A topological proof of the insolvability of the quintic

Yunhai Xiang

June 23, 2022

Short Attention Span Math Seminars, Pure Math Club, University of Waterloo

Table of contents

Introduction

In middle school, we learned the solutions of the quadratic equation

$$ax^2 + bx + c = 0$$

where $a, b, c \in \mathbb{C}$ with $a \neq 0$, are given by the quadratic formula

$$\lambda_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

which only used field operations $+, -, \times, /$ and the square root $\sqrt{\cdot}$.

For the cubic and quartic equations

$$ax^3 + bx^2 + cx + d = 0$$
$$ax^4 + bx^3 + cx^2 + dx + e = 0$$

where $a, b, c, d, e \in \mathbb{C}$ and $a \neq 0$, there are cubic and quartic formulas using only field operations $+, -, \times, /$ and radicals $\sqrt{\cdot}$, $\sqrt[3]{\cdot}$, and $\sqrt[4]{\cdot}$, albeit much more complicated than the quadratic formula.

The solution to $ax^3 + bx^2 + cx + d = 0$ is given by

$$x = \sqrt[3]{\left(\frac{-b^3}{27a^3} + \frac{bc}{6a^2} - \frac{d}{2a}\right) + \sqrt{\left(\frac{-b^3}{27a^3} + \frac{bc}{6a^2} - \frac{d}{2a}\right)^2 + \left(\frac{c}{3a} - \frac{b^2}{9a^2}\right)^3}}$$
$$+ \sqrt[3]{\left(\frac{-b^3}{27a^3} + \frac{bc}{6a^2} - \frac{d}{2a}\right) - \sqrt{\left(\frac{-b^3}{27a^3} + \frac{bc}{6a^2} - \frac{d}{2a}\right)^2 + \left(\frac{c}{3a} - \frac{b^2}{9a^2}\right)^3}} - \frac{b}{3a}.$$

The solution to $ax^4 + bx^3 + cx^2 + dx + e = 0$ is an even longer formula.

Theorem 1.1 (Fundamental theorem of algebra)

For $n > 0$*, the equation*

$$z^n + a_{n-1}z^{n-1} + a_{n-2}z^{n-2} + \cdots + a_1 z + a_0 = 0$$

where all $a_i \in \mathbb{C}$*, has exactly* n *solutions in* $\mathbb{C}$ *counting multiplicity.*

Question: For $n \geq 5$, is there a general formula for

$$z^n + a_{n-1}z^{n-1} + a_{n-2}z^{n-2} + \cdots + a_1 z + a_0 = 0$$

on $a_i \in \mathbb{C}$ using only (a finite number of) $+, -, \times, /$ and $\sqrt{\cdot}, \sqrt[3]{\cdot}, \ldots$?

Theorem 1.2 (Abel–Ruffini)

For $n \geq 5$, there is no general formula for

$$z^n + a_{n-1}z^{n-1} + a_{n-2}z^{n-2} + \cdots + a_1 z + a_0 = 0$$

on $a_i \in \mathbb{C}$ using only (a finite number of) $+, -, \times, /$ *and* $\sqrt{\cdot}, \sqrt[3]{\cdot}, \ldots$

The typical proof of this theorem uses heavy machinery from Galois theory, but there is in fact a far more elementary but much less well known proof due to V.I. Arnold, using nothing but basic knowledge of complex numbers and topology.

Reasons I prefer Arnold's proof over the classical Galois theory proof.

(i) It is more elementary,

(ii) It is more visual,

(iii) It is a stronger result in some sense,

(iv) It helps you to **really** understand the classical Galois theory proof.

Toy examples

Question: Can we distinguish between i and $-i$ canonically?

Remember there are two square roots of -1, either can be defined as i.

Way too often, we use notation $\mathbb{C}$ to mean "$\mathbb{C}$ with a choice of i". This is an abuse of notation that goes unnoticed due to its subtlety.

