \pi(Y, \psi(x))$ is an isomorphism.*

The last piece of the puzzle is the fundamental group of the circle S^1, which turns out to be $\mathbb{Z}$. Let's consider a map e

$$\mathbb{R} \to S^1$$
$$t \to e^{2\pi i t}$$

Note that $e^{-1}(1) = \mathbb{Z} \subset \mathbb{R}$. If we are given $f : I \to S$ with $f(0) = f(1) = 1$, there is a unique map $\tilde{f} : I \to \mathbb{R}$ with $\tilde{f}(0) = 0$ and $e\tilde{f} = f$. $\tilde{f}$ is the lifting map of f. The integer $\tilde{f}(1) \in e^{-1}(1) = \mathbb{Z}$ is defined to be degree of f. If f_0 and f_1 are equivalent paths in S^1, then $\tilde{f}_0(1) = \tilde{f}_1(1)$. As a result, the function $\pi(S^1, 1) \to \mathbb{Z}$, where $[f] \mapsto degree(f)$, is isomorphism, which means the fundamental group of the circle is the set of integers.

Algebraic Proof for the Main Theorem

Proof. Suppose to the contrary that $x \neq f(x)$ for all $x \in D^2$. Then, we may define a function $\psi : D^2 \to S^1$ by setting $\psi(x)$ to be the point on S^1 obtained from the intersection of the line segment from $f(x)$ to x extended to meet S^1. We want to show ψ is continuous. Let's write ψ explicitely in coordinates, $y = \psi(x)$. The condition the ray meets the boundary is

$$|y + t(x - y)|^2 = 1.$$

It is a quadratic equation with the solution in

$$t_\pm = \frac{-2(x-y)y \pm \sqrt{4((x-y)y)^2 - 4|x-y|^2(|y|^2 - 1)}}{2|x-y|^2}$$

We only interested in the solution $y + t_+(x - y)$. Therefore, ϕ is continuous. Define $i : S^1 \to D^2$, denote the inclusion, then $\psi i = Id$ and we have the commutative diagram.

$$\begin{array}{ccc} S^1 & \xrightarrow{Id} & S^1 \\ & \searrow^{i} & \uparrow{\scriptstyle \psi} \\ & & D^2 \end{array}$$

This leads to another commutative diagram,

$$\begin{array}{ccc} \pi(S^1, 1) & \xrightarrow{Id} & \pi(S^1, 1) \\ & \searrow^{i_*} & \uparrow{\scriptstyle \psi_*} \\ & & \pi(D^2, 1) \end{array}$$

where ψ_* and i_* denote induced homomorphism. But $\pi(D^2, 1) = 0$ since D^2 is contractible, and so we get another commutative diagram due to isomorphism.

$$\begin{array}{ccc} \mathbb{Z} & \xrightarrow{Id} & \mathbb{Z} \\ & \searrow^{i_*\psi_*} & \uparrow \\ & & 0 \end{array}$$

This is impossible. Therefore, we prove the Brouwer's Fixed Point Theorem in two dimension. □

Game Theory Preliminaries

Now we move on to application to the game theory. We want to introduce the abstract notion of a normal form game, the following definitions are from [2]. We will use prisoner's dilemma to illustrate those definitions. In prisoner's dilemma, two prisoners are interrogated separately. If both of them confess, they get sentence for 3 years. If one confess, the other does not, the people who confess gets 1 years of sentence, the other gets 10 year. If both of them do not confess, they are innocent.

Definition (Normal-form game)**.** *A (finite, n-person) normal-form game is a tuple (N, A, O, μ, u), where*

- *N is a finite set of n players, indexed by i.*
- *$A = (A_1, \ldots, A_n)$, where A_i is a finite set of actions (or pure strategies; we will use the terms interchangeably) available to player i. Each vector $a = (a_1, \ldots, a_n) \in A$ is called an action profile (or pure strategy profile);*
- *O is a set of outcomes;*
- *$\mu : A \to O$ determines the outcomes as a function of the action profile; and*
- *$u = (u_1, \ldots, u_n)$ where $u_i : O \to \mathbb{R}$ is a real valued utility function for player i*

In prisoner's dilemma, $N = 2$. A is confess or not confess. O is what happens if both of prisoners confess, not confess and etc. In this way, μ and u is easy to understand. While players can select a single action to play, which is the pure strategy, they can also follow another type of strategy: randomizing over the set of available actions according to some probability distribution. Such strategy is called mixed strategy. We can define mixed strategy as follows. In prisoner's dilemma, mixed strategy can be like one person has 50% chance confessing 50% chance not confessing.

Definition (Mixed Strategy)**.** *Let $(N, (A_1, \ldots, A_n), O, \mu, u)$ be a normal form game, and for any set X let $\prod(X)$ be the set of all probability distributions over X. Then, the set of mixed strategies for player i is $S_i = \prod(A_i)$.*

Definition (Mixed Strategy Profile)**.** *The set of mixed strategy profiles is simply the Cartesian product of the individual mixed strategy sets, $S_1 \times \cdots \times S_n$.*

By $s_i(a_i)$ we denote the probability that an action a_i will be played under mixed strategy s_i. The subset of actions that are assigned positive probability by the mixed strategy s_i is called the support of s_i.

Definition (Support)**.** *The support of a mixed strategy s_i for a player i is the set of pure strategies $\{a_i | s_i(a_i) > 0\}$.*

Now, we want to introduce the concept of expected utility, a basic notion in decision theory,

Definition (Expected Utility of a Mixed Strategy)**.** *Given a normal form game (N, A, u), the expected utility u_i for player i of the mixed strategy profile $s = (s_1, \ldots, s_n)$ is defined as*

$$u_i(s) = \sum_{a \in A} u_i(a) \prod_{j=1}^{n} s_j(a_j)$$

Then, we want to look at games from an individual agent's point of view. This is going to lead us to the most influential concept in game theory, the Nash Equilibrium. Assume an agent knew how others were going to play, his strategy becomes simple. Define $s_{-i} = (s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_n)$, a strategy profile s without i's strategy. We can write $s = (s_i, s_{-i})$. If the agents other than i were commit to play s_{-i}, a utility-maximizing agent I would face the problem of determining his best response.

Definition (Best Response)**.** *Player i's best response to the strategy profile s_{-i} is a mixed strategy $s_i^* \in S_i$ such that $u_i(s_i^*, s_{-1}) \geq u_i(s_i, s_{-i})$ for all strategies $s_i \in S_i$. Obviously, in prison's dilemma, confess is the best response.*

Finally, we will go to the most important definition, the Nah equilibrium. Its existence is one of the most well known application of Brouwer's Fixed Point Theorem.

Definition (Nash Equilibrium)**.** *A strategy profile $s = (s_1, \ldots, s_n)$ is a Nash equilibrium if for all agents i, s_i is a best response to s_{-i}*

When both prisoners confess, the game attain the Nash equilibrium.

Existence of Nash Equilibrium

Theorem (Nash 1951)**.** *Every game with a finite number of players and action profiles has at least one Nash equilibrium*

Proof. Given a strategy profile $s \in S$, for all $i \in N$ and $a_i \in A_i$ we define

$$\psi_{i,a_i}(s) = max\{0, u_i(a_i.s_{-i}) - u_i(s)\}$$

a function denoting the change of utility after each iteration of strategy. We then define the function $f : S \to S$ by $f(s) = s'$, where

$$s_i'(a_i) = \frac{s_i(a_i + \psi_{i,a_i}(s))}{\sum_{b_i \in A} s_i(b_i) + \psi_{i,b_i}(s)} = \frac{s_i(a_i + \psi_{i,a_i}(s))}{\sum_{b_i \in A} \psi_{i,b_i}(s) + 1}$$

Intuitively, this function maps a strategy profile s to a new strategy profile s' in which each agent's actions that are better responses to s receive increased probability mass.
The function f is continuous since each ψ_{i,a_i} is continuous. Since S is convex and compact and $f : S \to S$, by Brouwer's Fixed Point Theorem, f must have at least one fixed point.
First if s is a Nash equilibrium then all ψ's are 0, making s a fixed point of f.
Conversely, consider an arbitrary fixed point of f, s. If s is a fixed point, then $s'(a_i') = s(a_i')$. It follows that $\psi_{i,b_i}(s) = 0$, which only happens when no player can enhance their utility. Therefore, $s' = f(s)$ is the Nash equilibrium □

References

[1] Allen Hatcher. *Algebraic Topology*. 2001.

[2] Alber Xin Jiang and Kevin Leyton-Brown. "A Tutorial on the Proof of the Existence of Nash Equilibrium". In: (2007).

[3] Czes Kosniowski. *A First Course in Algebraic Topology*. Cambridge University Press, 1980.

A Bite-sized Introduction to Fluid Mechanics

Carlos Ortiz, advised by Pranav Arrepu

2022 Math Directed Reading Program, UC Santa Barbara

Introduction

The focus of fluid mechanics is the measurement of **observables** related to a fluid. Liquids and gases are examples of fluids, and their observables include temperature, pressure and density, to name a few. To approach this focus through first principles, the description of fluids is idealised by the notion of a *continuum*, which neglects the microscopic structure of fluids as separate molecules. "Infinitesimal" volume-elements of the fluid (called *fluid parcels*) are then understood to be large enough to contain many molecules, but small relative to the variation in the length-scale of the fluid properties. In this way, observables are understood to be averaged values over a fluid parcel. With this discussion, we are ready to form a mathematical approach to fluid theory.

To "Think" Eulerian or Lagrangian

There are two natural methods for studying fluid properties. In the **Lagrangian** approach, we follow a fluid parcel as it moves and measure the observables along the motion of the parcel. Suppose, at some initial time t_0, a fluid occupies an open set S_0 of $\mathbb{R}^n$. We could then label a *fluid particle* on the fluid (say with some dye) at the position $\mathbf{a} \in S_0$, and follow the particle over time. At some later time t, the fluid occupies the set S_t, and the particle's position is given by $\mathbf{X}(\mathbf{a}, t) \in S_t$. The Lagrangian coordinate $\mathbf{X}(\mathbf{a}, t)$ depends on the time as well as on the initial position to distinguish between fluid particles.

In the Eulerian approach, we instead consider a fixed point $\mathbf{x}$ in space and measure the fluid properties at this point as functions of time (being careful to ensure that $\mathbf{x}$ remains in S_t – otherwise no fluid is at the point!). An observable, q, is then a function of position and time: $q = q(\mathbf{x}, t)$.

Surely, there must be some relation between the two methods! In fact, the most obvious one is the concept of velocity: we have

$$\frac{\partial \mathbf{X}}{\partial t} = \mathbf{u}(\mathbf{X}(\mathbf{a}, t), t), \tag{1}$$

where $\mathbf{u}$ is the **flow velocity** in the Eulerian viewpoint at the Lagrangian coordinate $\mathbf{X}$. This follows merely by construction. But what about other observables? In the Lagrangian perspective, any observable attached to a fluid parcel just depends on the time explicitly. In the Eulerian viewpoint, however, the observables depend on time explicitly and implicitly through the position. By the Chain rule,

$$\frac{\mathrm{d}\, q(\mathbf{X}(\mathbf{a}, t), t)}{\mathrm{d}\, t} = \frac{\partial q}{\partial t} + \frac{\partial \mathbf{X}}{\partial t} \nabla q = \partial_t q + \mathbf{u} \cdot \nabla q. \tag{2}$$

The special operator

$$\frac{\mathrm{D}}{\mathrm{D}t} \equiv \frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla, \tag{3}$$

is called the **material derivative**, and denotes the Lagrangian time derivative in Eulerian variables.

