Fundamental Theorem of Algebra
Complex Analytic and Topological Proofs

Zhang Ruiyang
ruiyang.zhang.20@ucl.ac.uk

Contents

1 Introduction

This notes is prepared as the supplementary material for a talk I gave at the Undergraduate Mathematics Colloquium at UCL. It assumes little knowledge from the readers, and it will be good to have some basic understanding of Analysis and Topology. Additionally, knowledge of Complex Analysis and Algebraic Topology will be very helpful, and may make much of the content of this notes trivial. Hopefully, it could serve as good expository material for novices of Mathematics. Another aim of this notes is to explore how different branches of Mathematics can come together to show the same result, illustrating the elegance of Mathematics.

–

The **Fundamental Theorem of Algebra**, or FTA, is a theorem that shows the existence of root of a non-constant polynomial. To be more specific, it is about the existence of complex root of a non-constant polynomial with complex coefficients. Moreover, the number of roots is the same as the degree of the polynomial. The first proper proof was provided by Gauss, and as time progresses, other Mathematicians provided other proofs of the same theorem. In this notes, I will be presenting two (not original) proofs of FTA, the first one is using Complex Analysis, and the second involves Algebraic Topology.

Recall that a complex polynomial

$$P(z) = a_n z^n + \cdots + a_1 z + a_0$$

where the coefficients $a_0, \cdots, a_n$ are complex numbers, and $z \in \mathbb{C}$. The highest power of z with non-zero coefficient is known as the degree. A root, or a zero, of the polynomial is a value $z_o \in \mathbb{C}$ such that $P(z_0) = 0$. The existence of root can be extended to show that the number of roots is the same as the degree. For a polynomial $P(z)$ with degree n, if z_0 is a root of it, we can factorise it and get $P(z) = (z - z_0)Q(z)$ for a polynomial $Q(z)$ of degree $n - 1$. We can continue this process of factorising and eventually get n roots.

Now, why do we have different kinds of proof for this theorem of Algebra? One reason is that polynomial represents different things in different branches of Mathematics. In this notes, we presented the Complex Analytic proof and the Algebraic Topological proof. In Complex Analysis, a polynomial can be treated as a function that maps from $\mathbb{C}$ to $\mathbb{C}$, and therefore we can use properties of functions on it. In Algebraic Topology, we will plot a polynomial and treat it as a shape. In Topology, a big area of interest is on the topological invariants, the properties that will remain unchanged after homeomorphisms. Homeomorphism, roughly speaking, is the action of continuous deforming. For example, the classic example is how in the topological viewpoint, a mug and a donut are the same, since they can be continuously deformed into each other.

For those who want to discover more about FTA, for example the alternative proofs, you may refer to the book 'The Fundamental Theorem of Algebra' by Fine and Rosenberger [2], or this post on MathOverflow.

2 Complex Analysis Proof

In this Chapter, we will be presenting the proof of FTA using Complex Analysis. As we assume no prior knowledge beyond Mathematical Analysis and aim to make it as self-contained as possible, we will start from the very beginning of Complex Analysis, and slowly move our way to the proof of FTA. As the goal is to prove FTA, we will keep anything less relevant to the proof to a minimum. Here, we will be looking at things such as Cauchy Integral Theorem, Cauchy Integral Theorem for derivatives, Cauchy Inequalities, Liouville Theorem and so on. For readers who are interested to have a more complete understanding of Undergraduate Complex Analysis, I would recommend you to read a proper textbook, such as Serge Lang [4], or Elias Stein [6] (both under the title 'Complex Analysis').

2.1 Complex Analysis Basics

From Mathematical Analysis, we learned about how to do calculus with real-valued function $f : \mathbb{R} \to \mathbb{R}$. Now, for Complex Analysis, we will be working with complex-valued functions $f : \mathbb{C} \to \mathbb{C}$. This is going to be different from the real-valued cases, but it is not that different. Notice that we can write $z \in \mathbb{C}$ as $x + iy$ for $x, y \in \mathbb{R}$, we can somewhat think about $\mathbb{C}$ as $\mathbb{R}^2$. This may help us with some intuitions.

We will start with some of the basics of complex-valued functions. Recall that $f : \mathbb{R} \to \mathbb{R}$ will be **differentiable** at a point x if

$$\lim_{h\to 0} \frac{f(x+h) - f(x)}{h}$$

exists. Similarly, we will say $f : \mathbb{C} \to \mathbb{C}$ is differentiable, or more specifically **complex differentiable**, at a point z if both

$$\lim_{h\to 0} \frac{f(z+h) - f(z)}{h} \quad \text{and} \quad \lim_{h\to 0} \frac{f(z+ih) - f(z)}{ih}$$

exist and are equal for $h \in \mathbb{R}$. This does not seem to be too accessible. Let us see if we can simplify it a bit.