There is no "nice" or "canonical" order of roots for $z^2 = -1$ or any algebraic equation $z^n + a_{n-1}z^{n-1} + a_{n-2}z^{n-2} + \cdots + a_1 z + a_0 = 0$.

We pick an order for the roots of $z^2 = -1$, say $\lambda_1 = i$ and $\lambda_2 = -i$.

Consider the roots of $z^2 = e^{i\theta}$. Observe how the they change as θ goes from π to 3π continuously. Note that $e^{i\pi} = e^{i(3\pi)} = -1$.

We see that $e^{i\theta}$ moves along a loop based at -1. A **loop** based at $p \in \mathbb{C}$ is just a continuous function $\gamma : [0, 1] \to \mathbb{C}$ s.t. $\gamma(0) = \gamma(1) = p$.

What happened? The roots swap places!

In general, for $z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0 = 0$, pick an order of its roots $\lambda_1, \ldots, \lambda_n$, then a permutation $\sigma \in S_n$ of λ_i is induced by moving each coefficient a_i along some loop based at a_i.

In this example of $z^2 = -1$, we moved a_0 along the loop $-e^{i\theta}$ with θ from π to 3π, inducing (1 2) on the roots $\lambda_1 = i$ and $\lambda_2 = -i$.

This proves there is no quadratic formula only using $+, -, \times, /$. (Why?)

In the previous example, the loop $e^{i\theta}$ for θ from π to 3π is no longer a loop under $\sqrt{\cdot}$, since it induces a nontrivial permutation (1 2).

Question: What are the loops that **remain loops under radicals**, i.e., the permutation they induce on $\sqrt[k]{\cdot}$ is the identity for each k?

Let γ a loop, then let γ^{-1} denote its **inverse loop**, i.e. the same loop but going backwards. Let τ be a loop based at the same point as γ, then let $\gamma \cdot \tau$ denote their **concatenated loop**, i.e. the loop that goes along γ and then goes along τ.

Question: Suppose that γ does not pass through 0, do you see why the loop $\gamma \cdot \gamma^{-1}$ remains a loop under radicals?

Let γ be a loop not passing through 0. Its **winding index** about 0 is an integer: the number of times it wraps around the anticlockwise direction about 0 subtracted by that of clockwise ones.

For experts: the winding index of γ about 0 is given by

$$\mathrm{Ind}_\gamma(0) = \frac{1}{2\pi i} \oint_\gamma \frac{dz}{z}$$

It's straightforward to see that γ remains a loop under radicals if it has winding index 0 about 0, and $\gamma \cdot \gamma^{-1}$ has winding index 0 about 0.

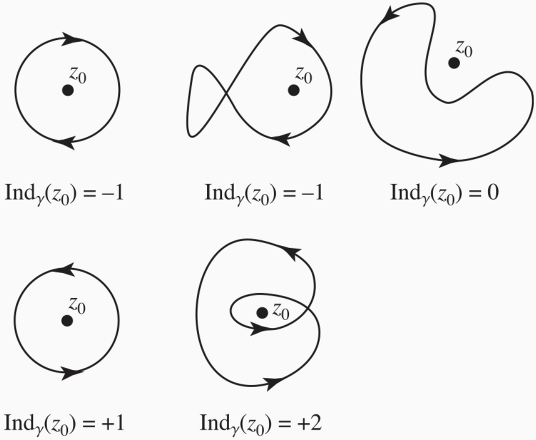

Figure 1: Winding index

Theorem 2.3 (Vieta's formulas)

If $z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0 = 0$ *has solutions* $\lambda_1, \ldots, \lambda_n$, *then*

$$a_{n-k} = (-1)^k \sum_{1 \le i_1 < \cdots < i_k \le n} \left(\prod_{j=1}^{k} \lambda_{i_j} \right)$$

for $k = 1, \ldots, n$.

For any permutation $\sigma \in S_n$ of roots $\lambda_1, \ldots, \lambda_n$, we can move the roots continuously so that they end up at the positions after the permutation. By Vieta's formulas, this produces loops of coefficients that induces σ, and we can always make it so that the loops avoid passing through 0.

