A PROBABILISTIC PROOF
OF THE FUNDAMENTAL THEOREM OF ALGEBRA

MIHAI N. PASCU

(Communicated by Richard C. Bradley)

I dedicate this paper to my dear friend M. K.

Abstract. We use Lévy’s theorem on invariance of planar Brownian motion under conformal maps and the support theorem for Brownian motion to show that the range of a non-constant polynomial of a complex variable consists of the whole complex plane. In particular, we obtain a probabilistic proof of the fundamental theorem of algebra.

1. Introduction

The Fundamental Theorem of Algebra is a well-established result in mathematics, and there are several proofs of it in the mathematical literature. In the present paper we give a new probabilistic proof of it, as follows:

Theorem 2.4. If \( f : \mathbb{C} \to \mathbb{C}, f(z) = a_n z^n + a_{n-1} z^{n-1} + \ldots + a_1 z + a_0 \) (\( a_n \neq 0 \), \( a_0, \ldots, a_n \in \mathbb{C} \)) is a polynomial of degree \( n \geq 1 \), then \( f(\mathbb{C}) = \mathbb{C} \).

In particular, the equation \( f(z) = 0 \) has at least one root in \( \mathbb{C} \).

The purpose of this paper is two-fold: on the one hand, we show how probabilistic methods can be used in order to give a simple proof of this particular result, and on the other hand, we present a method which might be suitable for deriving new properties about the range of analytic functions.

Probabilistic methods have proved to be a useful tool in Analysis. One could, for example, refer to monographs such as [2] and [3]; among others, they contain probabilistic proofs of analytic results such as: Privalov’s and Plessner’s Theorem ([3], pp. 132 – 134), Koebe’s \( 1/4 \) Theorem ([2], pp. 323 – 326) and Picard’s Theorem ([3], pp. 139 - 142 or [2], pp. 320 – 322), to mention just a few well-known results related to the range of analytic functions (incidentally, both proofs of Picard’s Theorem mentioned above have gaps; see the web site of the second author for a revised proof). For some recent activity in this area, one could see for example [4].

The present proof can also be modified to give a proof of Picard’s Theorem in the particular case of a \( p \)-valued entire function (see the remark preceding Theorem 2.5),

Received by the editors October 10, 2003 and, in revised form, February 4, 2004.
2000 Mathematics Subject Classification. Primary 30C15; Secondary 60J65.
Key words and phrases. Brownian motion, Lévy’s theorem, support theorem.
This work was supported in part by NSF grant # 0203961 - DMS.

\( \copyright \)2004 American Mathematical Society
Reverts to public domain 28 years from publication

1707

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use
and shows explicitly the reason for which the $p$-valuedness is a sufficient condition for an entire function to omit no values in $\mathbb{C}$.

While interesting in their own, probabilistic methods shed a new perspective on the subject and are able sometimes to provide answers to open problems in Analysis (see for example a partial resolution of the Hot Spots conjecture by probabilistic arguments in [5]).

2. A PROBABILISTIC PROOF OF THE FUNDAMENTAL THEOREM OF ALGEBRA

2.1. Preliminaries. We begin by reviewing some ingredients needed for the proof.

Recall that a complex map $f : \mathbb{C} \to \mathbb{C}$ is called entire if it is analytic in $\mathbb{C}$.

For a given integer $p \geq 1$, a complex map $f : D \to \mathbb{C}$ is said to be $p$-valued if for any $w \in \mathbb{C}$ there are at most $p$ solutions in $D$ to the equation $f(z) = w$, and there exists $w_0 \in \mathbb{C}$ for which the equation $f(z) = w_0$ has exactly $p$ roots in $D$.

We recall the following result on local correspondence for analytic maps (see [1], p. 131):

**Theorem 2.1.** Suppose that $f(z)$ is analytic at $z_0$, $f(z_0) = w_0$ and that $f(z) - w_0$ has a zero of order $n$ at $z_0$. If $\varepsilon > 0$ is sufficiently small, there exists $\delta > 0$ such that for all $w$ with $|w - w_0| < \delta$ the equation $f(z) = w$ has exactly $n$ roots in the disk $|z - z_0| < \varepsilon$.

We will need the following support theorem for Brownian motion (see [2], p. 59, for a proof), showing that with positive probability the paths of a Brownian motion follow the graph of a continuous function:

**Theorem 2.2.** If $\varphi : [0, t] \to \mathbb{R}^d$ is continuous ($d \geq 1$), $B_t$ is a $d$-dimensional Brownian motion starting at $B_0 = \varphi(0)$ and $\varepsilon > 0$, then there exists $c > 0$ such that

$$P^{\varphi(0)}\left(\sup_{s \leq t} \|B_s - \varphi(s)\| < \varepsilon\right) > c,$$

where $c$ can be taken to depend only on $t$, $\varepsilon$ and the modulus of continuity of $\varphi$.