As we know that we can write $z \in \mathbb{C}$ as $x + iy$ for $x, y \in \mathbb{R}$, we can tweak things a bit, and rewrite $f(z) \in \mathbb{C}$ as $u(x, y) + iv(x, y)$ where $u(x, y), v(x, y) \in \mathbb{R}^2$ and $x, y \in \mathbb{R}$. If we assume $u(x, y)$ and $v(x, y)$ are real differentiable at that point z, we will have

$$\begin{aligned}
\lim_{h\to 0} \frac{f(z+h) - f(z)}{h} &= \lim_{h\to 0} \frac{u(x+h,y) + iv(x+h,y) - u(x,y) - iv(x,y)}{h} \\
&= \lim_{h\to 0} \frac{u(x+h,y) - u(x,y)}{h} + \lim_{h\to 0} \frac{iv(x+h,y) - iv(x,y)}{h} \\
&= \frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x}
\end{aligned}$$

and

$$\begin{aligned}
\lim_{h\to 0}\frac{f(z+ih)-f(z)}{ih} &= \lim_{h\to 0}\frac{u(x,y+h)+iv(x,y+h)-u(x,y)-iv(x,y)}{ih} \\
&= \lim_{h\to 0}\frac{u(x,y+h)-u(x,y)}{ih} + \lim_{h\to 0}\frac{iv(x,y+h)-iv(x,y)}{ih} \\
&= -i\frac{\partial u}{\partial y}+\frac{\partial v}{\partial y}.
\end{aligned}$$

For these two to be equal, we need to have $\frac{\partial u}{\partial x}+i\frac{\partial v}{\partial x}=-i\frac{\partial u}{\partial y}+\frac{\partial v}{\partial y}$, or simply

$$\begin{cases} \frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \\ \frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}. \end{cases}$$

This is known as the **Cauchy-Riemann Equations**. The Cauchy-Riemann equations provide us with an alternative way to check for complex-differentiability. If we can write $f(z)$ as $u(x,y)+iv(x,y)$ for u,v differentiable functions on $\mathbb{R}^2$ at $z=x_0+iy_0$, as long as u,v satisfy the Cauchy-Riemann equations for (u_0,v_0), it is complex differentiable. If this is true for every point of the domain, the function is **holomorphic**. If a holomorphic function has domain of the entire $\mathbb{C}$, it is **entire**

—

We have quite a bit of differentiation. Now let us move on to integration. Just like what we did for differentiation, let us recall how we did integration for real-valued function $f:\mathbb{R}\to\mathbb{R}$.

For (Riemann) integrable functions $f:[a,b]\to\mathbb{R}$, since the domain is a subset of $\mathbb{R}$, we can visualise the domain as a portion from the real line. For complex-valued function $f:\mathbb{C}\to\mathbb{C}$, the domain will be a subset of $\mathbb{C}$, and instead of for example a segment of the real line, we will have a more complicated domain like a path or a loop. We have not properly defined what a path or a loop may be, but by common sense we could have a rough picture of a smooth curve connecting two points for a path, and a circle-ish curve starting and ending at the same point for a loop, all happening in the complex plane. Now, let us try to formally define them.

Definition 2.1. A **smooth path** in the complex plane is a map $\gamma:[A,B]\to\mathbb{C}$ that is continuously differentiable on $[A,B]$.

Sometimes a smooth path is too much to ask for, and we will have a slightly weaker form that will allow us to do integration just fine.

Definition 2.2. A continuous map $\gamma:[A,B]\to\mathbb{C}$, where $[A,B]\subset\mathbb{R}$ is a closed interval, is called a **piecewise smooth path** in C if there exists a finite partition $A=A_0<A_1<\cdots<A_n=B$ such that the restriction of γ to each segment $[A_j,A_{j+1}]$ is a smooth path.

In the following, we will use 'path' when we really mean 'piecewise smooth path'. Now, we are ready to define integration over path.

Definition 2.3. Let $\gamma : [A, B] \to \mathbb{C}$ be a piecewise smooth path and f be a continuous complex-valued function defined on the set $\gamma([A, B]) \subset \mathbb{C}$. Then, the **integral** of f over γ is the number

$$\int_\gamma f(z)dz = \int_A^B f(\gamma(t))\gamma'(t)dt.$$

Remark. The equality is simply a change of variable using the fact that $z = \gamma(t)$.

Notice that we are using a parametrisation of the path to define the integral. The choice of parametrisation is mostly independent to the integral.

Proposition 2.4. Assume that the conditions of Definition 2.3 are satisfied and $\varphi : [A_1, B_1] \to [A, B]$ is a bijective differentiable function whose derivative is everywhere positive. Then,

$$\int_{\gamma\circ\varphi} f(z)dz = \int_\gamma f(z)dz.$$

If the derivative is everywhere negative, then

$$\int_{\gamma\circ\varphi} f(z)dz = -\int_\gamma f(z)dz.$$

Proof. If φ is increasing with $\varphi(A_1) = A$ and $\varphi(B_1) = B$, then we will have

$$\begin{aligned}\int_\gamma f(z)dz &= \int_A^B f(\gamma(t))\gamma'(t)dt \\ &= \int_{A_1}^{B_1} f(\gamma(\varphi(u)))\gamma'(\varphi(u))\phi'(u)dt \\ &= \int_{A_1}^{B_1} f(\gamma\circ\varphi(u))(\gamma\circ\varphi)'(u)du \\ &= \int_{\gamma\circ\varphi} f(z)dz.\end{aligned}$$

The decreasing case can be shown in a similar manner. □

We will now introduce another definition.

Definition 2.5. A **piecewise smooth curve** is a subset in $\mathbb{C}$ of the form $\gamma([A, B])$ where $\gamma : [A, B] \to \mathbb{C}$ is an injective piecewise smooth path and $\gamma'(t) \neq 0 \ \forall \ t \in [A, B]$ except for possibly finitely many points.