Using these ideas, we show that any cubic formula must use **nested radicals**, i.e. radicals inside radicals, such as $\sqrt[3]{a_0 + \sqrt{a_0^2 + a_1}}$.

Question: Assume that $a_0, a_1, a_2 \neq 0$ and the equation has no repeated solutions. If we could find three loops based at a_0, a_1, a_2 not passing through 0 that remain loops under radicals but induces a nontrivial permutation on the roots $\lambda_1, \lambda_2, \lambda_3$, what would that imply?

Pick loops $\gamma_0, \gamma_1, \gamma_2$ and loops τ_0, τ_1, τ_2 based at a_0, a_1, a_2 respectively, such that none of γ_i or τ_i pass through 0 and γ_i and τ_i induce (1 2 3) and (1 2) respectively on the roots. Define the **commutator loop**

$$[\gamma_i, \tau_i] = \gamma_i \cdot \tau_i \cdot \gamma_i^{-1} \cdot \tau_i^{-1}$$

for each i. The commutators have winding index 0 about 0, so they remain loops under radicals, but they induce a nontrivial permutation

$$[(1\ 2), (1\ 2\ 3)] = (1\ 2)^{-1}(1\ 2\ 3)^{-1}(1\ 2)(1\ 2\ 3) = (2\ 1\ 3)$$

which is the **commutator permutation** of (1 2) and (1 2 3).

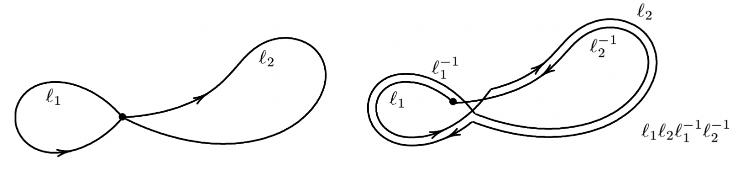

Figure 2: Commutator of loops ℓ_1 and ℓ_2

The proof

Assume for sake of contradiction that there is a quintic formula.

Question: For some values $a_i \neq 0$ for each i such that the equation has no repeated roots, can we find loops based at each a_i that remain loops under the quintic formula, but induce a nontrivial permutation on the roots $\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5$?

One would naturally think about taking commutators, commutators of commutators, commutators of commutators of commutators, and so on.

Fact 3.4

degree	*# perm.*	*# comm.*	*# comm. of comm.*	*# comm. of comm. of comm.*
$n = 1$	*1*	*1*	*1*	*1*
$n = 2$	*2*	*1*	*1*	*1*
$n = 3$	*6*	*3*	*1*	*1*
$n = 4$	*24*	*12*	*4*	*1*
$n = 5$	*120*	*60*	*60*	*60*

In degree 5*, there are* 60 *permutations that are commutators of other permutations, and the commutators of these elements are themselves.*

Finally, for the pièce de résistance, if the quintic formula has n levels of nested radicals, we can always pick a nontrivial permutation expressed by n levels of commutators by **Fact 3.4**. Using these permutations, we can induce loops which remain loops under the formula (Why?). This means that the quintic formula must have an infinite level of nested radicals,

$$\sqrt[k_1]{f_1 + \sqrt[k_2]{f_2 + \sqrt[k_3]{f_3 + \sqrt[k_4]{\cdots}}}}$$

which is a contradiction!

As an exercise, show that there is no general formula for any algebraic equation of degree ≥ 5 using only field operations, radicals, or any continuous complex function (e.g. $\sin z, \cos z, e^z$).

Rigorous formulation

Technical obstacles in the proof

(i) How do we rigorously define a general formula?

(ii) How do we rigorously define induced permutation on roots?