Also, we will need Lévy’s theorem on conformal invariance of planar Brownian motion, showing that the image of a planar Brownian motion under an entire map is a time change of another Brownian motion (see [2], p. 310, for a proof). We have:

**Theorem 2.3.** Let $f : \mathbb{C} \to \mathbb{C}$ be an entire map and $B_t$ a 2-dimensional Brownian motion starting at $B_0 = x$. Then $f(B_{\alpha_t})$ is a 2-dimensional Brownian motion starting at $f(B_0) = f(x)$, where

$$\alpha_t = \inf\{s : A_s \geq t\} \text{ and } A_t = \int_0^t |f'(B_s)|^2 ds.$$

Recall that given a domain $D$, a closed curve $\gamma \subset D$ is said to be homotopic to zero in $D$ if the curve $\gamma$ can be deformed continuously in $D$ to a constant curve. It is known that the homotopy is a topological property (it is preserved under continuous mappings).

We denote by $D(z, r)$ the open disk centered at $z \in \mathbb{C}$ of radius $r > 0$. 

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use
2.2. Main results. We are now ready to prove the Fundamental Theorem of Algebra, as follows:

**Theorem 2.4.** If \( f : \mathbb{C} \to \mathbb{C} \), \( f(z) = a_n z^n + a_{n-1} z^{n-1} + \ldots + a_1 z + a_0 \) (\( a_n \neq 0 \), \( a_0, \ldots, a_n \in \mathbb{C} \)) is a polynomial of degree \( n \geq 1 \), then \( f(\mathbb{C}) = \mathbb{C} \).

In particular, the equation \( f(z) = 0 \) has at least one root in \( \mathbb{C} \).

The idea of the proof is that if \( f \) omits a certain value in \( \mathbb{C} \), then by the support theorem, a planar Brownian motion will wind around this value with positive probability; if we choose this Brownian motion to be the image under \( f \) of a Brownian motion in the domain of \( f \), this will allow us to construct (on a set of positive probability) a closed curve \( \gamma \) in the domain of \( f \) that is homotopic to zero, while its image \( \Gamma = f(\gamma) \) is not, thus obtaining a contradiction. Here are the details:

**Proof.** Assume \( f \) never assumes the value \( w_0 \in \mathbb{C} \), that is, the equation \( f(z) = w_0 \) has no roots in \( \mathbb{C} \).

Since \( f \) is a (non-constant) polynomial of degree \( n \geq 1 \), there exists \( p \leq n \) such that \( f \) is a \( p \)-valent function. Therefore, for any \( w \in \mathbb{C} \), the equation \( f(z) = w \) has at most \( p \) roots in \( \mathbb{C} \), and there exists \( w_1 \in \mathbb{C} \) such that the equation \( f(z) = w_1 \) has exactly \( p \) roots in \( \mathbb{C} \).

By replacing \( f(z) \) with

\[
\tilde{f}(z) = \frac{f(z) - w_0}{w_1 - w_0},
\]

if necessary, we may assume that \( w_0 = 0 \) and \( w_1 = 1 \), that is, the equation \( f(z) = 0 \) has no roots in \( \mathbb{C} \), and the equation \( f(z) = 1 \) has exactly \( p \) roots in \( \mathbb{C} \).

Let \( z_1, z_2, \ldots, z_m \) be the distinct roots of \( f(z) = 1 \) in \( \mathbb{C} \), with multiplicities \( n_1, n_2, \ldots, n_m \) (and therefore \( n_1 + n_2 + \ldots + n_m = p \)).

By the continuity of \( f \), we can choose \( \varepsilon > 0 \) small enough so that it satisfies

\[
|f(z) - 1| < \frac{1}{4} \text{ for all } z \in \bigcup_{i=1}^{m} D(z_i, \varepsilon).
\]

By choosing a smaller \( \varepsilon > 0 \), if necessary, and using Theorem 2.1 it follows that there exist \( \delta_1, \delta_2, \ldots, \delta_m > 0 \) such that for \( |w - 1| < \delta_i \), the equation \( f(z) = w \) has exactly \( n_i \) roots in \( |z - z_i| < \varepsilon \), for all \( i \in \{1, \ldots, n\} \). Since \( n_1 + n_2 + \ldots + n_m = p \) and \( f \) is \( p \)-valent (thus \( f(z) = w \) cannot have more than \( p \) roots in \( \mathbb{C} \)), it follows that these roots are all the roots of the equation \( f(z) = w \) in \( \mathbb{C} \). We obtained the following:

\[
|f(z) - 1| < \delta \Rightarrow z \in \bigcup_{i=1}^{m} D(z_i, \varepsilon),
\]

where \( \delta = \min\{\delta_1, \delta_2, \ldots, \delta_m\} \); without loss of generality, we can also assume \( \delta < \frac{1}{4} \).