A curve will be called closed if we have $\gamma(A) = \gamma(B)$. The direction of a closed curve will matter here. We will call the counterclockwise direction as positive, and clockwise as negative. Readers can refer to the way sign of the angles is determined in the polar coordinates.

We will now try to do an example of path integral.

Example 2.6. If γ is a positively oriented circle of radius $r > 0$ centred at $a \in \mathbb{C}$, then

$$\int_\gamma \frac{dz}{z-a} = 2\pi i.$$

A way to parametrise the circle is $\gamma(t) = a + re^{it}$ where $t \in [0, 2\pi]$. Then, we have $z = a + re^{it}$ and $dz = ire^{it}dt$. So,

$$\begin{aligned}\int_\gamma \frac{dz}{z-a} &= \int_0^{2\pi} \frac{ire^{it}dt}{re^{it}} \\ &= \int_0^{2\pi} idt \\ &= 2\pi i.\end{aligned}$$

This example will appear again in the following sections.

Next, we will study some of the properties of integration over curves. We first make a definition of length. For path $\gamma : [A, B] \to \mathbb{C}$, the **length** of it is defined to be $\text{length}(\gamma) = \int_A^B |\gamma'(t)|dt$ for a particular parametrisation of the curve. Then, we have the following inequality.

Proposition 2.7. For a continuous function defined on a curve $\gamma([A, B])$ with $\gamma : [A, B] \to \mathbb{C}$, we have

$$\Big|\int_\gamma f(z)dz\Big| \leq \sup_{z\in\gamma([A,B])} |f(z)| \cdot \text{length}(\gamma).$$

Proof. Let $M = \sup_{z\in\gamma([A,B])} |f(z)|$. We have

$$\begin{aligned}\Big|\int_\gamma f(z)dz\Big| &= \Big|\int_A^B f(\gamma(t))\gamma'(t)dt\Big| \\ &\leq \int_A^B \Big|f(\gamma(t))\gamma'(t)\Big|dt \\ &\leq \int_A^B M\Big|\gamma'(t)\Big|dt \\ &\leq M \cdot \int_A^B \Big|\gamma'(t)\Big|dt = M \cdot \text{length}(\gamma)\end{aligned}$$

□

–

Recall that we have a Fundamental Theorem of Calculus for real-valued functions. We will have that for complex-valued ones.

Definition 2.8. let $U \subset \mathbb{C}$ be an open set and $f : U \to \mathbb{C}$ be a continuous function. We say that a function $F : U \to \mathbb{C}$ is an **antiderivative** (or **primitive**) of f if F is holomorphic on U and $F'(z) = f(z)$ for every $z \in U$.

Theorem 2.9. (Fundamental Theorem of Calculus for Complex Functions) Let $U \subset \mathbb{C}$ be an open set and $f : U \to \mathbb{C}$ be a continuous function that has an antiderivative F in U. If γ is a path in U joining points $p \in U$ and $q \in U$, then

$$\int_\gamma f(z)dz = F(q) - F(p).$$

Proof. Let $\gamma : [A, B] \to U$ be a path such that $\gamma(A) = p$ and $\gamma(B) = q$. Then,

$$\begin{aligned} \int_\gamma f(z)dz &= \int_\gamma F'(z)dz \\ &= \int_A^B F'(\gamma(t))\gamma'(t)dt \\ &= \int_A^B \frac{d}{dt}F(\gamma(t))dt \\ &= F(\gamma(B)) - F(\gamma(A)) \\ &= F(q) - F(p). \end{aligned}$$

□

Corollary 2.10. If a continuous function f defined on an open set $U \subset \mathbb{C}$ has an antiderivative on this set, then the integral of f over every closed path in U vanishes.

This is obvious to see, as $F(q) = F(p)$ for a closed path. Also, if you ever wonder why the integral over the circle in the example above does not vanish, check the conditions of this Corollary and see if they are all satisfied, especially when $z = a$.

2.2 Cauchy's Theorem

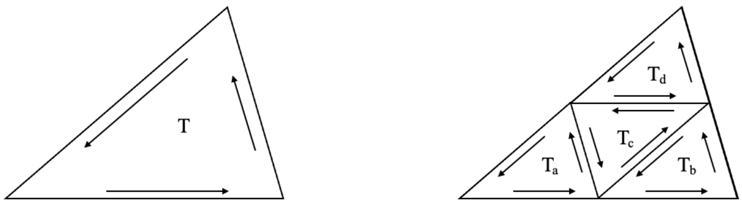
From the end of the previous section, we saw that a function defined on an open set with an antiderivative will have vanishing integral over any closed path. Having an antiderivative everywhere on an open set is quite a strong condition, and we do not always have that. The generalisation that remove the condition on existence of antiderivative is the Cauchy's Theorem. However, due to the expository nature of this text, we will not be showing and proving the full form of this theorem, and instead we will only be looking at the minimum form to proceed with our proof of the Fundamental Theorem of Algebra. For those readers who would like to know the complete version, do learn some Topology (especially homotopy) and refer to a Complex Analysis textbook. Let us start with the following theorem.

Theorem 2.11. (Goursat's Theorem) If U is an open set in $\mathbb{C}$, and $T \subset U$ is a triangle whose interior is also contained in U, then

$$\int_{\partial T} f(z)dz = 0$$

wherever f is holomorphic in U.