A Useful Identity: The Derivative of the Determinant of the Jacobian

One useful way to describe how fluid parcels transform is by the Jacobian $\mathbf{J}$. (For simplicity, I will work in three-dimensions, and I'll make heavy use of indicial notation and Einstein summation convention.) Recall in the Lagrangian viewpoint that a fluid moves from a set S_0 to S_t, which can be understood to occur via a map $M_t : S_0 \to S_t$. We now introduce the Jacobian of this map, whose elements are given by

$$\mathbf{J}_{ij} = \left.\frac{\partial x_i}{\partial a_j}\right|_t. \tag{4}$$

The determinant of the Jacobian can be written succinctly using the completely antisymmetric tensor, ϵ_{ijk}, as

$$J = \epsilon_{ijk} \mathbf{J}_{1i} \mathbf{J}_{2j} \mathbf{J}_{3k}. \tag{5}$$

With this, we can describe the deformation of a fluid parcel from $S_0 \to S_t$ by

$$\int_{S_t} \mathrm{d}^3\mathbf{x} = \int_{S_0} J \, \mathrm{d}^3\mathbf{a}. \tag{6}$$

We might be interested in observing how volume integrals like (6) change over time. This raises an interim problem: what is the material derivative of J? Well, by the product rule,

$$\frac{\mathrm{D}\, J}{\mathrm{D}\, t} = \epsilon_{ijk} \left(\frac{\mathrm{D}\, \mathbf{J}_{1i}}{\mathrm{D}\, t} \mathbf{J}_{2j} \mathbf{J}_{3k} + \mathbf{J}_{1i} \frac{\mathrm{D}\, \mathbf{J}_{2j}}{\mathrm{D}\, t} \mathbf{J}_{3k} + \mathbf{J}_{1i} \mathbf{J}_{2j} \frac{\mathrm{D}\, \mathbf{J}_{3k}}{\mathrm{D}\, t} \right). \tag{7}$$

Consider the derivative in the first term on (7):

$$\frac{\mathrm{D}\, \mathbf{J}_{1i}}{\mathrm{D}\, t} = \frac{\mathrm{D}}{\mathrm{D}\, t} \left(\frac{\partial x_1}{\partial a_i} \right) = \frac{\partial}{\partial a_i} \frac{\mathrm{D}\, x_1}{\mathrm{D}\, t} = \frac{\partial u_1}{\partial a_i} = \frac{\partial u_1}{\partial x_l} \frac{\partial x_l}{\partial a_i} = \frac{\partial u_1}{\partial x_l} \mathbf{J}_{li}. \tag{8}$$

We interchange the derivatives since the initial position $\mathbf{a}$ is time-independent. The first term of (7) now expands completely as

$$\epsilon_{ijk} \frac{\mathrm{D}\, \mathbf{J}_{1i}}{\mathrm{D}\, t} \mathbf{J}_{2j} \mathbf{J}_{3k} = \epsilon_{ijk} \frac{\partial u_1}{\partial x_l} \mathbf{J}_{li} \mathbf{J}_{2j} \mathbf{J}_{3k}. \tag{9}$$

For $l \neq 1$, $\epsilon_{ijk} = 0$ by definition since then $i = j$ or $i = k$. An analogous approach can be made for the other two terms of (7), admitting one final term by virtue of index repetition. Thus, reducing (7) gives

$$\frac{\mathrm{D}\, J}{\mathrm{D}t} = \epsilon_{ijk} \mathbf{J}_{1i} \mathbf{J}_{2j} \mathbf{J}_{3k} \left(\frac{\partial u_l}{\partial x_l} \right) = J(\nabla \cdot \mathbf{u}). \tag{10}$$

With this result, we can observe how a fluid property $q(\mathbf{X}(\mathbf{a}, t), t)$ changes with time over a fluid parcel:

$$\frac{\mathrm{D}}{\mathrm{D}\, t} \int_{S_t} q(\mathbf{X}(\mathbf{a}, t), t) \, \mathrm{d}^3\mathbf{x} = \frac{\mathrm{D}}{\mathrm{D}\, t} \int_{S_0} q J \, \mathrm{d}^3\mathbf{a} = \int_{S_0} \left(\frac{\mathrm{D}\, q}{\mathrm{D}\, t} J + q \frac{\mathrm{D}\, J}{\mathrm{D}\, t} \right) \mathrm{d}^3\mathbf{a} \tag{11}$$

$$= \int_{S_0} \left(\frac{\mathrm{D}\, q}{\mathrm{D}\, t} + q \nabla \cdot \mathbf{u} \right) J \, \mathrm{d}^3\mathbf{a} = \int_{S_t} \left(\frac{\mathrm{D}\, q}{\mathrm{D}\, t} + q \nabla \cdot \mathbf{u} \right) \mathrm{d}^3\mathbf{x} \tag{12}$$

$$\implies \frac{\mathrm{D}}{\mathrm{D}\, t} \int_{S_t} q(\mathbf{X}(\mathbf{a}, t), t) \, \mathrm{d}^3\mathbf{x} = \int_{S_t} \left(\frac{\mathrm{D}\, q}{\mathrm{D}\, t} + q \nabla \cdot \mathbf{u} \right) \mathrm{d}^3\mathbf{x}. \tag{13}$$

The result of equation (13) is known as the *Reynolds Transport Theorem*. When the observable $q = \rho(\mathbf{X}(\mathbf{a}, t), t)$, we have

$$\frac{\mathrm{D}}{\mathrm{D}\, t} \int_{S_t} \rho \, \mathrm{d}^3\mathbf{x} = \int_{S_t} \left(\frac{\mathrm{D}\, \rho}{\mathrm{D}\, t} + \rho \nabla \cdot \mathbf{u} \right) \mathrm{d}^3\mathbf{x}. \tag{14}$$

If we assume that mass is conserved, then (14) must be zero. The equation above must hold for all fluid parcels, which is only true if the integrand itself is zero:

$$\frac{\mathrm{D}\, \rho}{\mathrm{D}\, t} + \rho \nabla \cdot \mathbf{u} = 0. \tag{15}$$

The Euler Equations of Motion

The result (15) is known as the *continuity equation*, and, together with conservation of momentum, we can arrive at the so-called Euler equations. Newton's second law relates the material derivative of the momentum of a fluid to the net external force on the fluid. The net force per unit volume can be expressed generally as

$$F_i = f_i + \frac{\partial \sigma_{ij}}{\partial x_j}, \tag{16}$$

where σ_{ij} is the stress tensor and f_i is some external body force. Assuming an ideal fluid, the stress tensor is completely diagonal with $\sigma_{ij} = -p\delta_{ij}$, so that $\nabla \cdot \sigma = -\nabla p$, where p is the pressure exerted normal to the surface of the fluid. Hence, by Newton's second law,

$$\int_{S_t} \rho \frac{\mathrm{D}\, \mathbf{u}}{\mathrm{D}\, t} \, \mathrm{d}^3\mathbf{x} = \int_{S_t} \mathbf{f} \, \mathrm{d}^3\mathbf{x} + \int_{\partial S_t} (\boldsymbol{\sigma} \cdot \hat{n}) \, \mathrm{d}^2\mathbf{x}. \tag{17}$$

By the Divergence theorem, the far-right integral becomes $\int_{S_t} (\nabla \cdot \boldsymbol{\sigma}) \, \mathrm{d}^3\mathbf{x}$, so that

$$\int_{S_t} \rho \frac{\mathrm{D}\, \mathbf{u}}{\mathrm{D}\, t} \, \mathrm{d}^3\mathbf{x} = \int_{S_t} (\mathbf{f} + \nabla \cdot \boldsymbol{\sigma}) \, \mathrm{d}^3\mathbf{x} = \int_{S_t} (\mathbf{f} - \nabla p) \, \mathrm{d}^3\mathbf{x}. \tag{18}$$

Since this must hold for all such fluid parcels, we arrive at the second Euler equation:

$$\rho \frac{\mathrm{D}\, \mathbf{u}}{\mathrm{D}\, t} = \mathbf{f} - \nabla p. \tag{19}$$

The Navier-Stokes Equations of Motion for Viscous Fluids

The Euler equations assume stresses incident only normal to the surface of a fluid. Real fluids, however, are hardly as ideal. We can correct the equations of motion by modifying the stress tensor σ_{ij} to contain additional stresses unrelated to pressure. By Cauchy's Theorem (see [2]), one can prove that these non-pressure forces, represented by the *deviatoric tensor*, act linearly on the normal vector. Then we can split up the stress tensor into a sum of two terms: (i) the stresses normal to the surfaces of the fluid given by the pressure; and (ii) the stresses acting at arbitrary directions along the surface of the fluid. Mathematically, $\sigma_{ij} = -p\delta_{ij} + d_{ij}$, where d_{ij} denotes the deviatoric tensor. The implementation of this stress tensor to obtain the equations of motion is analogous to the method in obtaining the Euler equations. But before that, we need to first obtain the form of the deviatoric tensor.

By a physical argument, we find that the deviatoric, and hence the stress tensor, is symmetric. From Figure 1, one can see that the torque about the z-axis of a cube centered at the origin is $\alpha^3(\sigma_{21} - \sigma_{12})$, where α is the side-length of the cube. From elementary physics, we know that the moment of inertia of such a cube is of order α^4, so that the angular acceleration is proportional to $\alpha^{-1}(\sigma_{21} - \sigma_{12})$. In the limit of a fluid parcel $\alpha \to 0$, the angular acceleration remains finite only if $\sigma_{21} = \sigma_{12}$. A similar computation of the torque about the other axes allows one to conclude that the tensor is symmetric.

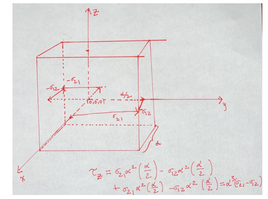


Figure 1.