A **general formula** for $E \subseteq \mathbb{C}^n$ is an ordered list of rational functions (over $\mathbb{Q}$) $f_1, \ldots, f_m$ of $n, n+1, \ldots, n+m-1$ variables resp. and an ordered list of $k_1, \ldots, k_m \in \mathbb{Z}^+$ s.t. for all $(a_0, \ldots, a_{n-1}) \in E$, if λ is a root of $a^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0 = 0$, then there exists some $z_1, \ldots, z_m \in \mathbb{C}$ such that $z_m = \lambda$ and

$$\begin{aligned} z_1^{k_1} &= f_1(a_0, \ldots, a_{n-1}) \\ z_2^{k_2} &= f_2(a_0, \ldots, a_{n-1}, z_1) \\ z_3^{k_3} &= f_3(a_0, \ldots, a_{n-1}, z_1, z_2) \\ &\vdots \\ z_m^{k_m} &= f_m(a_0, \ldots, a_{n-1}, z_1, \ldots, z_{m-1}) \end{aligned}$$

We say E is **solvable by radicals** if it has a general formula.

An **algebraic function** $y = f(x)$ is defined by

$$F(x, y) = y^n + g_{n-1}(x)y^{n-1} + \cdots + g_0(x) = 0$$

where g_i are polynomials. Pick $x = a$ such that $F(a, y)$ has distinct roots $y = z_1, \ldots, z_n$. By implicit function theorem, exists open nbhd U_a of a s.t. $F(x, y) = 0$ has distinct roots for $x \in U_a$, which defines functions $f_{a,1}(x), \ldots, f_{a,n}(x)$ on U_a. WLOG, we can choose U_a as the common convergence disc of the Taylor expansions of $f_{a,i}(x)$. A pair $(f_{a,i}, U_a)$ is called a **chart** or an **analytic element**.

Given (f_a, U_a) where f_a defined on U_a has convergent Taylor series at a. We prolong (f_a, U_a) along a path γ (not passing through any singular points), covered by finite number of U_{a_i} where $a_i \in \gamma$ which agree on intersections, to obtain a chart (f_b, U_b) at the end. This is called the **analytic continuation** of (f_a, U_a) along γ.

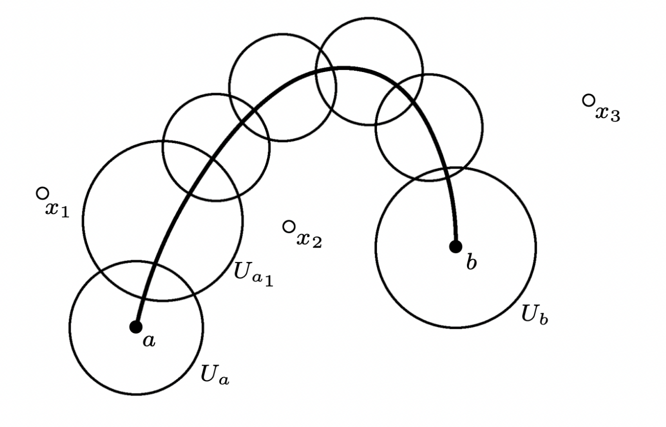

Figure 3: Analytic continuation along a path

Theorem 4.5 (Monodromy theorem)

If the paths γ_1, γ_2 (starting from a point a and ending at a point b) are homotopic (we can continuously deform them to each other), then the analytic continuation of (f_a, U_a) along γ_1 is the same the the analytic continuation of (f_a, U_a) along γ_2.

This guarantees the uniqueness of the analytic continuation.

The union of all charts and all analytic continuations of them along all possible paths is called the the **Riemann surface** of f, which is a natural covering space of the complex plane minus the singular points.

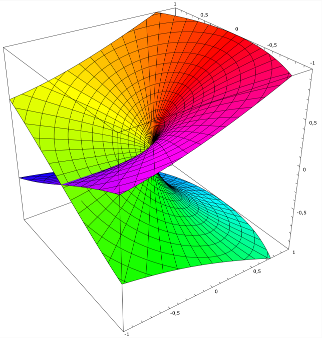

Figure 4: Riemann surface of $\sqrt{x}$

If γ is a loop in $\mathbb{C}$ (not passing through singular points), then the analytic continuation of $(f_{a,i}, U_a)$ along γ leads to some $(f_{a,j}, U_a)$. The permutation that arises this way is the induced permutation.