Consider now a Brownian motion \( B_t \) in \( \mathbb{C} \) starting at \( B_0 = z_1 \).

Since \( f \) is entire, by Theorem 2.3 it follows that \( W_t = f(B_{\alpha t}) \) is a Brownian motion starting at \( f(B_0) = f(z_1) = 1 \), where the time change \( \alpha_t \) is the inverse of the nondecreasing process

\[
A_t = \int_0^t |f'(B_s)|^2 \, ds.
\]
The support theorem for Brownian motion (Theorem 2.2) shows that with positive probability we have:

\[(2.6)\quad |W_t - e^{2\pi it}| < \delta, \text{ for all } t \in [0, m].\]

Fix now an \(\omega\) in the sample space \(\Omega\) for which (2.6) holds. Since \(|f(B_{\alpha_j}(\omega)) - 1| = |W_j(\omega) - 1| = |W_j(\omega) - e^{2\pi it_j}| < \delta\) for all \(j \in \{0, 1, \ldots, m\}\), and using (2.3), it follows that \(B_{\alpha_j}(\omega) \in \bigcup_{i=1}^{m} D(z_i, \varepsilon)\), for all \(j \in \{0, 1, \ldots, m\}\).

The box principle shows that there exist \(0 \leq j < k \leq m\) such that \(B_{\alpha_j}(\omega), B_{\alpha_k}(\omega) \in D(z_l, \varepsilon)\), for some \(1 \leq l \leq m\).

Consider the closed curve \(\gamma\) formed by concatenation of \([z_l, B_{\alpha_j}(\omega)], B_{\alpha_k}(\omega) \ (j \leq t \leq k)\) and \([B_{\alpha_k}(\omega), z_l]\), and let \(\Gamma = f(\gamma)\) be the image of \(\gamma\) under \(f\) (we denoted by \([x, y]\) the straight line segment with endpoints \(x\) and \(y\)).

By construction \([z_l, B_{\alpha_j}(\omega)], [B_{\alpha_k}(\omega), z_l] \subset D(z_l, \varepsilon) \subset \bigcup_{i=1}^{m} D(z_i, \varepsilon)\), and using (2.3) it follows that \(f([z_l, B_{\alpha_j}(\omega)]), f([B_{\alpha_k}(\omega), z_l]) \subset D(1, 1/\delta)\). By the choice of \(\omega\) it follows that \(W_t(\omega) = f(B_{\alpha_j}(\omega)), j \leq t \leq k\), lies in the \(\delta\)-tube about the unit circle centered at the origin, and since \(\delta < 1/\pi\), it follows that the index of the curve \(\Gamma\), with respect to 0 is the same as the index of the curve \(e^{2\pi it}, j \leq t \leq k\), with respect to this point:

\[(2.7)\quad n(\Gamma, 0) = n\left(\left(e^{2\pi it}\right)_{j \leq t \leq k}, 0\right) = k - j \neq 0,\]

which shows that \(\Gamma\) is not homotopic to zero in \(f(\mathbb{C})\) (recall that \(0 \notin f(\mathbb{C})\)).

However, this contradicts the fact that \(\Gamma = f(\gamma)\) is the image under the continuous function \(f\) of the curve \(\gamma\), which is homotopic to zero in the domain \(\mathbb{C}\) of \(f\) and should be therefore also homotopic to zero in the range of \(f\).

The contradiction obtained shows that \(f\) cannot omit the value 0, and therefore (see the remarks at the beginning of the proof) the range of \(f\) must consist of all of \(\mathbb{C}\), that is, \(f(\mathbb{C}) = \mathbb{C}\).

A careful examination of the proof above shows that we did not explicitly use the fact that \(f\) is a polynomial; the only part where we have used the fact that \(f\) is a polynomial was in showing that \(f\) is a \(p\)-valent (entire) function in \(\mathbb{C}\). Therefore, the above proof also shows the following:

**Theorem 2.5.** If \(f : \mathbb{C} \to \mathbb{C}\) is an entire \(p\)-valent function \((p \geq 1)\), then \(f\) does not omit any value in \(\mathbb{C}\).

Unfortunately, the class of \(p\)-valent entire functions coincides with the class of polynomials of degree \(p\), so the above theorem is equivalent to Theorem 2.4. However, one could use the ideas in the proof above and apply them, for example, to the (larger) class of \(p\)-valent analytic functions in the unit disk, in order to obtain similar statements about the range of such functions.

**References**


Department of Mathematics, Purdue University, West Lafayette, Indiana 47907-2067
E-mail address: pascu@math.purdue.edu

Current address: Faculty of Mathematics and Computer Science, “Transilvania” University of Brașov, Str. Iuliu Maniu Nr. 50, Brașov, Jud. Brașov – COD 2200, Romania
E-mail address: mihai.pascu@unitbv.ro