For momentum to be conserved, the force on the fluid must be proportional to the gradient of the velocity. See the discussion on section 6.1 of [1]. As a result, the deviatoric tensor is a linear function of the deformation tensor $D_{ij} = (1/2)(\partial u_i/\partial x_j + \partial u_j/\partial x_i)$. This also means that d_{ij} and D_{ij} are simultaneously diagonalisable. By permuting the eigenvalues under rotations, the requirement that d_{ij} be isotropic forces the deviatoric tensor to take the form

$$d_{ij} = \lambda(\nabla \cdot \mathbf{u})\delta_{ij} + \mu \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right). \tag{20}$$

With the full form of the deviatoric tensor, we can derive the equations of motion. By Newton's second law, we obtain (17) except now with our corrected stress tensor. The only new addition is the deviatoric tensor term $\int_{\partial S_t} (\mathbf{d} \cdot \hat{n}) \, \mathrm{d}^2\mathbf{x}$, which, by the Divergence theorem, becomes $\int_{S_t} (\nabla \cdot \mathbf{d}) \, \mathrm{d}^3\mathbf{x}$. Consider just one component (using index notation and Einstein summation convention):

$$(\nabla \cdot \mathbf{d})_i = \frac{\partial d_{ij}}{\partial x_j} = \lambda \delta_{ij} \frac{\partial}{\partial x_j} (\nabla \cdot \mathbf{u}) + \mu \frac{\partial}{\partial x_j} \frac{\partial u_j}{\partial x_i} + \mu \frac{\partial^2 u_i}{\partial x_j^2} \tag{21}$$

$$= \lambda \delta_{ij} \frac{\partial}{\partial x_j} (\nabla \cdot \mathbf{u}) + \mu \frac{\partial}{\partial x_i} \frac{\partial u_j}{\partial x_j} + \mu \nabla^2 u_i \tag{22}$$

$$= \lambda \left(\delta_{i1} \frac{\partial}{\partial x_1} + \delta_{i2} \frac{\partial}{\partial x_2} + \delta_{i3} \frac{\partial}{\partial x_3} \right) (\nabla \cdot \mathbf{u}) + \mu \frac{\partial}{\partial x_i} (\nabla \cdot \mathbf{u}) + \mu \nabla^2 u_i \tag{23}$$

Upon specifying the index i, the first and second terms of (23) are really the same thing, up to the viscosity coefficients λ and μ. From this, we obtain that

$$\nabla \cdot \mathbf{d} = (\lambda + \mu) \nabla (\nabla \cdot \mathbf{u}) + \mu \nabla^2 \mathbf{u}. \tag{24}$$

Returning to Newton's second law, we have

$$\int_{S_t} \rho \frac{\mathrm{D}\, \mathbf{u}}{\mathrm{D}\, t} \, \mathrm{d}^3\mathbf{x} = \int_{S_t} \left(\mathbf{f} - \nabla p + (\lambda + \mu) \nabla (\nabla \cdot \mathbf{u}) + \mu \nabla^2 \mathbf{u} \right) \mathrm{d}^3\mathbf{x}. \tag{25}$$

Since (25) must hold for any volume, it follows that

$$\rho \frac{\mathrm{D}\, \mathbf{u}}{\mathrm{D}\, t} = \mathbf{f} - \nabla p + (\lambda + \mu) \nabla (\nabla \cdot \mathbf{u}) + \mu \nabla^2 \mathbf{u}. \tag{26}$$

And so, we arrive at the *Navier-Stokes* equations for a viscous fluid!

References

[1] Stephen Childress. *An introduction to theoretical fluid mechanics*, volume 19. American Mathematical Soc., 2009.

[2] Alexandre Joel Chorin and Jerrold E Marsden. *A mathematical introduction to fluid mechanics*, volume 3. Springer, 1990.

[3] Jerrold E Marsden and Thomas JR Hughes. *Mathematical foundations of elasticity*. Courier Corporation, 1994.

Lebesgue Measure and Integration

Vardan Martirosyan Jordan Russo

University of California, Santa Barbara

Abstract and Background

Traditionally, the integral of a non-negative single valued function is defined to be the area under the smooth curve of the function, from a start point a to an end point b on the real number line. In undergraduate courses, this concept is formalized as the Riemman Integral. After proving numerous results and theorems relating to Riemman Integration, as well as extending it to multiple dimensions, it is shown that the Riemman Integral has some limitations: namely, there are severe issues when dealing with point-wise limits, and integrating sequences. To resolve these issues, the concept of the Lebesgue Measure and Lebesgue Integration is introduced at the end of undergraduate and the beginning of graduate courses. For our project, we studied the Lebesgue Measure and Lebesgue Integration from the textbook "Real Analysis" by H. Royden and P. Fitzpatrick.

Definitions

We first define several important terms:

- **Open, Non-Empty, and Bounded Sets** A *open* set is a set such that, for any point in the set, and any given distance, a point of the set can be found between the given point and distance. A *non-empty* set is a set that has at least one element contained within it. A *bounded* set is a set that is of a finite size.
- **Complement of a Set** Let E be a set of points. The *complement* of E, denoted by E^c, is the set of points that are not in E. We note that $E \cap E^c$, the intersection of E and it's complement, is the empty set $\emptyset$. Additionally, the union of E and E^c is all of the points U that are being looked at.

Length

Consider the extended real number line, which spans $\mathbb{R}$, the set of real numbers, combined with $-\infty$ and $+\infty$. Let I be an interval on the extended real number line. We define the **length of** I to be the difference of it's endpoints if it is bounded, and to be ∞ if it is unbounded. We call the length function a **set** function, which is a function that assigns an extended real number to each set in a collection of sets.

Outer Measure

Before being able to define the Lebesgue Measure, we first have to define a separate measure, called the **outer measure**. Let A be a set of real numbers. Consider the countable collections $\{I_k\}_{k=1}^{\infty}$ of nonempty, open, bounded intervals that cover A. For each collection, consider the sums of the lengths of the intervals within the collection. Since the lengths are forced to be positive numbers, each sum is uniquely defined independently of the order of the terms. We can then define the **outer measure** of A, $m^*(A)$, to be the infimum of all types of sums:

$$m^*(A) = \inf\left\{\sum_{k=1}^{\infty} l(I_k) | A \subseteq \cup_{k=1}^{\infty} I_k\right\} \tag{1}$$

Measurable Functions and the Lebesgue Measure

Let E be a set. We define E to be measurable if for any set A, we have the following to be true:

$$m^*(A) = m^*(A \cap E) + m^*(A \cap E^c) \tag{2}$$

All sets that satisfy the above equation make up a Borel sigma algebra. Then, the **Lebesgue Measure** is the restriction of the set function outer measure to this class of measurable sets. We denote the Lebesgue measure by m, and write that $m(E) = m^*(E)$. We note that the Lebesgue measure is not defined on all subsets of $\mathbb{R}$: only those that satisfy the above equation. (A proof of why not all subsets of $\mathbb{R}$ are measurable comes from Vitali's Theorem).

Properties of the Lebesgue Measure

There are several key properties that the Lebesgue Measure contains, which we will now describe:

- **The Measure of An Interval is it's Length** Let I be an arbitrary non-empty interval. Then, I is Lebesgue Measurable and:
$$m(I) = l(I) \tag{3}$$
where l is the 'set' length function described earlier.
- **Lebesgue Measure is Translation Invariant** Let E be a Lebesgue measurable set, and y be any number. Then, the translation of E by y, $E + y = \{x + y | x \in E\}$, is also Lebesgue measurable and:
$$m(E + y) = m(E) \tag{4}$$
We display a picture to illustrate this property:

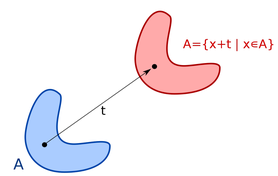


Figure 1. A picture displaying what translation invariance looks like. The picture comes from Wikipedia.

- **Lebesgue Measurable is Countably Additive Over Countable Disjoint Unions of Sets** Let $\{E_k\}_{k=1}^{\infty}$ be a countable disjoint collection of Lebesgue measurable sets. Then, we have that:
$$m(\cup_{k=1}^{\infty} E_k) = \sum_{k=1}^{\infty} m(E_k) \tag{5}$$
We note that one of the key differences between the outer measure defined earlier and the Lebesgue Measure is that in the equation above, the outer measure has sub-additive property, which is less powerful than the additive property stated above.

Lebesgue Measurable Functions

An extended real-valued function defined on a set E is said to be **Lebesgue measurable**, provided it's domain E is measurable, and it satisfies one of the following two conditions:

- For each real number c, the set $\{x \in E | f(x) > c\}$ is measurable.
- For each real number c, the set $\{x \in E | f(x) \geq c\}$ is measurable.

Characterizations and Properties of Measurable Functions

- A function f is measurable if and only if for each open set O, the inverse image of O under f is measurable.
- A real valued function that is continuous on it's measurable domain is measurable.
- A monotone function that is defined on an interval is measurable.
- Linear combinations, products, and compositions of finite measurable functions on the same set E are also measurable on the set E.
- A non-negative measurable function is the limit of a sequence of simple functions.

Lebesgue Integration

Characteristic Functions For any set A, we define the characteristic function of A on the real numbers, denoted by χ_A, as:

$$\chi_A(x) = \begin{cases} 1, & \text{if } x \in A \\ 0, & \text{if } x \notin A \end{cases} \tag{6}$$

Simple Functions Let ϕ be a real valued function defined on a measurable set E. It is called a simple function if it is measurable and takes a finite number of values. Any simple function can be represented as a linear combination of characteristic functions:

$$\psi = \sum_{k=1}^{n} c_k \cdot \chi_{E_k}, \quad E_k = \{x \in E |\ \psi(x) = c_k\} \tag{7}$$

Integration of Simple Functions For a simple function ψ defined on a set E where $m(E) < \infty$, we defined the integral of ψ over E by:

$$\int_E \psi = \sum_{i=1}^{n} a_i \cdot m(E_i) \tag{8}$$

Lebesgue Integration For a bounded real-valued function f defined on a set E where $m(E) < \infty$, we define the lower and upper Lebesgue Integral, respectively, of f over E to be:

$$\sup\left\{\int_E \psi \;\middle|\; \psi \text{ simple},\ \psi \leq f \text{ on } E\right\} \quad \text{and} \quad \inf\left\{\int_E \phi \;\middle|\; \phi \text{ simple},\ f \leq \phi \text{ on } E\right\} \tag{9}$$

We say that f is **Lebesgue Integrable** over E when its lower and upper Lebesgue integrals over E are equal. We call the common value the **Lebesgue Integral** of f over E, and denote it by:

$$\int_E f \tag{10}$$

Advantages over the Riemman Integral

Monotone convergence Theorem Suppose we have a sequence of non-negative measurable functions $\{f_n\}$ on a measurable set X such that f_n converges pointwise to f almost everywhere and $f_1 \leq f_2 \leq \cdots \leq f_n$. The Monotone Convergence Theorem gives us the following property for Lebesgue integration:

$$\lim_{n\to\infty} \int_X f_n = \int_X \lim_{n\to\infty} f_n = \int_X f \tag{11}$$

Under the Riemman Integral, the ability to move the limit inside the integral requires uniform convergence, while under the Lebesgue Intregral, we only require pointwise convergence. We now give an example to illustrate the use of the Monotone Convergence Theorem. Let a_{ij} be an non-negative real valued sequence of numbers. Then, we have that:

$$\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} a_{ij} = \sum_{j=0}^{\infty}\sum_{i=0}^{\infty} a_{ij} \tag{12}$$

Acknowledgements

We would like to Pranav Arrepu for his guidance in helping us understand the material during this program. Additionally, we would like to thank the Directed Reading Program at UCSB for giving us the opportunity to participate in this program.

Sticking a 'PINN' in It: A Physics-Informed Neural Network Approach to PDEs

Yan Lashchev [1] Bill Nguyen [1] Mentor: Rafael Lainez Reyes [1]

[1]Department of Mathematics, University of California - Santa Barbara

UCSB

Abstract

Recent successes in neural networks have greatly encouraged their use in solving classical problems in applied mathematics, as the networks allow for rapid prototyping with usable estimations. This holds especially true in areas involving high dimensional partial differential equations (PDEs), such as quantum physics and fluid dynamics. Here, we present a neural network architecture, the physics-informed neural network (PINN), and implement a specific method, the continuous time approach.

Background

We describe the PINN approach for approximating the solution

$$u : [0,T] \times \mathcal{D} \to \mathbb{R} \quad (\star)$$

of an evolution equation

$$\partial_t u(t,x) + \mathcal{N}[u](t,x) = 0, \qquad (t,x) \in (0,T] \times \mathcal{D}, \tag{1a}$$
$$u(0,x) = u_0(x), \qquad x \in \mathcal{D}, \tag{1b}$$

where $\mathcal{N}$ is a differential operator acting on u, $\mathcal{D} \subset \mathbb{R}^d$ a bounded domain, T denotes the final time and $u_0 : \mathcal{D} \to \mathbb{R}$ the prescribed initial data. Based on the literature review conducted, we restrict our discussion to the Dirichlet case and define

$$u(t,x) = u_b(t,x), \quad (t,x) \in (0,T] \times \partial\mathcal{D}, \tag{1c}$$

where $\partial\mathcal{D}$ denotes the boundary of the domain $\mathcal{D}$ and $u_b : (0,T] \times \partial\mathcal{D} \to \mathbb{R}$ the given boundary data. The method constructs a neural network approximation $u_\theta(t,x) \approx u(t,x)$ of the solution of (1), where $u_\theta : [0,T] \times \mathcal{D} \to \mathbb{R}$ denotes a function realized by a neural network with parameters θ.

