CONSTANTS ARE DEFINABLE IN RINGS OF ANALYTIC FUNCTIONS

TANELI HUUSKONEN

(Communicated by Andreas R. Blass)

Abstract. The analytic functions defined in a fixed domain form a ring with pointwise addition and multiplication. We describe a way to define constants in the ring using a first-order formula which is independent of the domain.

1. Preliminaries

Our motivation comes from [BHR], which is an impressive study of an otherwise fairly obscure area lying between mathematical logic and complex function theory. (For some further developments, see [Hu] and [Ru].) As this area could conceivably interest analysts with little background in logic and logicians with little background in analysis, we have tried to be fairly generous with our explanations.

Let $G \subseteq \mathbb{C}$ be a domain in the complex plane, i.e., an open connected set. The analytic functions defined on $G$, together with pointwise addition and multiplication and the constant functions 0 and 1, form a ring, which is denoted $H(G)$. The constant function $i$ is also included in the structure to distinguish it from $-i$. Defining multiplication by complex constants turns the ring into a complex algebra, $H_C(G)$. Two domains $G$ and $G'$ are conformally equivalent if and only if the corresponding function rings $H(G)$ and $H(G')$ are isomorphic. The "if" direction is proved in [Be], for instance; the "only if" is quite easy (see [BHR], for example).

Let us define formal languages for describing the structures $H(G)$ and $H_C(G)$; they are called the ring language and the algebra language, respectively. A first-order atomic formula of the ring language is an equation both sides of which are expressions formed of variables ($x, y, z, \ldots$) and constants (0, 1, $i$) using addition and multiplication. In the algebra language, we also allow atomic formulas of the form $\text{Const}(x)$, where $x$ is any variable. The formula $\text{Const}(x)$ is defined to be true iff the value of $x$ is a constant function. The word “first-order” refers to the fact that the variables denote...
members of the ring, never its subsets or the like. A first-order formula is obtained from atomic formulas by propositional connectives (¬, ∨, ∧, →, ↔) and quantifiers (∃, ∀) by well-known rules, using brackets where necessary to avoid ambiguity. A sentence is a formula all of whose variables are bound by quantifiers. The theory of a structure is the set of sentences which are true in it.

Much of the work in [BHR] is based on the assumption that constant functions are definable in the language used. This condition is trivially met in the algebra language, which is indeed shown to be very expressive. Just to mention a few examples, they prove that there is an algebra language sentence which is true in \( H_C(G) \) iff \( G \) is finitely connected; that every sentence \( \varphi \) in second-order number theory can be effectively translated into an algebra language sentence \( \varphi' \) such that \( \varphi \) is true in the standard natural numbers iff \( \varphi' \) is true in \( H_C(D) \), where \( D \) denotes the unit disc; that every finitely connected domain \( G \) is determined up to conformal equivalence by the first-order theory of \( H_C(G) \); and finally, that it is consistent with the standard ZFC axioms of set theory that every domain be determined up to conformal equivalence by the first-order theory of the corresponding algebra.

We present a way of defining the constants by a ring language formula, thus answering open question Q7 mentioned in Chapter 10 of [BHR]. Moreover, this enables us to apply the theorems listed above directly to the ring language, thereby settling open questions Q3, Q4, Q8, and Q14 in addition. Our result will allow for a generalization of the main theorem of [Ru] as well, as noted by its author.

I have been informed by several people that Jan Denef had previously found a definition of the constants in the unit disc, but that he had never published this result. His method did not generalize to other domains, though.

2. Defining the constants

This section contains the actual definition of the constants. Throughout the section \( G \) will denote a fixed domain. For \( z \in \mathbb{C}, \ r > 0 \), we write \( U_r(z) \) for the \( z \)-centred open disc of radius \( r \).

2.1. Definition [BHR]. An analytic function \( f \) defined on \( G \) is a point function if it is not a unit (i.e., it has no multiplicative inverse) and not the product of two nonunits.