Proof. First assume that $T \subset \mathbb{C}$ is an arbitrary triangle and $f : T \to \mathbb{C}$ is an arbitrary continuous function. The midsegments divide T into four congruent triangles T_a, T_b, T_c, and T_d as shown below.



Then, we have

$$\int_{\partial T} f(z)dz = \int_{\partial T_a} f(z)dz + \int_{\partial T_b} f(z)dz + \int_{\partial T_c} f(z)dz + \int_{\partial T_d} f(z)dz$$

where the ∂ represents the boundary of the shape. This is true by looking at the direction of the arrows. The boundaries of T_c will go in the opposite direction as the sides of the other three, so they will cancel out eventually.

We will try to prove by contradiction. Assume that the integral of f over the boundary of T does not vanish, and set $\left| \int_{\partial T} f(z)dz \right| = C > 0$ for some C. Next, we know that at least one of the four smaller triangles (denote by T_1) will have

$$\left| \int_{\partial T_1} f(z)dz \right| \geq \frac{C}{4}$$

or else the quality above will not hold. We will then repeat the process of dividing a triangle into four to T_1, and get a T_2 that satisfies

$$\left| \int_{\partial T_2} f(z)dz \right| \geq \frac{C}{4^2}.$$

By iterating this process, we will get a sequence of triangles with $T = T_0 \supset T_1 \supset \cdots \supset T_n \supset \cdots$ where

$$\left| \int_{\partial T_n} f(z)dz \right| \geq \frac{C}{4^n}.$$

It is also easy to spot that there exists a point $a \in U$ in every T_n. Also, if we let the perimeter of T to be p, the perimeter of T_n will then be $p/2^n$ since we are bisecting each of the sides to form the smaller triangles. Now, since f is holomorphic in all of U, it must has a derivative at a, and we can then have a function φ by $f(z) = f(a) + f'(a)(z-a) + \varphi(z)(z-a)$ where $\varphi(z) \to 0$ as $z \to a$. Notice that both $f(a)$ and $f'(a)(z-a)$ have antiderivative in U, so by Corollary 2.10, we have

$$\int_{\partial T_n} f(z)dz = \int_{\partial T_n} (f(a) + f'(a)(z-a) + \varphi(z)(z-a))dz = \int_{\partial T_n} \varphi(z)(z-a)dz.$$

Some more, as $a \in T_n$ and z is on the boundary of T_n, we have $|z-a| \le p/2^n$. Then, by Proposition 2.7, we can have the estimate

$$\Big|\int_{\partial T_n} \varphi(z)(z-a)dz\Big| \le \varepsilon_n \frac{p}{2^n} \cdot \frac{p}{2^n}$$

where $\varepsilon_n = \sup_{z\in T_n} |\varphi(z)| \to 0$ as $n \to \infty$ by the way φ is defined.

Combining what we have gotten so far, we will have

$$\Big|\int_{\partial T} f(z)dz\Big| \le 4^n \Big|\int_{\partial T_n} f(z)dz\Big| \le \varepsilon_n p \cdot p.$$

By having $n \to \infty$, the absolute of the integral will be 0, since $\varphi_n \to 0$. Thus we complete the proof. □

Now, we will try to generalise the theorem to a broader range of shapes. But first, we need another result.

Proposition 2.12. Let $U \subset \mathbb{C}$ be a convex open set and $f : U \to \mathbb{C}$ be a holomorphic function. Then f has an antiderivative.

Proof. First, we will use the notation $[a, b]$ to denote the line segment from a to b here.

Choose a point $z_0 \in U$ and define a function $F : U \to \mathbb{C}$ by the formula $F(z) = \int_{[z_0,z]} f(t)dt$. We will then show that this F is an antiderivative of f.

Let $a \in U$, and $a + h \in U$. Let T be a triangle with vertices z_0, a and $a + h$. By convexity, it is contained in U. Then, using Goursat's Theorem, we have

$$0 = \int_{\partial T} f(t)dt = \int_{[z_0,a]} f(t)dt + \int_{[a,a+h]} f(t)dt + \int_{[a+h,z_0]} f(t)dt$$

when we also have

$$F(a+h) - F(a) = -\int_{[z_0,a+h]} f(t)dt - \int_{[z_0,a]} f(t)dt = \int_{[a,a+h]} f(t)dt.$$

Note that

$$\int_{[a,a+h]} f(t)dt = \int_{[a,a+h]} f(a)dt + \int_{[a,a+h]} (f(t) - f(a))dt = f(a) \cdot h + \int_{[a,a+h]} (f(t) - f(a))dt.$$

Using a similar estimation as the proof of Goursat's Theorem, we get

$$\begin{aligned}\Big|\frac{F(a+h) - F(a)}{h} - f(a)\Big| &= \frac{\Big|\int_{[a,a+h]}(f(t) - f(a))dt\Big|}{|h|} \\ &\le \frac{|h| \cdot \sup_{t\in[a,a+h]} |f(t) - t(a)|}{|h|} \\ &= \sup_{t\in[a,a+h]} |f(t) - t(a)|.\end{aligned}$$

Since the function f is holomorphic and then continuous, the right-hand side above will tend to zero as $h \to 0$, which means the derivative of F at a is $f(a)$, so the proof is completed.

□

With this result, we can improve Goursat's Theorem and get a form of Cauchy's Theorem for convex open set.