Continuous Time Approach

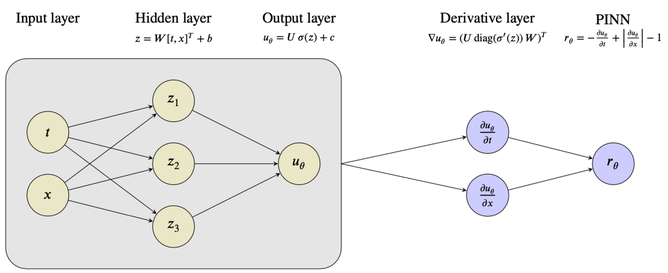


Figure 1) Neural network architecture of the PINN approach

The (strong) residual of a given neural network approximation of $(\star)$ with respect to the PINN approach above is

$$r_\theta(t,x) := \partial_t u_\theta(t,x) + \mathcal{N}\left[u_\theta\right](t,x) \tag{2}$$

These networks are compositions of alternating affine linear $W^\ell \cdot + b^\ell$ and nonlinear functions $\sigma^\ell(\cdot)$ called activations, i.e.,

$$u_\theta(z) := W^L \sigma^L \left(W^{L-1}\sigma^{L-1}\left(\cdots \sigma^1\left(W^0 z + b^0\right)\cdots\right) + b^{L-1}\right) + b^L,$$

where W^ℓ and b^ℓ are weight matrices and bias vectors, and $z = [t,x]^T$.

PINN Approach

For the solution of the PDE (1) now proceeds by minimization of the loss functional

$$\phi_\theta(X) := \phi_\theta^r\left(X^r\right) + \phi_\theta^0\left(X^0\right) + \phi_\theta^b\left(X^b\right), \tag{3}$$

where X denotes the collection of training data and the loss function ϕ_θ contains the following terms:

The Mean Squared Residual

$$\phi_\theta^r\left(X^r\right) := \frac{1}{N_r}\sum_{i=1}^{N_r}\left|r_\theta\left(t_i^r, x_i^r\right)\right|^2$$

in a number of collocation points $X^r := \left\{\left(t_i^r, x_i^r\right)\right\}_{i=1}^{N_r} \subset (0,T] \times \mathcal{D}$, where r_θ is the physics-informed neural network (2),

The Mean Squared Misfit *w.r.t* Initial and Boundary Conditions

$$\phi_\theta^0\left(X^0\right) := \frac{1}{N_0}\sum_{i=1}^{N_0}\left|u_\theta\left(t_i^0, x_i^0\right) - u_0\left(x_i^0\right)\right|^2 \quad \text{and} \quad \phi_\theta^b\left(X^b\right) := \frac{1}{N_b}\sum_{i=1}^{N_b}\left|u_\theta\left(t_i^b, x_i^b\right) - u_b\left(t_i^b, x_i^b\right)\right|^2$$

in a number of points $X^0 := \left\{\left(t_i^0, x_i^0\right)\right\}_{i=1}^{N_0} \subset \{0\} \times \mathcal{D}$ and $X^b := \left\{\left(t_i^b, x_i^b\right)\right\}_{i=1}^{N_b} \subset (0,T] \times \partial\mathcal{D}$, where u_θ is the neural network approximation of the solution $u : [0,T] \times \mathcal{D} \to \mathbb{R}$.

Example: Heat Equation

A classical problem in the domain of PDEs, the heat equation governs the temperature distribution of a rod of length l :

$$u_t = k u_{xx} \qquad (t,x) \in \mathbb{R}^+ \times (0,l)$$
$$u(t,0) = u(t,l) = 0 \qquad t \geq 0$$
$$u(0,x) = f(x) \qquad x \in (0,l).$$

If k, called the conductivity is a constant the rod is isotropic; if $k = k(x)$ it is anisotropic or heterogeneous medium.

Application

With respect to the fitting, we choose $k = 1, l = \pi$, and $f(x) = \sin(3x)$ for the demo of the PINN.

We assume that the collocation points X_r as well as the points for the initial time and boundary data X_0 and X_b are generated by random sampling from a uniform distribution.

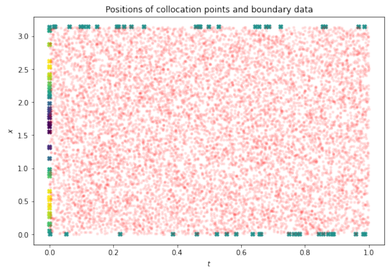
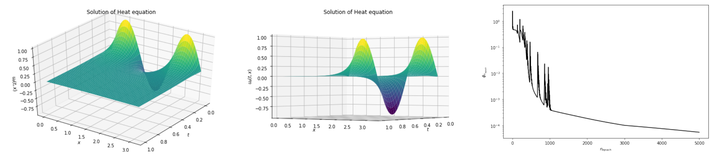


Figure 2) Plot of the collocation points ($N = 10,000$)

PINN Approximation and Evolution of Loss



(a) View 1 (b) View 2 (c) Loss

Test

The chosen problem can be solved via separation of variables. The idea is to assume the solution $u = u(t,x)$ can be written as

$$u(t,x) = F(t)G(x)$$

If we compute the corresponding partial derivatives and replace in the PDE, we get

$$\frac{F'(t)}{F(t)} = \frac{G''(x)}{G(x)}$$

The only way this equality is true for all t and x is if

$$F'(t) = \lambda F(t) \quad \text{and} \quad G''(x) = \lambda G(x)$$

The boundary condition becomes

$$G(0) = G(\pi) = 0$$

We can easily solve these ordinary differential equations. By considering the cases $\lambda > 0, \lambda = 0$ and $\lambda < 0$, we conclude $\lambda = -n^2, n \in \mathbb{N}$ and (up to constants)

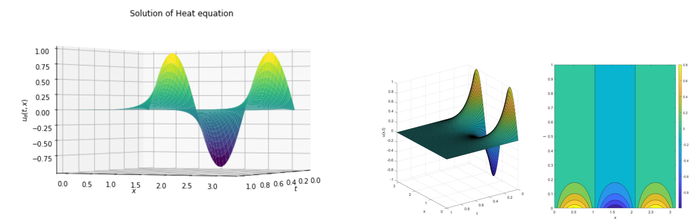
$$F(t) = \exp\left(-n^2 t\right) \quad \text{and} \quad G(x) = \sin(nx)$$

Since the equation is linear, by the principle of superposition $u(t,x) = \sum_{n=1}^{\infty} c_n \exp\left(-n^2 t\right)\sin(nx)$

Finally, since $u(0,x) = \sin(3x) = \sum_{n=1}^{\infty} c_n \sin(nx)$ with $c_3 = 1$ and $c_n = 0$ if $n \neq 3$. Hence,

$$u(t,x) = \exp(-9t)\sin(3x)$$

True Solution



(a) PINN (b) True Solution

References

[1] Jan Blechschmidt Oliver Ernst. Three ways to solve partial differential equations with neural networks — a review. *GAMM-Mitteilungen*, 44(2), 2021.

[2] Peter Olver. Introduction to partial differential equations. *Springer*, 2020.

Primitive Parking Functions and Non-Crossing Partitions

Yanru Liu, Mentored by **Sam Sehayek**

2022 Mathematics Direct Reading Program. University of California, Santa Barbara

Parking Functions

Imagine living on a one-way street that dead-ends with n parking spots available. You and your neighbors have n cars in total, and everyone has their preferred spot to park. Without reversing, does there exist a solution that everyone can park without collision? In mathematics, this real life dilemma is called the parking problem. Consider this set up:

- There are n cars and n parking spots on a straight street (n is a positive integer, $n \in \mathbb{Z}^+$; and i denotes the i-th spot, $i \in \{1, \cdots, n\}$)
- C_i is the i-th car to park, having preferred spot $\alpha_i \in \{1, \cdots, n\}$. More than one car can have the same preference.
- If the preferred spot had already been occupied, then the car will move forward and park in the next available spot. No backward movement allowed.

If all n cars can be parked under these conditions, then the preference list $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n)$ is a parking function.
Equivalently, an n-tuple of integers $\alpha = (\alpha_1, \cdots, \alpha_n)$ is a **parking function** if and only if $\beta_i \leq i$, where $\beta = \{\beta_1, \cdots, \beta_n\}$ is a reordering of α into weakly increasing order. i.e. $\beta_1 \leq \cdots \leq \beta_n$.
For n cars, how many parking functions are there?

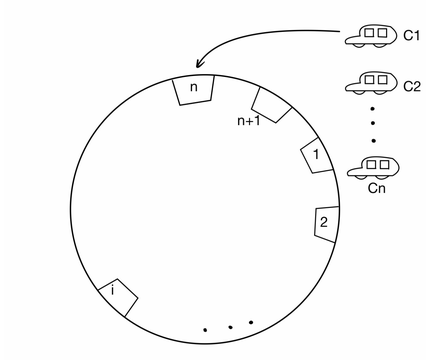


Fig. 1: Visual Representation of H.Pollack's Proof.

Regard the elements of the group $G = \mathbb{Z}/(n+1)\mathbb{Z}$ as being the integers $0, 1, \cdots, n$. Let H be the (cyclic) subgroup of order $n+1$ of the group G^n generated by $(1, 1, \cdots, 1)$. Each coset of H contains exactly one parking function. Let $f(n)$ be the number of parking functions of length n, hence we have

Theorem 1 (Konheim and Weiss, 1966)**.** *The number of parking functions of length n is*

$$f(n) = (n+1)^{n-1}.$$

Primitive Parking Functions

A parking function is called a **primitive parking function** if it is already in a weakly increasing order.
There is a well-known bijection between parking functions and labeled Dyke paths, wherein each distinct labeling of the same Dyke path corresponds to a permutation of the parking function. Thus, the primitive parking functions with length $n-1$ are in bijection with Dyke paths and can be enumerated by the n-th Catalan number:

$$C_n = \frac{1}{n+1}\binom{2n}{n}.$$

Non-Crossing Partitions

A **partition** of a finite set S is a collection $\{B_1, \cdots, B_k\}$ of nonempty subset $B_i \subseteq S$ s.t. $B_1 \cup \cdots \cup B_k = S$ and $B_i \cap B_j = \emptyset$ if $i \neq j$. And in our research of primitive parking functions, we especially care about a special one: the non-crossing partition. A **non-crossing partition** of set $\{1, \cdots, n\}$ is a partition $\{B_1, \cdots, B_k\}$ of $\{1, \cdots, n\}$ s.t. for $a < b < c < d$, $a, c \in B_i$, and $b, d \in B_j \Rightarrow i = j$.

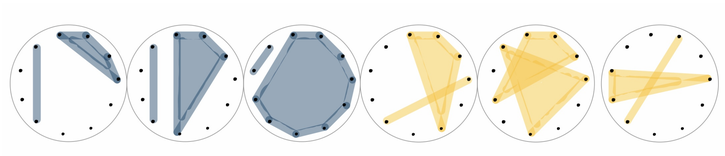


Fig. 2: Non-Crossing Partitions (Blue) & Crossing Partitions (Yellow) of $\{1, \cdots, 11\}$.