It is obvious how to define a point function formally in the ring language. A point function derives its name from the fact that there is exactly one point in \( G \) where the function gets the value zero; the zero is of order one. Therefore any two point functions are divisible by each other iff they have the same zero point. In general, any function is divisible by a point function iff it has a zero at the zero of the latter. Hence, point functions can be used to code points of \( G \), and other functions to code sets of points which have no limit points in \( G \). The zero point of a point function \( p \) is denoted by \( Z(p) \), and \( P_{z_0} \) denotes the function \( f : z \mapsto z - z_0 \); this is a point function having its zero at \( z_0 \), if \( z_0 \in G \).

2.2. Definition. Let \( p_0 \) and \( p_1 \) be point functions. A function \( q \) is \( (p_0, p_1) \)-quasilinear if \( q \) is a point function dividing \( p_0 \) and \( q - 1 \) is a point function dividing \( p_1 \).
Clearly a \((p_0, p_1)\)-quasilinear function is \((q_0, q_1)\)-quasilinear if \(Z(p_0) = Z(q_0)\) and \(Z(p_1) = Z(q_1)\). Therefore we may sometimes abuse the notation and write \((z_0, z_1)\)-quasilinear instead of \((P_{z_0}, P_{z_1})\)-quasilinear. Note that the linear transformation \(L_{z_0, z_1}: z \mapsto (z - z_0)/(z_1 - z_0)\) is \((z_0, z_1)\)-quasilinear.

A family \(M\) of analytic functions defined in a domain \(B\) is called normal if every sequence from \(M\) contains a subsequence which converges uniformly in compact subsets of \(B\) to an analytic function or to infinity.

2.3. Theorem \([Mo]\). Let \(M\) be a family of functions which are analytic in a domain \(B\). Suppose that there are two distinct finite complex values \(a\) and \(b\) such that no function \(f \in M\) assumes the value \(a\) anywhere in \(B\) and no single function \(f \in M\) assumes the value \(b\) in more than \(p\) points in \(B\), where \(p \in \mathbb{N}\). Then \(M\) is normal in \(B\).

Proof. See \([Mo, \text{Chapter II, §§36-37}]\); quoted in \([Go, \text{Chapter II, §7, Theorem 2}]\).

2.4. Lemma. For all real numbers \(\varepsilon > 0\) there is \(R > 0\) such that if \(f\) is analytic in \(U_R(0)\), satisfying \(f(z) = 0\) iff \(z = 0\) and \(f(z) = 1\) iff \(z = 1\), then \(|f(z) - z| < \varepsilon\) for all \(z \in U_1(0)\).

Proof. Assume that the claim does not hold for some \(\varepsilon\). For \(n \in \mathbb{N}\), \(n > 1\), let \(f_n\) be a counterexample defined in \(U_n(0)\). Now the set \(\{f_n \mid U_2(0) \setminus \{0\}: n > 1\}\) is a normal family by Theorem 2.3, since all the functions omit the value 0 and get the value 1 only once in \(U_2(0) \setminus \{0\}\). Hence there is a subsequence \(\langle f_{n_i} \mid U_2: i \in \mathbb{N}\rangle\) which converges to an analytic function \(f\) in \(U_2(0) \setminus \{0\}\); it cannot converge to infinity, since \(f_{n_i}(1) = 1\) for all \(i\). Similarly, the set

\[F_r = \{f_{n_i} \mid U_r(0) \setminus \{1\}: i \in \mathbb{N}, \ n_i > r\}\]

is a normal family for every \(r > 0\). Therefore the function \(f\) has an analytic continuation to every \(U_r(0)\), defined as the limit of a convergent subsequence of \(F_r\). Note that 0 belongs to the domain of every member of \(F_r\) for all \(r\).

By the uniqueness of an analytic continuation, \(f\) can be continued analytically to the whole complex plane. We denote the continuation by \(f\) as well, abusing the notation. Now \(f(0) = 0\) and \(f(1) = 1\); moreover, \(f(z) \neq 0, 1\) for \(z \neq 0, 1\), since \(f_{n_i}(z) \neq 0, 1\) for \(z \neq 0, 1\), by a well-known argument, which can be found, inter alia, in \([Go, \text{Chapter I, §2, Theorem 2}]\). According to Picard's Great Theorem (see \([Go, \text{Chapter II, §7, Theorem 4}]\), for example), this implies that \(f\) cannot have an essential singularity at infinity; thus \(f\) must be a rational function, and hence, being entire, a polynomial. A polynomial of degree greater than one assumes all complex values more than once, with at most one exception. Therefore \(f\) must be the identity, but this contradicts the assumption that every \(f_{n_i}\) is a counterexample to the original claim.