Theorem 2.13. (Cauchy's Theorem for convex open set) Let $U \subset \mathbb{C}$ be a convex open set and $f : U \to \mathbb{C}$ be a holomorphic function. Then, (a) the integral of f over every closed path in U vanishes; (b) if $p, q \in U$ and γ_1, γ_2 are two paths in U joining p and q, then the integrals of f over γ_1 and γ_2 will be identical.

To generalise this version of the Cauchy's Theorem even further, we will need to know more Topology, mostly homotopy. We do not need much of that to continue this proof to the Fundamental Theorem of Algebra, so we will omit it here. We just need one more corollary to finish this section.

Corollary 2.14. Suppose f is holomorphic in an open set containing the circle C and its interior, then $\int_C f(z)dz = 0$.

Proof. Let D be a disc (open circle) with boundary circle C. Then, there exists a slightly larger disc D; which contains D and so that f is holomorphic on D'. D' is an open convex set, so we can apply Cauchy's Theorem for convex open set to complete the proof. □

2.3 Cauchy Integral Theorem

Theorem 2.15. (Cauchy Integral Theorem) Suppose f is holomorphic in an open set that ocntians the closure of a disc D. If C denotes the boundary circle of this discs with the positive orientation, then

$$f(w) = \frac{1}{2\pi i} \int_C \frac{f(z)}{z - w} dz$$

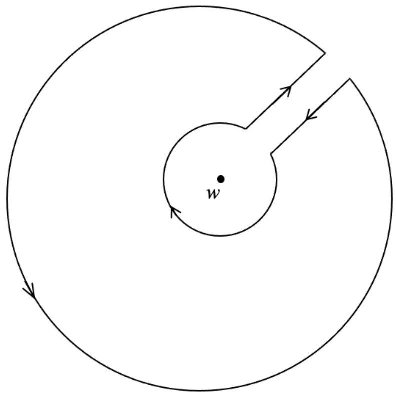
for any $w \in D$.

Proof. Fix $w \in D$ and consider the keyhole $\Gamma_{\delta,\varepsilon}$ which omits the point w as below.

Here, we let δ be the width of the corridor and ε be the radius of the small circle centred at w. Since the function $F(z) = \frac{f(z)}{z-w}$ is holomorphic everywhere except w, we will have

$$\int_{\Gamma_{\delta,\varepsilon}} F(z)dz = 0$$

by Cauchy's theorem. Now, we make the corridor narrower by letting $\delta \to 0$, and eventually these two sides will cancel out over the integrals. The remaining part consists of two curves,



one is the large circle C with positive orientation, and the other is a small circle, called C_ε, centred at w with radius ε and negative orientation. The integral then becomes

$$F(z) = \frac{f(z) - f(w)}{z - w} + \frac{f(w)}{z - w}.$$

Notice that the integral of the first term on the right will go to 0 as $\varepsilon \to 0$. Thus, we have

$$\begin{aligned} \int_{C_\varepsilon} \frac{f(w)}{z-w} dz &= f(w) \int_{C_\varepsilon} \frac{dz}{z-w} \\ &= -f(w) 2\pi i. \end{aligned}$$

So, we have

$$0 = \int_{\Gamma_{\delta,\varepsilon}} F(z)dz = \int_C \frac{f(z)}{z-w} dz - f(w)2\pi i,$$

or simply $f(w) = \frac{1}{2\pi i} \int_C \frac{f(z)}{z-w} dz$. □

If we think about the this theorem for a while, we will realise how remarkable it is! The theorem is saying that if we know the values of f on the boundary curve C, then we can know <u>everything</u> about f inside C. Next, we will have a corollary of it.

Corollary 2.16. (Cauchy Integral Theorem for derivatives) If f is holomorphic in an open set U, then f has infinitely many complex derivatives in U. Moreover, if $C \subset U$ is a circle whose interior is also contained in U, then

$$f^{(n)}(w) = \frac{n!}{2\pi i} \int_C \frac{f(z)}{(z-w)^{n+1}} dz$$

for all w in the interior of C.

This is quite easy to show using induction. This will be left as an exercise for the readers. ⊙$_\omega$⊙

After knowing the formula for derivatives, we can find a bound for it.

Corollary 2.17. (Cauchy Inequalities) If f is holomorphic in an open set that contains the closure of a disc D centred at z_0 and of radius R, then

$$|f^{(n)}(z_0)| \leq \frac{n!M_R}{R^n}$$

where $M_R = \sup_{z \in C} |f(z)|$ denotes the supremum of $|f|$ on the boundary circle C.

Proof. Applying the Cauchy Integral Theorem for derivatives, we have

$$\begin{aligned} |f^{(n)}(z_0)| &\leq \frac{n!}{2\pi} \int_C \frac{|f(w)|}{|w - z_0|^{n+1}} |dw| \\ &\leq \frac{n!}{2\pi} \frac{M_R}{R^{n+1}} \int_C |dw| \\ &= \frac{n!}{2\pi} \frac{M_R}{R^{n+1}} \cdot 2\pi R \\ &= \frac{n!M_R}{R^n}. \end{aligned}$$

□

2.4 The Final Touch

Theorem 2.18. (Liouville's Theorem) Assume $f(z)$ is entire and suppoes it is bounded in the complex plane, namely $|f(z)| < M$ for all $z \in \mathbb{C}$, then $f(z)$ is constant.