A **maximal chain** of non-crossing partitions of $\{1, \cdots, n+1\}$ is a sequence $\pi_0, \cdots, \pi_n$ of noncrossing partitions s.t. π_i is obtained from π_{i-1} by merging two blocks of π_{i-1} into a single block.
A maximal chain of $[n+1]$ has n merging steps. If we pick a label for each step, there are exactly n labels. Thus, it's possible for us to connect parking function with maximal chains.

Theorem 2. *There is a bijection between parking functions of length n and maximal chains of NC_{n+1}.*

Here is the algorithm: Let A and B be the two blocks we're going to merge at stage i, and A contains the smallest element in the disjoint union $A \cup B$. The label for this stage is the largest element in A which is smaller than all elements in B.

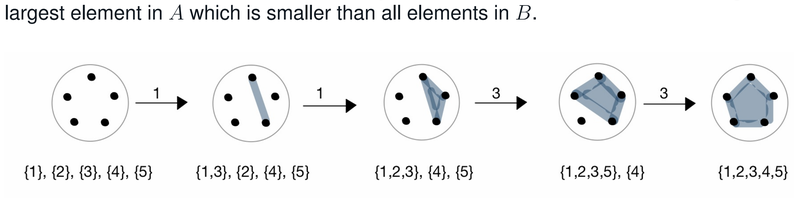


Fig. 3: One of the Maximal Chain of $\{1, \cdots, 5\}$ and Its Associated Parking Function $(1, 1, 3, 3)$.
(The top is 1 and the label goes in clockwise direction)

Every maximal chain is associating with a parking function, and only some of them are associating with the primitive ones.

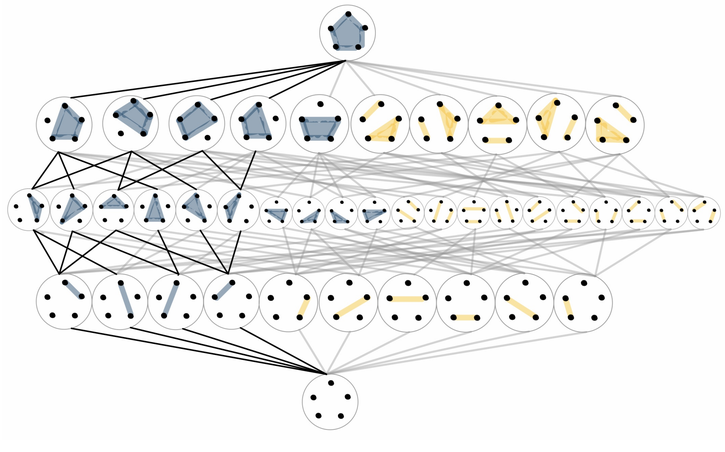


Fig. 4: Primitive Parking Functions with Length 4 (black) in the Maximal Chains of Noncrossing Partitions of $\{1, \cdots, 5\}$ (grey).

A New Proposition

For maximal chains corresponding to the primitive parking functions, do they form a certain pattern?

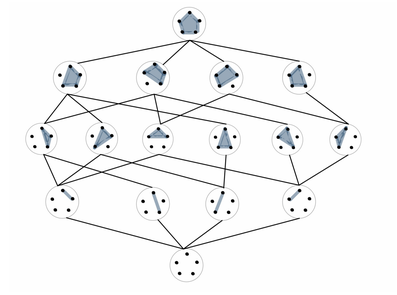
Lemma 3. *The chain starts by merging 1 with some other element.*

This lemma is trivial as the primitive parking function is always starting with 1, and only the merging between 1 and some other element gives the label 1.

Lemma 4. *The primitive parking functions are always adding one single block to the other block.*

This proposition can be verified via figure 4. From these lemmas, we can prove:

Proposition 5. *The subdiagram consisting only of nodes and edges from the primitive parking functions inside the non-crossing partition lattice is a coarsening of the Boolean lattice of size equal to the length of these parking functions.*



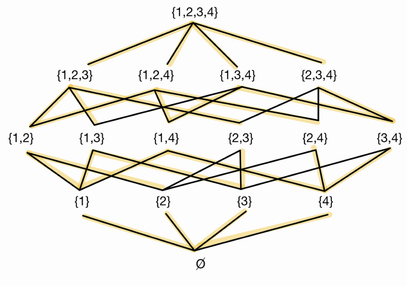


Fig. 6: Maximal Chains for Primitive Parking Functions of Length 4 (Up)
& Boolean Lattice of size 4 (Bottom).

In other words, maximal chains of primitive parking functions of length n are look just like the Boolean lattice of the same size with some relations removed.

Acknowledgements

I would like to thank my mentor Sam Sehayek, his knowledge and enthusiasm in mathematics deeply impressed me; he is such a good mentor and friend on my way of learning mathematics.

Reference

Richard P. Stanley. *Enumerative Combinatorics*. Cambridge Press, 1998.

A Brief Introduction to Network Theory

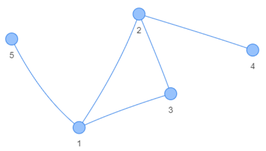
Ponokela DeMarzo

University of California, Santa Barbara

What is a Network?

A **network** is a collection of nodes where pairs of nodes may be connected by edges.

Networks can be visualized by drawing their graph structure, but they are also commonly represented by their adjacency matrix. The **adjacency matrix** for a network consisting of n nodes is an n x n matrix $\mathbf{A}$ with entries A_{ij} = the number of edges connecting node i and j.



$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Figure 1: The graph and adjacency matrix representations of the same network.

Edges can be **weighted**, where each edge is assigned a value that represents the "strength" of the connection, as well as **directed**, where a connection from node i to j does not imply a connection from j to i. Networks are also not limited to only one type of node; however, since many of the forms of analysis change, we will not be discussing them.

Node Centrality

A natural question arising from the gathered network data is determining the importance, or **centrality**, of each node. This can give us an idea of which nodes have more influence over a network. The four main centrality measures are defined below:

- **Degree Centrality:** This is perhaps the simplest measure of centrality, calculated by counting the number of edges attached to the node in question. In terms of the adjacency matrix $\mathbf{A}$, the degree centrality of a node i can be defined as follows:

$$k_i = \sum_j A_{ij}.$$

- **Eigenvector Centrality:** Unlike degree centrality, eigenvector centrality is primarily concerned with the *quality* of connections, not *quantity*. To measure centrality this way, the centrality of a node i will be proportional to the centrality of its neighbors, and thus is defined recursively like so:

$$x_i = \kappa^{-1} \sum_j A_{ij} x_j.$$

Rewritten in matrix notation, this equation becomes

$$\mathbf{x} = \kappa^{-1}\mathbf{A}\mathbf{x}, \text{ or } \mathbf{A}\mathbf{x} = \kappa\mathbf{x}.$$

In this form, it is clear that $\mathbf{x}$ is an eigenvector of $\mathbf{A}$; however, since there may be multiple eigenvectors, we generally define $\mathbf{x}$ and κ to be the leading eigenvector and eigenvalue.

- **Closeness:** This is a measure of the average distance from a node to other nodes. Suppose that d_{ij} is the shortest distance from node i to node j. Then the average distance from i to every other node is

$$\ell_i = \tfrac{1}{n-1} \sum_j d_{ij}.$$

Since we want to consider nodes that are on average closer to all other nodes as being more central, we define the closeness centrality as the inverse of ℓ_i so

$$C_i = \tfrac{1}{\ell_i} = \frac{n-1}{\left(\sum_j d_{ij}\right)}.$$

- **Betweenness:** This measures how often a given node lies on a shortest path between other nodes. Let n_{st}^i be the number of shortest paths from s to t that pass through i, and let g_{st} be the total number of shortest paths from s to t. We can then define the betweenness centrality of a node i as follows:

$$x_i = \tfrac{1}{n^2} \sum_{st} \tfrac{n_{st}^i}{g_{st}}.$$

A Social Network Example

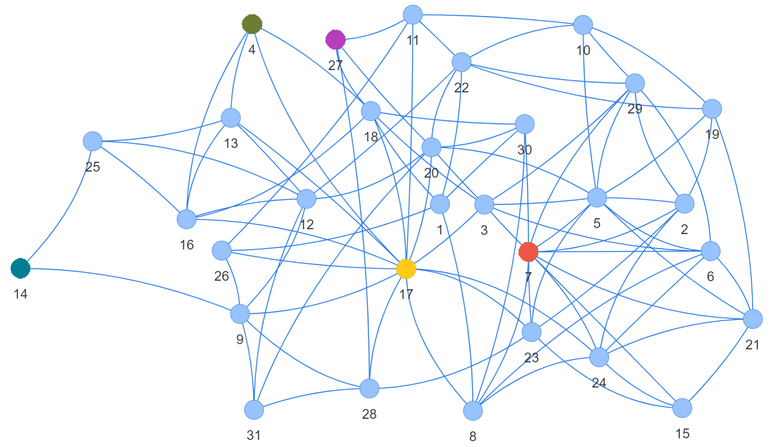


Figure 2: A social network constructed from anonymized friendship data collected by surveying a high school math classroom. The density of the network is 20.2% and the clustering coefficient is 37.2%. The largest core is a 4-core which includes all nodes except 14 and 25.

Network Analysis

Node	Degree	Eigenvector	Closeness	Betweenness	Local Clustering
1	5	0.1049	0.4615	0.0209	0.2
2	6	0.2153	0.4347	0.0116	0.4666
3	7	0.2595	0.5454	0.0525	0.4285
4	4	0.0911	0.4347	**0.0018**	**0.8333**
5	10	0.3201	0.5084	0.0657	0.3777
6	7	0.2605	0.4477	0.0107	0.5714
7	11	**0.3491**	0.4838	0.0573	0.3818
8	6	0.1972	0.4918	0.0285	0.4
9	6	0.0967	0.4687	0.0671	0.2666
10	5	0.1390	0.4347	0.0131	0.5
11	5	0.1054	0.4687	0.0324	0.2
12	7	0.1116	0.4838	0.0767	0.2857
13	5	0.0918	0.4477	0.0169	0.6
14	**2**	**0.0207**	**0.3370**	0.0021	**0**
15	4	0.1432	0.4	0.0019	0.6666
16	6	0.1111	0.4761	0.0297	0.5333
17	**13**	0.2783	**0.625**	**0.2732**	0.1538
18	7	0.1537	0.5	0.0474	0.2380
19	5	0.1485	0.4225	0.0156	0.4
20	8	0.2040	0.5555	0.0926	0.1428
21	6	0.2093	0.4285	0.0125	0.5333
22	7	0.1516	0.5084	0.0817	0.1904
23	6	0.2025	0.5084	0.0603	0.3333
24	7	0.2372	0.5	0.0573	0.4285
25	4	0.0481	0.375	0.0135	0.5
26	4	0.0840	0.4615	0.0137	0.3333
27	4	0.0815	0.4411	0.0102	**0**
28	5	0.1052	0.4687	0.0316	0.3
29	7	0.2432	0.4687	0.0247	0.4761
30	5	0.1447	0.4687	0.0171	0.3
31	4	0.0742	0.4285	0.0065	0.5

Table 1: The centrality and clustering measures for each node in Figure 2. The largest and smallest values in each column are bolded and the corresponding nodes are highlighted. Node 17 has the highest degree, closeness, and betweenness centralities, however, node 7 has the highest eigenvector centrality due to the importance of its neighbors (e.g., nodes 5 and 6). Node 4 has the highest clustering coefficient, indicating a tight-knit friend group, but has a low betweenness centrality because it is somewhat redundant in the network. Node 14 is the most isolated.