These permutations form the **monodormy group** of f, denoted $\mathrm{Mon}(f)$, which can be identified as the image of the natural map

$$\pi_1(\mathbb{C} \setminus \{\textit{singular pts}\}, a) \to S_{\{z_1,\ldots,z_n\}}$$

known as the **monodromy representation**, where $\pi_1(X, a)$ is the group of loops in X at a up to homotopy, called the **fundamental group**.

Definition 4.6 (Commutator subgroup)

The **commutator subgroup** of a group G is

$$[G, G] = \langle [g, h] : g, h \in G \rangle$$

with the operation inherited from G.

Definition 4.7 (Solvable groups)

A group G is **solvable** if the **derived series** of subgroups

$$G = G^{(0)} \trianglerighteq G^{(1)} \trianglerighteq G^{(2)} \trianglerighteq \cdots$$

where $G^{(k)} = [G^{(k-1)}, G^{(k-1)}]$, terminates in the trivial group.

The monodromy group of a typical algebraic function is S_n where n is the degree, which is not solvable for $n \geq 5$, since the derived series

$$S_5 \trianglerighteq A_5 \trianglerighteq A_5 \trianglerighteq A_5 \trianglerighteq A_5 \cdots$$

contradicting the fact that a monodromy group of an algebraic function expressed in radicals is solvable.

In fact $\mathrm{Mon}(f)$ is isomorphic to a certain Galois group.

Further topics

Galois theory and topology are related to each other in numerous ways.

(i) Each Galois group is naturally a topological group with the **Krull topology**, which is discrete for finite Galois groups, and given by

$$\mathrm{Gal}(K/k) = \varprojlim_{\substack{K/L/k \\ L/k \text{ finite Galois}}} \mathrm{Gal}(L/k)$$

for infinite ones. This topology is Hausdorff and compact.

(ii) The **fundamental theorem of covering spaces** gives an order preserving correspondence between subgroups of fundamental groups (up to conjugacy) and covering spaces, which is uncanily similar to the **fundamental theorem of Galois theory**.

Some further connections between Galois theory and topology:

(i) Topological Galois theory: Arnold's idea was explored further by A. Khovanskii and other mathematicians on problems regarding solvability of differential equations, integrals, etc.

(ii) Grothendieck's Galois theory: Inspired by the similarity of Galois groups and fundamental groups, Grothendieck developed an analogue of them called the étale fundamental group of schemes.

(iii) Grothendieck–Teichmüller theory: The Grothendieck–Teichmüller conjecture states that the Grothendieck–Teichmüller group is isomorphic to $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. Grothendieck wanted "a purely topological characterization of the Galois group, a purely arithmetical object." This relates to ideas like Belyi pairs and dessin d'enfant.

Thank you for listening!

References

i. L. Goldmakher (2013). *Arnold's elementary proof of the insolvability of the quintic*

ii. H. Zoladek (2000). *The topological proof of Abel-Ruffini theorem*

iii. F. Akalin (2016). *Why is the quintic unsolvable?*

iv. V. Kalicki, J. Morales, R. Ostrander (2019). *Visulization of Abel's impossibility theorem*

v. V.B. Alekseev (2004). *Abel's Theorem in problems and solutions*

vi. D. Miller (2014). *A brief tour of Grothendieck-Teichmüller theory*

vii. A. Khovanskii (2014). *Topological Galois theory*

viii. F. Neumann, S. Schroll (2018). *Galois covers, Grothendieck-Teichmüller theory and Dessins d'Enfants*

ix. D. Fuchs, S. Tabachnikov (2007). *Mathematical Omnibus: Thirty lectures on classical mathematics*