Proof. For any circle of radius R around z_0, the Cauchy inequality tells us that $|f'(z_0)| \leq \frac{M}{R}$ when we set $n = 1$. But, R can be any number, so we can let it be extremely large and get $|f'(z_0)| = 0$ for every $z_0 \in \mathbb{C}$. If the derivative is 0, the function is a constant. □

Finally, we have all we need to prove the Fundamental Theorem of Algebra.

Theorem 2.19. (Fundamental Theorem of Algebra) Every non-constant polynomial $P(z) = a_n z^n + \cdots + a_0$ with complex coefficients has a root in $\mathbb{C}$.

Proof. If P has no roots, then $\frac{1}{P(z)}$ is a bounded holomorphic function, since we can have

$$\frac{P(z)}{z^n} = a_n + (\frac{a_{n-1}}{n} + \cdots + \frac{a_0}{z^n})$$

whenever $z \neq 0$. Each term in the bracket will go to 0 as $|z| \to \infty$, and we can say that there exists $R > 0$ such that if $c = \frac{|a_n|}{2}|$, then $|P(z)| \geq c|z|^n$ whenever $|z| > R$ and $\frac{1}{P(z)}$ is therefore bounded. Also, since P is continuous and has no roots in the disc $|z| \leq R$, $\frac{1}{P(z)}$ is bounded too. Thus, the claim earlier of $\frac{1}{P(z)}$ is bounded is shown. Then, according to Liouville's Theorem, $\frac{1}{P}$ is a constant. This contradicts with the condition that P is non-constant. Thus, by contradiction, the theorem is proved.

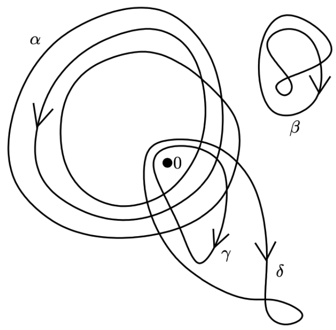
□

3 Topology Proof

In this Chapter, we will be presenting another proof of FTA using Topology, and specifically winding number. We will start with a sketch proof that lacks some rigour and proper definitions of terms. We will then fill up the gaps and complete the proof in the following sections.

3.1 Sketch Proof

To start, we first need the concept of 'winding number'. It is quite easy to see what this term means via a graphical example.



Winding number of a loop γ is the number of times that γ winds anticlockwise around a point, for example 0. We will denote this number by $\#\gamma$, and we can see that in the above figure, $\#\alpha = 3$, $\#\beta = 0$, $\#\gamma = -1$, and $\#\delta = -1$. More over, if two loops can be continuously deformed into each other without crossing zero, these two loops will have the same winding number. This relationship between the two loops is known as **homotopty**, and we will go into more detail later on this Chapter. In this example, γ and δ are homotopic, so they have the same winding number -1.

Next, take a complex polynomial $f(z) = a_0 + a_1 z + \cdots + a_n z^n$ of degree n, and suppose that f has no complex root. Then, we take an $R \in [0, \infty)$. As the input z goes around in circle $\{z||z| = R\}$ in the anticlockwise direction, $f(z)$ will form a loop γ_R in $\mathbb{C}$ and it does not cross 0, since $f(z)$ has no complex root and no $z_0 \in \mathbb{C}$ will let $f(z_0) = 0$.

(1) We know that as R increases, γ_R will only change continuously, since f is a continuous function. So, all possible γ_R can be deformed into each other and are thus homotopic, which implies that $\#\gamma_R$ is independent of the value of R.

(2) When $|z|$ is large, $f(z)$ will behave like $a_n z^n$, then the remaining terms will be negligible in this case. For z travelling on the circle $\{z||z| = R\}$ in the anticlockwise direction, it will wind around 0 n times. So, for big enough R, $\#\gamma_R = n$.

(3) When $R = 0$, γ_0 will be a constant at $f(0) \neq 0$, so $\#\gamma_0 = 0$.

If we compile all these information, we will get $\#\gamma_0 = \#\gamma_R = 0 \ \forall R \in [0, \infty)$, so $n = 0$ and f is a constant. This contradicts with the condition of FTA, so there exists $z_0 \in \mathbb{C}$ as the root of any non-constant complex polynomial.

3.2 Algebraic Topology Basics

Now, let us begin to define things properly and rigorously. Topology deals with different shapes, and one of the simplest shape is that of path and loop. It is essentially the same concept as the path and loop we defined earlier in the previous Chapter, but there is some very minor difference in their definitions. A **path** in a space X is a continuous map $f : I \to X$ where I is the unit interval $[0, 1]$. If the starting and ending points of the path are identical, i.e. $f(0) = f(1)$, then it becomes a **loop**.

Next, we will start to define what we mean by a continuous deformation. A continuous deformation, or a **homotopy**, of paths in X is a family $f_s : I \to X$, $0leS \leq 1$ such that (1) the endpoints $f_s(0) = x_0$ and $f_s(1) = x_1$ are independent of t, meaning that they are all fixed regardless of the value of t. (2) the associated map $F : I \times I \to X$ defined by $F(s, t) = f_s(t)$ is continuous. When two paths f_0 and f_1 are connected in this way by a homotopy f_s, they are said to be **homotopic**, and we denote this relationship by $f_0 \simeq f_1$.