Why are Networks Useful?

Networks are a powerful analytical tool which are used across many different disciplines with a multitude of applications. Networks are an elegant representation of almost any system which consists of *objects* and *connections between those objects*, and when modeled this way, we can perform well-defined and meaningful calculations to analyze its structure.

There are four primary categories in which we can sort networks: technological, information, social, and biological. Technological networks are physical networks which are typically responsible for the transfer of data or materials, such as the Internet, waterlines, or commercial airline flights. Information networks can model the interaction of ideas, and are used to represent structures such as the World Wide Web or citation networks for academic papers. Social networks are used to model people and their interactions, such as friendships in a workplace or followers on social media. Even systems such as metabolic processes and food chains can be modeled by biological networks.

Network Structure

Beyond simply ranking nodes in accordance to their centralities, it also often important to be able to describe the overall structure of the network. Below are some basic, yet useful, measures of network structure:

- **Density:** In a simple network consisting of n nodes, we can calculate the maximum possible number of edges by counting the number of pairs of nodes given by $\binom{n}{2}$. The density of a network is the proportion of existing edges to the maximum possible edges.
- **k-cores:** A k-core is a connected set of nodes where each node is connected to at least k other nodes in the set.

Figure 3: The nodes 1-4 form a 3-core, and the nodes 1-5 form a 2-core.

- **k-components:** A k-component is a set of nodes where each node is reachable from each of the others by k node-independent paths.

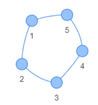


Figure 4: This network forms a 2-component.

- **Local Clustering Coefficient:** The local clustering coefficient is the proportion of the number of neighbors of a node i that are neighbors themselves. Visually, this can be thought of as the fraction of closed triangles out of all possible triangles with i as a vertex.

Figure 5: The local clustering coefficient of node 1 is $1/3$.

- **Clustering Coefficient:** This is a generalization of the local clustering coefficient to the whole network. For the whole network, it is the proportion of connected triples that are also closed triangles.

References

[1] Mark Newman.
Networks, Second Edition.
Oxford University Press, 2018.

Guided by

Sanjay Kumar
UCSB

Differential Forms and Maxwell's Equations

Samuel Zhang - Mentored by Yusen Xia

Department of Mathematics - University of California, Santa Barbara

Introduction

Differential manifolds are topological spaces that are locally homeomorphic to a vector space so that one may perform calculus on it, and a differential form allows one to define integrals over such manifolds. This poster is meant to revisit the Maxwell's Equations using such languages

Smooth Manifolds and Tangent Map

- A topological space M is called an *n-dimensional manifold* if $\forall p \in M$ there is a homeomorphism $F : U \to O$ such that $U \subset \mathbb{R}^n$ is non-empty and open, and $O \subset M$ is an open subset containing p. Such an F is called a *local parametrization* around p
- An n-dimensional manifold is called an *n-dimensional smooth manifold* if there is a collection of local parametrizations $F_\alpha : U_\alpha \to O_\alpha$ such that
  - $\cup U_\alpha = M$ (such parametrizations cover all of M)
  - Any transition map $F_\alpha^{-1} \circ F_\beta$ is smooth on their domain

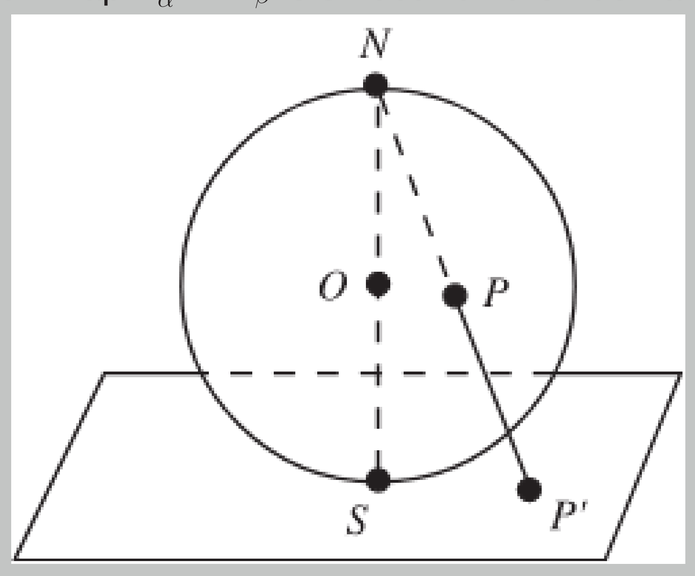


Figure 1:The stereographic projection of a sphere, which is a parametrization of a sphere except at the north pole. The sphere is an example of a 2-dimensional smooth manifold

- Given a smooth n-dimensional manifold M and a local parametrization $F : U \to M$, the *tangent space* at p is defined as $T_pM = span\{\frac{\partial}{\partial u_1}(p), ..., \frac{\partial}{\partial u_n}(p)\}$, with each $\frac{\partial}{\partial u_i}$ being the differential operator with respect to $F(u_1, ..., u_n)$ at p
- For a vector space V, define its *dual space* $V^* = \{T | T : V \to \mathbb{R}\}$
  - Moreover, V^* is a vector space itself. Given a basis of V as $\{e_1, ..., e_n\}$, a basis of V^* is $\{e^1, ..., e^n\}$ such that $e^i(e_j) = \delta_{ij}$
  - One can define the *cotangent space* of M at p as $T_p^*M = (T_pM)^*$ and any $v^* \in T_p^*M$ is called a *cotangent vector* of M at p. In the same fashion, one can express $T_p^*M = span\{du^1, ..., du^n\}$ such that $du^i[\frac{\partial}{\partial u_j}(p)] = \delta_{ij}$

Tensor Product and Wedge Product

- For V, W being two vector spaces, $T \in V^*$ and $S \in W^*$, the *tensor product* between T and S is defined as $T \otimes S : V \times W \to \mathbb{R}$ such that $(T \otimes S)(X, Y) = T(X)S(Y)$
- In the same setup, the *wedge product* is defined as $T \wedge S = T \otimes S - S \otimes T$
  - One can see that a wedge product is alternating; $T \wedge S = -S \wedge T$
  - Also noted that $T \wedge T = 0$

Differential Form

- Let M be a smooth manifold. The *smooth differential k-form* w on M is defined as $w : T_pM \times T_pM \times ... \times T_pM$(k times)$\to \mathbb{R}$ such that for any local parametrization $F : U \to M$, $w = \sum_{i_1,...,i_k=1}^n w_{i_1 i_2 ... i_k} du^{i_1} \wedge ... \wedge du^{i_k}$. The $w_{i_1 i_2 ... i_k}$'s are scalar functions locally defined in $F(U)$ and are called the *local components* of w
  - For example, in $\int_a^b f(x)dx$ the $f(x)dx$ is a differential 1-form

Exterior Derivative

- Given a smooth differential k-form w, its *exterior derivative* is defined as

$$dw = \sum_{i_1,...,i_k=1}^{n} dw_{i_1 i_2 ... i_k} \wedge du^{i_1} \wedge ... \wedge du^{i_k}$$
$$= \sum_{i_1,...,i_k=1}^{n} \sum_{j=1}^{n} \frac{\partial w_{i_1 i_2 ... i_k}}{\partial u_j} du^j \wedge du^{i_1} \wedge ... \wedge du^{i_k}$$

- Given smooth differential k-forms w, η on a smooth manifold M and a smooth scalar function f,
  - $d(w + \eta) = dw + d\eta$
  - $d(fw) = df \wedge w + d\eta$
  - $d^2w = d(dw) = 0$
- A connection between a differential form and exterior derivative in $\mathbb{R}^3$ and usual multivariable calculus is shown below:

Differential Form on $\mathbb{R}^3$	**Multivariable Calculus**
$f(x, y, z)$	$f(x, y, z)$
$w = Pdx + Qdy + Rdz$	$F = P\hat{i} + Q\hat{j} + R\hat{k}$
$\beta = Ady \wedge dz + Bdz \wedge dx + Cdx \wedge dy$	$G = A\hat{i} + B\hat{j} + C\hat{k}$
df	∇f
dw	$\nabla \times F$
$d\beta$	$\nabla \cdot G$
$d^2f = 0$	$\nabla \times \nabla f = 0$
$d^2w = 0$	$\nabla \cdot (\nabla \times F) = 0$

Revisit Maxwell's Equations

- The Maxwell's Equations can be written in differential equations as:

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}$$
$$\nabla \cdot \mathbf{B} = 0$$
$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$$
$$\nabla \times \mathbf{B} = \mu_0 \mathbf{j} + \frac{1}{c^2}\frac{\partial \mathbf{E}}{\partial t}$$

- The first equation is the Gauss's law on electric field, the second equation is a statement that an magnetic monopole does not exist (it has been predicted in several models but not yet verified), the third equation is the law of electromagnetic induction, and the fourth equation is the Ampere's circuit law with Maxwell's correction
- Denote $(t, x, y, z) \in \mathbb{R}^4$ as (x_0, x_1, x_2, x_3) and take w be a k-form on $\mathbb{R}^4$ (here $k = 0, 1, 2, 3$ or 4). Define the *Hodge-star* map from a k-form to (4-k)-form such that $w \wedge *w = dt \wedge dx \wedge dy \wedge dz$, or $-dt \wedge dx \wedge dy \wedge dz$ if w contains a dt term (this is known as the volume form of the Minkowski spacetime)
- Express $\mathbf{E}, \mathbf{B}, \mathbf{J}$ as

$$\mathbf{E} = E_x dx + E_y dy + E_z dz$$
$$\mathbf{B} = B_x dy \wedge dz + B_y dz \wedge dx + B_z dx \wedge dy$$
$$\mathbf{J} = -(J_x dy + J_y dz \wedge dx + J_z dx \wedge dy) \wedge dt + \rho dx \wedge dy \wedge dz$$

Define $\mathbf{F} \equiv \mathbf{B} + \mathbf{E} \wedge dt$. Together with the Hodge-star map, one can rewrite the Maxwell's equations as:

$$d\mathbf{F} = 0$$
$$d(*\mathbf{F}) = \mathbf{J}$$

References

[1] T.H. Fong.
Differentiable Manifolds & Riemannian Geometry.
Hong Kong University of Science and Technology, 7th edition, 2021.

UNIVERSAL APPROXIMATION THEOREM

Xin Wang, Ruifan Wang

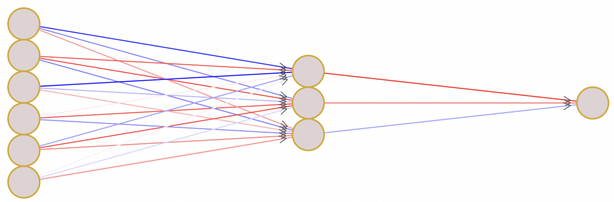
University of California, Santa Barbara

UC SANTA BARBARA

Multiplayer Feedforward Neural Networks

A neural network is the interconnection of unit models characterized by a threshold value θ, a univariate **activation function** $\sigma : R \to R$, and a vector of weights $w = w_1, ..., w_n$. Here, the value of n is determined by the dimension of the input-vector $x = x_1, ..., x_n$. When we feed x into a unit, it computes $\sigma(w \cdot x - \theta)$ and shoots the result to the next unit. A single hidden layer feedward neural network represents a $f : R^n \to R$ function

$$f(x) = \sum_{j=1}^{k} \beta_j \cdot \sigma(w_j \cdot x - \theta_j)$$



Input Layer $\in \mathbb{R}^6$ Hidden Layer $\in \mathbb{R}^3$ Output Layer $\in \mathbb{R}^1$

The most important application of neural networks is in machine learning, where neural networks are "trained" to approximate a function. Thus, a fundamental question for neural networks is whether they can approximate reasonable functions to an arbitrary degree of accuracy. This depends on the activation function σ and is the subject of many papers, including the paper studied for this project.

Nonpolynomial Activation Function

Leshno et al. proved in their paper "Multilayer Feedforward Networks With a Nonpolynomial Activation Function Can Approximate Any Function" [2] that, under modest assumptions, a broad class of activation functions are suitable for building neural networks to approximate continuous functions. We studied this paper to understand the mathematics underlying the result.

Definition (Notion of approximation)**.** We say that a set F of functions in $L^{\infty}_{loc}(R^n)$ is **dense** in $C(R^n)$ if for every function $g \in C(\mathbb{R}^n)$ and for every compact set $K \subset \mathbb{R}^n$, there exists a sequence of functions $f_j \in F$ such that

$$\lim_{j \to \infty} ||g - f_j||_{L^{\infty}(K)} = 0$$

Colloquially, $\{f_j\}$ approximates g "arbitrarily well."

Definition. The **admissible class of activation functions** which Leshno et al. denote by M is the set of locally bounded functions with a "small" number of discontinuities: if $\sigma \in M$ and K is the collection of discontinuities of σ, then $\overline{K}$ has zero Lebesgue measure.

Neural networks arise from the collection,

$$\Sigma_n = \text{span}\{\sigma(w \cdot x + \theta) : w \in \mathbb{R}^n, \theta \in \mathbb{R}\}$$

and the main result of the paper is that Σ_n is dense in $C(\mathbb{R}^n)$ if and only if σ is not an **algebraic polynomial**. This is a novel conclusion since the condition is very simple.

Reduced Case

There are two directions to prove, one of which is not difficult: *Σ_n is dense in $C(\mathbb{R}^n)$, then σ is not a polynomial.* The rest of the proof aims to show that if σ is not a polynomial, then Σ_n is dense in $C(\mathbb{R}^n)$. The proof relies on some analysis tricks, which we summarize here. Some common techniques to prove results like this are,

- Reduce the dimension of the space(s) considered.
- Prove the result for well-behaved functions first.
- Use a "smoothing" technique to deal with functions lacking regularity.

The complexity of the problem is reduced by showing first that if $\overline{\Sigma_1} = C(\mathbb{R})$, then $\overline{\Sigma_n} = C(\mathbb{R}^n)$. Then, Leshno et al. prove $\overline{\Sigma_1} = C(\mathbb{R})$ in the case that $\sigma \in C^{\infty}$.

To show $\overline{\Sigma_1} = C(\mathbb{R})$ when $\sigma \in C^{\infty}$, Leshno et al. show that $\overline{\Sigma_1}$ contains all polynomials. The result follows then as a consequence of Weierstrass's Theorem:

Theorem (Weierstrass's Theorem[3])**.** *If f is a continuous function on a compact set K, there exists a sequence of polynomials P_n such that*

$$\lim_{n \to \infty} P_n(x) = f(x)$$

uniformly on K.

It follows that $\overline{\Sigma_1}$ contains $C(K)$, where all $K \subset \mathbb{R}$. Hence, Σ_1 is dense in $C(\mathbb{R})$.

Generalized Case

From above steps, the "dense" argument can be easily achieved when σ is smooth. In this section, the author generalizes the problem to the entire class of admissible activation function by considering σ that is not smooth. The purpose of **convolution** $\sigma * \varphi$ is to deal with the *discontinuities* and points where σ *is not differentiable*. In a way, the convolution can overcome the limited differentiability of σ. We will discuss the merit of convolution in the next section.

By convolving σ with functions $\varphi \in C_0^{\infty}$, the general case follows as a consequence of the work for the reduced case: $\overline{\Sigma_1}$ is dense in $C(\mathbb{R})$ so long as $\sigma * \varphi$ is not a polynomial for some test function φ. The authors deal with this caveat using advanced techniques.

Basically, Leshno et al. must knows what is the condition that makes $\sigma * \varphi$ a polynomial for every test function φ. It turns out that this only occurs if σ is a polynomial almost everywhere, which rules out any strange conditions where $\sigma * \varphi$ is a polynomial for some $\varphi \in C_0^{\infty}$, yet σ is not a polynomial. Their key argument is to show that if $\sigma * \varphi$ is a polynomial for every test function φ, then the degree of $\sigma * \varphi$ is bounded by some m for every φ. From here, they conclude that since $\sigma * \varphi$ is a polynomial of degree at most m for every test function φ, andσ itself must be (almost everywhere) a polynomial of degree at most m.

Convolution Applied to a Specific Example

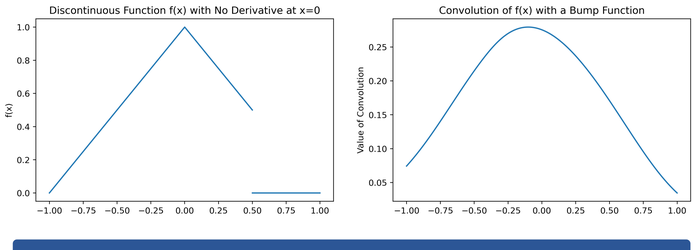
To illustrate the utility of the convolution, let $f(x) := 1 - |x|$ when $-1 \leq x \leq 1/2$ and 0 otherwise. Let $g(x)$ be a bump function, where $g(x) = e^{-1/(1-x^2)}$ for $|x| \leq 1$ and 0 otherwise. The convolution: $(f * g)(x)$, is defined as,

$$(f * g)(x) = \int_{-\infty}^{\infty} f(x - y)g(y)dy = \int_{-\infty}^{\infty} f(y)g(x - y)dy$$

The equality of the integrals above follows by a change of variables. Since f and g are supported on a compact set, we will write the convolution as,

$$(f * g)(x) = \int_{-1}^{1/2} (1 - |y|)e^{-1/(1-(x-y)^2)}dy$$

Then, we may use a numerical integrator to evaluate the convolution.
Here is a visual comparison between the discontinuous function $f(x)$ and the convolution of $f(x)$ with $g(x)$.



Remarks

In 1991, Hornik showed that the multilayer feed-ward architecture gives neural networks the potential of being universal approximators[1]. Leshno et al. led the study to a new dimension and discovered that a neural network does not need a continuous activation function to approximate some real-world functions in an arbitrary accuracy. This endows the neural network a biological interpretation because a real neuron is unlikely to have a continuous activation function. Later in the history, mathematicians extends the **Universal Approximation Theorem** by studying discontinuous functions, noncompact domains, and so on.

Acknowledgement

We especially appreciate our mentor Zach Wagner for enlightenment and two quarters of patient teaching for this project. We also thank every member contributed to the Directed Reading Program.

References

[1] Kurt Hornik. "Approximation capabilities of multilayer feedforward networks". In: *Neural Networks* 4.2 (1991), pp. 251–257.

[2] Moshe Leshno et al. "Multilayer feedforward networks with a nonpolynomial activation function can approximate any function". In: *Neural Networks* 6.6 (1993), pp. 861–867.

[3] Walter Rudin. *Principle of Mathematical Analysis*. McGraw-Hill, Inc., 1964, p. 159.

An Introduction to Cryptography

Lainey Watlington

University of California - Santa Barbara

The Beginnings of Cryptography

Cryptography is the study of methods of sending messages in a disguised form so that only the intended recipients can remove the disguise and read the message.

At the most basic level, a **cryptosystem** is the process of converting plaintext to a cyphertext using encryption and subsequently converting that cyphertext back to plaintext using decryption.

One of the earliest cryptosystems was created using **digraphs**, which map two characters in a message to a number. Let us consider the 27 letter alphabet which contains letters A-Z and a blank. Then, given any message, the following digraph can be used as an enciphering function where x and y are two characters which occur in succession in the message:

$$27x + y = C$$

The deciphering function is given by:

$$\begin{cases} C \mod 27 = x \\ C - x = y \end{cases}$$

Most early cryptosystems were based on a similar idea of using a rule, or a **key**, to shift the letters in a message to a different location. The idea was that only the person with the key would be able to decipher the message.

Breaking a Cryptosystem

Cryptosystems were developed in order to help protect sensitive information. In modern times, cryptography is widely used in the field of cybersecurity to protect people's

- passwords
- credit card information
- identity information
- other sensitive forms of data

In an increasingly digital world, cryptosystems have become extremely important in protecting this information.

Cryptanalysis is the science of "breaking" the code of cryptosystems. People do this in order to gain access to data that is not intended for them. This begs the question, "How does one break a cryptosystem?". To do so, one needs two types of information

1. The general nature, or the **structure** of the system
2. The specific choice of certain parameters connected with the given cryptosystem, like the shift parameter, also known as the **enciphering key**

An Example in Python

Let us extend the idea of a digraph to a cryptosystem which enciphers a message of length n from an alphabet of any size. Let N represent the size of the alphabet. Then, the enciphering function will be represented by

$$N^{n-1}x_1 + N^{n-2}x_2 + \cdots + Nx_{n-1} + x_{n-1} = C$$

The Python code for an enciphering transformation of this form is as follows:

```
def ngraph(base, message):
    message = str(message)
    length = len(message)
    sum = 0
    n = 1
    while (length-n) >= 0:
        sum = sum + (base**(n-1))*letternumber(message[length-n])
        n = n + 1
    return(sum)
```

The deciphering transformation will subtract $C \mod N$ from C n times and update C after each iteration. The Python code for a deciphering transformation of this form is:

```
def decipherngraph(base, number):
    n = 0
    string = ""
    if number == 0:
        string = "a"
    while number > 0:
        x = number%27
        number = int((number-x)/27)
        print(number)
        string = numberletter(x) + string

    return(string)
```

Primality and Factorization

Cryptosystems have evolved over time to prevent people from breaking them.

- The easier it is to guess the enciphering key of a cryptosystem, the easier it is to break the cryptosystem.
- So, methods of creating difficult to guess keys were developed

Public Key Cryptography: the enciphering and deciphering algorithms are publicly known, but the enciphering and deciphering keys are concealed. Gaining access to the keys allows you to break the system.

How do we create difficult to guess keys?

- **Factoring primes** is really difficult once we start dealing with very large numbers. So, if we multiply two large primes together, factoring them becomes almost impossible without having access to a key.
- The **discrete logarithm** problem is an idea based on the fact that if we know $y = b^x$, it is extremely difficult to solve for

$$x = \log_b y$$

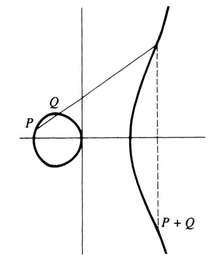
Fermat Factorization provides a way of "breaking" some public key cryptosystems. If two primes are close enough together, this algorithm allows one to efficiently calculate the two primes that have been multiplied together. This form of factorization is used to break RSA cryptosystems.

The Foundations of Modern Cryptography: Elliptic Curves

Elliptic Curve Cryptography

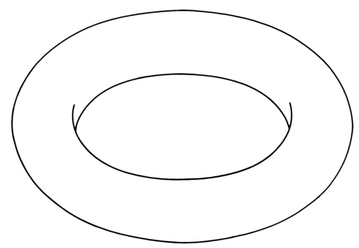
- An approach to public key cryptography which utilizes elliptic curves over finite fields to create keys.
- It is essentially impossible to find the discrete logarithm of a random element of an elliptic curve with respect to a publicly known base point.
- The larger the elliptic curve, the more secure the cryptosystem is since the discrete logarithm becomes more difficult to compute.