Example 3.1. (Linear Homotopies) Any two paths f_0 and f_1 in $\mathbb{R}^n$ having the same endpoints x_0 and x_1 are homotopic via the homotopy $f_s(t) = (1 - s)f_0(t) + sf_1(t)$. The endpoints will not be changed during this homotopy, and each point will travel along the line segment at a constant speed. This $f_s(t)$ is continuous since the vector addition and scalar multiplication of f_0 and f_1 are continuous. This construction can be generalised to any convex subspace $X \subset \mathbb{R}^n$, since this homotopy will work given the convexity of space.

We can find some algebraic structure of homotopies.

Proposition 3.2. The relation of homotopy on paths with fixed endpoints in any space is an equivalence relation.

Proof. Recall that an equivalence relation needs to satisfy three conditions (1) reflexivity, (2) symmetry, and (3) transitivity.

Reflexivity is easy to spot, since we can have the constant homotopy $f_s = f$ to construct the relation. Symmetry is also obvious, as we can get $f_1 \simeq f_0$ from $f_0 \simeq f_1$ with homotopy f_s by having the inverse homotopy f_{1-s}. Transitivity is the one that needs some work.

If we have $f_0 \simeq f_1$ via f_s, and $g_0 = f_1 \simeq g_1$ via g_s, we need to show that $f_0 \simeq g_1$ with homotopy h_s. This homotopy can be found by connecting f_s and g_s, by having f_{2s} for $0 \leq s < 1/2$, and g_{2s-1} for $1/2 \leq s \leq 1$. These two components agree on $s = 1/2$, since $g_0 = f_1$. The continuity is obvious since it is the union of two continuous functions with the same value at the intersection point. $\square$

Knowing the equivalence relation of the path homotopy, we will denote the equivalence class of a path f by homotopy by $[f]$, and it is called the **homotopy class** of f.

From the above proof, we can give a name to what we did to show the transitivity of homotopy. Given two paths $f, g : [0,1] \to X$ such that $f(1) = g(0)$, there is a **composition product path** $f \circ g$ that traverses first f and then g, defined by

$$f \circ g(t) = \begin{cases} f(2t), & 0 \leq t \leq \frac{1}{2} \\ g(2t-1), & \frac{1}{2} \leq t \leq 1. \end{cases}$$

Remark. Do take note that we are traversing first with f then with g, it is not how we normally think about compositions of functions. It is not the most intuitive way to think about it, but this has been the case for the textbooks I am referring to, so I guess that is the norm.

This product operation respects homotopy classes, since if $f_0 \simeq f_1$ and $g_0 \simeq g_1$ via the homotopy f_s and g_s respectively, and if we have $f_0(1) = g_0(0)$ so $f_0 \circ g_0$ is defined, $f_s \circ g_s$ is defined and provides a homotopy $f_0 \circ g_0 \simeq f_1 \circ g_1$. If instead of paths, we have a homotopy of loops, then the common starting and ending points x_0 is known as the **basepoint**. The set of all homotopy classes $[f]$ of loops $f : I \to X$ at the basepoint x_0 is denoted by $\pi_1(X, x_0)$. This set has some interesting algebraic structure too, and in particular:

Proposition 3.3. $\pi_1(X, x_0)$ is a group with respect to the product operation $[f] \circ [g] = [f \circ g]$.

Proof. Recall the axioms of the group structure, we need (1) associativity, (2) existence of identity, and (3) existence of inverse.

We first check if the operation is associative, meaning that $[\alpha \circ \beta] \circ [\gamma] = \alpha \circ [\beta \circ \gamma]$ for any three loops α, β, γ based at the same point p. We will do a reparametrisation of a path f to make it as $f\varphi$ where $\varphi : [0,1] \to [0,1]$ is a continuous map such that $\varphi(0) = 0$ and $\varphi(1) = 1$. The reparametrisation will preserve the homotopy class, as $f\varphi \simeq f$ via the homotopy $f\varphi_t$ where $\varphi_s(t) = (1-s)\varphi(t) + st$ such that $\varphi_0 = \varphi$ and $\varphi_1(t) = t$. This method should not be brand new any more, as we have just used this for the example on linear homotopies. With this in mind, we can do a reparametrisation to $(\alpha \circ \beta) \circ \gamma$ using

$$\varphi(t) = \begin{cases} 2t & 0 \leq t \leq \frac{1}{4} \\ t + \frac{1}{4} & \frac{1}{4} \leq t \leq \frac{1}{2} \\ \frac{t+1}{2} & \frac{1}{2} \leq t \leq 1 \end{cases}$$

, which will show us the associativity.

The identity element of the group will be the constant path defined to be $c(t) = f(1)$ for all $t \in [0,1]$. With this path, we can get $f \circ c \simeq f$ for a path $f : [0,1] \to X$. Also, we can get another constant path $c(t) = f(0)$ for all $t \in [0,1]$. With this path, we can get $c \circ f \simeq f$ for a path $f : [0,1] \to X$. Since we are dealing with loops, $f(0) = f(1)$ and the two constant paths are identical, making it the identity element.

For path $f(t)$, $f(1-t)$ will be the inverse path of it. The remaining details can be filled by the readers without much difficulties. □

This group $\pi_1(X, x_0)$ is called the **fundamental group** of X at the basepoint x_0. The subscript 1 of π is used to denote this group as the first group in a sequence of groups $\pi_n(X, x_n)$, called the homotopy groups. We will not be discussing them in this notes. Interested readers can refer to any textbook on Algebraic Topology, for example the classic by Allen Hatcher[3].

Let us think about what this group is trying to do for a bit. In Topology, we are frequently, in fact always, dealing with topological invariants - properties that will be preserved by homeomorphism. If you do not know already, a homeomorphism is a bijective map $f : X \to Y$ for topological spaces X and Y when both f and its inverse are continuous. One goal of Topology is to classify different shapes and spaces up to homeomorphism. If two shapes have different values for a particular topological invariant, they will not be the same in the topological sense. Among the various invariants, one common one, constructed by Poincaré, is the fundamental group. For example, do you think the fundamental group of a disc will be the same as that of an annulus? You can try to draw them out and play around with it, hopefully you will get the answer of "No". In the same example, do you think the basepoint you choose will affect the fundamental group? For points in a disc, do you think the fundamental group will be different for two distinct points? This brings us to the question on change of basepoint.

We define a **change of basepoint** map $\beta_h : \pi_1(X, x_0) \to \pi_1(X, x_1)$ by $\beta_h[f] = [h \circ f \circ h^{-1}]$, where $h : [0, 1] \to X$ is an invertible path from x_0 to x_1. We let the inverse of h be $h^{-1} = h(1-t)$ that goes from x_1 to x_0. We can then associate each loop f based at x_1 the loop $h \circ f \circ h^{-1}$ based at x_0.

Proposition 3.4. The map $\beta_h : \pi_1(X, x_0) \to \pi_1(X, x_1)$ is an isomorphism.

Proof. First, we see that β_h is a homomorphism since $\beta_h[f \circ g] = [h \circ f \circ g \circ h^{-1}] = [h \circ f \circ h^{-1} \circ h \circ g \circ h^{-1}] = \beta_h[f] \circ \beta_h[g]$. It is an isomorphism with inverse $\beta_{h^{-1}}$ since $\beta_h\beta_{h^{-1}}[f] = \beta_h[h^{-1} \circ f \circ h] = [h \circ h^{-1} \circ f \circ h \circ h^{-1}] = [f]$. Similarly, $\beta_{h^{-1}}\beta_h[f] = [f]$. Thus, the map is indeed an isomorphism. □

If X is path-connected, meaning that we can connect any pair of points in the space using a path, we can have the above map for every pair points. Since we are working algebraic objects, two fundamental groups being isomorphic means that they are the 'same'. So, the basepoint is not too important for a path-connected space X, and we can sometimes drop x in the notation and just write the fundamental group as $\pi_1(X)$.

Now, if a space is both path-connected and has trivial fundamental group of 0, then it is called **simply-connected**.

3.3 Fundamental Group of Circle

Theorem 3.5. $\pi_1(S^1)$ is an infinite cyclic group generated by the homotopy class of the loop $\omega(t) = (\cos 2\pi t, \sin 2\pi t)$ based at $(1, 0)$.

Remark. The proof of this theorem will be omitted here. Interested readers should refer to [1] or [3]. I will add this part later if there is such a demand. Do let me know via email if you would like to see the proof of it in this notes.

3.4 The Final Touch

Theorem 3.6. (Fundamental Theorem of Algebra) Every non-constant polynomial with coefficients in $\mathbb{C}$ has a root in $\mathbb{C}$.

Proof. We may assume the polynomial is of the form $p(z) = z^n + a_1 z^{n-1} + \cdots + a_n$. If $p(z)$ has no roots in $\mathbb{C}$, then for each real number $r \geq 0$, the formula

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

defines a loop in the unit circle $S^1 \subset \mathbb{C}$ bases at 1. As r varies, f_r is a homotopy of loops based at 1. Since f_0 is the trivial loop, we deduce that the class $[f_r] \in \pi_1(S^1)$ is zero for all r. Now, fix a large value of r, bigger than $|a_1| + \cdots + |a_n|$ and bigger than 1. Then, for $|z| = r$, we have

$$|z^n| \geq (|a_1| + \cdots + |a_n|)|z^{n-1}| > |a_1 z^{n-1}| + \cdots + |a_n| \geq |a_1 z^{n-1} + \cdots + a_n|.$$

This inequality is saying that the polynomial $p_t(z) = z^n + t(a_1 z^{n-1} + \cdots + a_n)$ has no roots on the circle $|z| = r$ when $0 \leq t \leq 1$. Replacing p by p_t in the formula for f_r to the loop $\omega_n(t) = e^{s\pi int}$. Using the fundamental group of circle, ω_n represents n times a generator of the infinite cyclic group $\pi_1(S^1)$. Since we have shown that $[\omega_n] = [f_r] = 0$, we conclude that $n = 0$. Thus, the only polynomials without roots in $\mathbb{C}$ are constants. $\square$

References

[1] M. A. Armstrong. *Basic Topology*, Undergraduate Texts in Mathematics. Springer, 1979.

[2] Benjamin Fine & Gerhard Rosenberger. *The Fundamental Theorem of Algebra*, Undergraduate Texts in Mathematics. Springer, 1997.

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

[4] Serge Lang. *Complex Analysis, Fourth Edition*, Graduate Texts in Mathematics 103. Springer, 1999.

[5] Serge Lvovski. *Principles of Complex Analysis*, Moscow Lectures 6. Springer, 2020.

[6] Elias M. Stein & Rami Shakarchi. *Complex Analysis,* Princeton Lectures in Analysis III. Princeton University Press, 2003.