An Elliptic Curve Over the Real Numbers



- Elliptic curves over the reals form an abelian group. Thus, if we perform operations on two elements of the curve, we will end up with another element on the curve.

An Elliptic Curve Over the Complex Numbers



- Elliptic curves over the complex numbers form a torus.
- We can think of plotting elements of the curve over the integer lattice and then connecting all of the edges together.

Acknowledgements

Reference Material: "A Course in Number Theory and Cryptography" by Neal Koblitz

Thank you to the UCSB Directed Reading Program and to my mentor Katherine Merkl for making this project possible.

Isoperimetric Inequalities

Tyler Guo Mentor: Malik Tuerkoen

2022 Mathematics Directed Reading Program. University of California-Santa Barbara

Introduction

The classical isoperimetric problem is stated as follows: Among all closed curves in the plane of fixed perimeter, which curve (if any) maximizes the area of its enclosed region? This is equivalent to the problem: Among all closed curves in the plane enclosing a fixed area, which curve (if any) minimizes the perimeter? The problem can be extended to regions and surfaces in $\mathbb{R}^n$. In this poster, we show that a sphere has the smallest surface area with given volume by developing certain isoperimetric inequalities relating to the $\mathcal{L}^n$ measure of a sets and its perimeter.

Definitions

(1) For a function $u \in L^1(\Omega, \mathbb{R})$, we define

$$\mathrm{Var}(u, \Omega) := \sup\{\int_\Omega u \operatorname{div} \varphi \, dx : \varphi \in C_c^1(\Omega, \mathbb{R}^N), \|\varphi\|_\infty \le 1\}.$$

We say u has bounded variation in Ω if $\mathrm{Var}(u, \Omega) < \infty$.
Moreover, we let $BV(\Omega)$ denote the space of functions $u \in L^1(\Omega)$ which have bounded variation in Ω.
We also set

$$BV_{\mathrm{loc}}(\Omega) := \{u \in L^1_{\mathrm{loc}}(\Omega) : \mathrm{Var}(u, \Omega') < \infty \text{ for every } \Omega' \subset\subset \Omega\}.$$

(2) For a Lebesgue measurable subset E of $\mathbb{R}^N$. The perimeter of E in Ω is defined by

$$P(E, \Omega) := \mathrm{Var}(1_E, \Omega).$$

We say E has finite perimeter in Ω if $1_E \in BV(\Omega)$; E has locally finite perimeter in Ω if $1_E \in BV_{\mathrm{loc}}(\Omega)$.

Examples

(1) The distribution function

$$F_\mu : \mathbb{R} \to \mathbb{R}, \quad F_\mu(t) = \mu((-\infty, t])$$

of a probability measure μ on $\mathcal{B}(\mathbb{R})$ is a function of bounded variation in $\mathbb{R}$.

(2) Suppose $\Omega \subset \mathbb{R}^N$ is bounded and $u \in C^1(\overline{\Omega})$. Then $u \in BV(\Omega)$ and

$$\mathrm{Var}(u, \Omega) = \int_\Omega |\nabla u| \, dx.$$

(3) Let $Q_N := [0, 1]^N \subset \mathbb{R}^N$. Then Q has finite perimeter in $\mathbb{R}^N$ given by

$$P(Q_N) = 2N.$$

(4) Let $E \subset \mathbb{R}^N$ be a bounded open set with C^1-boundary. Then E has locally finite perimeter in Ω given by

$$P(E, \Omega) = \mathrm{vol}_{N-1}(\partial E \cap \Omega).$$

Gagliardo's Lemma

Let $N \ge 2$. For $x \in \mathbb{R}^N$, $j = 1, \dots, N$ let $\hat{x_j} := (x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_N)$. Moreover, let $f_1, \dots, f_N \in L^{N-1}(\mathbb{R}^{N-1})$ be given, and let $f : \mathbb{R}^N \to \mathbb{R}$ be defined by $f(x) = f_1(\hat{x_1}) \cdots f_N(\hat{x_N})$. Then

$$f \in L^1(\mathbb{R}^N) \text{ and } \|f\|_{L^1(\mathbb{R}^N)} \le \prod_{j=1}^N \|f_j\|_{L^{N-1}(\mathbb{R}^{N-1})}.$$

Non-optimal Isoperimetric Inequality

Let $N \ge 2$. Then we have

$$P(E) \ge 2\sqrt{N}|E|^{\frac{N-1}{N}}$$

for all measurable subsets $E \subset \mathbb{R}^N$ with $|E| < \infty$.

Proof of the theorem

The inequality holds trivially if $P(E) = \infty$. Suppose $P(E) < \infty$. We claim that for $u \in BV(\mathbb{R}^N)$, $N \ge 2$, we have

$$\|u\|_{L^{\frac{N}{N-1}}(\mathbb{R}^N)} \le \frac{1}{2\sqrt{N}} \mathrm{Var}(u, \mathbb{R}^N).$$

We obtain the inequality by applying this result to the function 1_E.
Proof of Claim: By standard approximation arguments, one can show that there exists a sequence (u_n) such that $u_n \in BV(\mathbb{R}^N) \cap C_c^1(\mathbb{R}^N)$ satisfying

$$\|u - u_n\|_1 \to 0, \quad \mathrm{Var}(u_n, \mathbb{R}^N) \to \mathrm{Var}(u, \mathbb{R}^N).$$

Hence it suffices to consider $u \in C_c^1(\mathbb{R}^N)$. Integration parallel to the j-th coordinate axis yields

$$|u(x)| \le \frac{1}{2} \int_{\mathbb{R}} |\partial_j u(x_1, \dots, x_{j-1}, t, x_{j+1}, \dots, x_N)| \, dt := v_j(\hat{x}_j)$$

for $x \in \mathbb{R}^N$, $j = 1, \dots, N$.

We then apply the Gagliardo's Lemma to $v_j^{\frac{1}{N-1}} \in L^{N-1}(\mathbb{R}^{N-1})$ and obtain

$$\int_{\mathbb{R}^N} |u(x)|^{\frac{N}{N-1}} dx \le \int_{\mathbb{R}^N} \prod_{j=1}^N v_j^{\frac{1}{N-1}}(\hat{x}_j) \, dx \le \prod_{j=1}^N \left\| v_j^{\frac{1}{N-1}} \right\|_{L^{N-1}(\mathbb{R}^{N-1})} = (\prod_{j=1}^N \|v_j\|_{L^1(\mathbb{R}^{N-1})})^{\frac{1}{N-1}}$$

$$\le (\frac{1}{N} \sum_{j=1}^N \|v_j\|_{L^1(\mathbb{R}^{N-1})})^{\frac{N}{N-1}} = (\frac{1}{2N} \int_{\mathbb{R}^N} [\sum_{j=1}^N |\partial_j u(x)|] \, dx)^{\frac{N}{N-1}} \le (\frac{1}{2\sqrt{N}} \int_{\mathbb{R}^N} |\nabla u| \, dx)^{\frac{N}{N-1}}.$$

□

Optimal Isoperimetric Inequality

For any measurable subset $E \subset \mathbb{R}^N$ with $|E| < \infty$ we have

$$P(E) \ge N\omega_N^{\frac{1}{N}} |E|^{\frac{N-1}{N}},$$

where ω_N denotes the volume of the unit ball in $\mathbb{R}^N$, and the equality occurs if and only if E is a ball.

Proof of the theorem

Suppose $P(E) < \infty$, we have $1_E \in BV(\mathbb{R}^N)$. Let $E^* = B_r(0)$, where r is chosen such that $|E| = |E^*|$. Then one can show there exists a sequence of sets (E_n) with $P(E_n) \le P(E)$ and $\|1_{E_n} - 1_{E^*}\|_1 \to 0$. By lower semicontinuity,

$$P(E^*) \le \liminf_{n\to\infty} P(E_n) = \liminf_{n\to\infty} \mathrm{Var}(1_{E_n}, \mathbb{R}^N) \le \mathrm{Var}(1_E, \mathbb{R}^N) = P(E).$$

Moreover,

$$P(E^*) = \mathrm{vol}_{N-1}(\partial E^*) = N\omega_N r^{N-1} = N\omega_N^{\frac{1}{N}} |E^*|^{\frac{N-1}{N}} = N\omega_N^{\frac{1}{N}} |E|^{\frac{N-1}{N}}.$$

□

Coarea Formula for BV functions

Let $f \in BV(\Omega)$ be a nonnegative function, and put

$$E_t := \{x \in \Omega : f(x) > t\} \quad \text{for } t \ge 0.$$

Then

$$\mathrm{Var}(f, \Omega) = \int_0^\infty P(E_t, \Omega) \, dt.$$

Optimal Functional Isoperimetric Inequality

For $f \in BV(\mathbb{R}^N)$ we have

$$\mathrm{Var}(f, \mathbb{R}^N) \ge N\omega_N^{\frac{1}{N}} \|f\|_{L^{\frac{N}{N-1}}}.$$

Proof of the theorem

Since

$$\|f\|_{L^{\frac{N}{N-1}}} \le \|f^+\|_{L^{\frac{N}{N-1}}} + \|f^-\|_{L^{\frac{N}{N-1}}},$$

and one can show that for $f \in BV(\mathbb{R}^N)$,

$$\mathrm{Var}(f, \mathbb{R}^N) = \mathrm{Var}(f^+, \mathbb{R}) + \mathrm{Var}(f^-, \mathbb{R}).$$

Hence it suffice to consider the case where f is nonnegative. In this case, the Coarea formula and the isoperimetric inequality yields

$$\mathrm{Var}(f, \mathbb{R}^N) = \int_0^\infty P(E_t) \, dt \ge N\omega_N^{\frac{1}{N}} \int_0^\infty |E_t|^{\frac{N-1}{N}} \, dt.$$

We now define

$$\chi : [0, \infty) \to \mathbb{R}, \quad \chi(t) = \|\min\{f, t\}\|_{L^{\frac{N}{N-1}}}.$$

Then χ is continuous, nondecreasing and hence a.e. differentiable. Moreover, for $t, h > 0$, we have

$$0 \le \chi(t + h) - \chi(t) \le \|\min\{f, t + h\} - \min\{f, t\}\|_{L^{\frac{N}{N-1}}} \le \|1_{E_t} h\|_{L^{\frac{N}{N-1}}} = h|E_t|^{\frac{N-1}{N}},$$

which implies that χ is locally Lipschitz continuous on $(0, \infty)$ with $\chi'(t) \le |E_t|^{\frac{N-1}{N}}$ for a.e. $t > 0$. Hence χ satisfies the assumptions of the Fundamental theorem of calculus . Since $0 = \chi(0) = \lim_{t\to 0^+} \chi(t)$, it follows that

$$\|f\|_{L^{\frac{N}{N-1}}} = \lim_{b\to\infty} [\chi(b) - \chi(\frac{1}{b})] = \lim_{b\to\infty} \int_{\frac{1}{b}}^b \chi'(t) \, dt \le \int_0^\infty |E_t|^{\frac{N-1}{N}} \, dt.$$

□

Acknowledgements

I would like to thank my mentor Malik Tuerkoen for his guidance and insight. I would also like to thank the organizers of the UCSB Directed Reading Program for this wonderful opportunity.

References

[1] Lawrence Craig Evans and Ronald F Gariepy. *Measure Theory and Fine Properties of Functions, Revised Edition*. CRC Press LLC, Oakville, 2015.