2.5. Lemma. For every real number \(\varepsilon > 0\) and every \(z_0 \in G\), there is \(r_0 > 0\) such that \(U_{r_0}(z_0) \subseteq G\) and if \(0 < |z - z_0| < r < r_0\), \(p\) is \((z_0, z)\)-quasilinear, and \(f\) is analytic in \(G\) satisfying \(|f(w)| < 1 - \varepsilon\) for all \(w \in U_r(z_0)\), then there is some \(z' \in G\) such that \(f(z') = p(z')\) and \(|z' - z_0| < (1 - \varepsilon/2)|z - z_0|\).

Proof. Assume that \(\varepsilon > 0\) and \(z_0 \in G\) are such that the claim does not hold. Let \(R > 0\) be such that \(U_R(z_0) \subseteq G\). For each \(n \in \mathbb{N}\), \(n > 1\), let \(f_n, z_n, p_n\) form a counterexample showing that the value \(r_0 = R/n\) does not
satisfy the conditions. Define functions $g_n, u_n : U_n(0) \to \mathbb{C}$ as follows:

\[
\begin{align*}
  g_n(w) &= f_n(z_0 + (z_n - z_0)w), \\
  u_n(w) &= p_n(z_0 + (z_n - z_0)w).
\end{align*}
\]

Then $|g_n(w)| < 1 - \varepsilon$ for all $w \in U_1(0), \ n \in \mathbb{N}$. Moreover, by the previous lemma, it can be seen that the sequence $\{u_n \mid U_1(0) : n \in \mathbb{N}, \ n > 0\}$ converges uniformly to the identity. Now choose some $n \in \mathbb{N}$ such that $|u_n(w)-w| < \varepsilon/2$ for all $w \in U_1(0)$. Then the following inequality holds for all $w \in U_1(0)$:

\[
|w + g_n(w) - u_n(w)| \leq |w - u_n(w)| + |g_n(w)| < \varepsilon/2 + 1 - \varepsilon = 1 - \varepsilon/2.
\]

Thus, by Brouwer's fixed point theorem, $w + g_n(w) - u_n(w) = w$ for some $w$ such that $|w| < 1 - \varepsilon/2$. This implies that $g_n(w) = u_n(w)$, and hence $f_n(z') = p_n(z')$, where $z' = z_0 + (z_n - z_0)w$, contrary to the assumption.

2.6. **Definition.** Let $p$ be a point function and $f \in H(G)$. A function $s \in H(G)$ is $f$-closed around $p$, if $s$ is not identically zero and for every point function $t$ dividing $s$ there is a $(p, t)$-quasilinear function $u$ such that $u - f$ divides $s$.

The function $s$ codes a set of points, namely the set where it assumes the value zero. Obviously $s$ could be replaced by any other element of $H(G)$ having exactly the same zeroes, since all conditions imposed on $s$ deal with divisibility in the ring only. Similarly, $p$ codes a point. The intuitive content of the definition of $f$-closed is that whenever $x \in S$, there is $y \in S$ such that $y - z$ is approximately equal to $(x - z)f(z)$, where $z = Z(p)$ and $S$ is the set of zeroes of $s$. The approximation is only meaningful for $x$ lying close enough to $z$.

2.7. **Definition.** (i) Let $p$ be a point function. A unit $f$ is a local pseudoroot at $p$, if for every point function $q$ not dividing $p$ there is a function $s$ which is divisible by $q$ and both $f$-closed and $(1/f)$-closed around $p$.

(ii) A function is a global pseudoroot if it is a local pseudoroot at every point function.

The following lemmas will contain all essential things we need to know about local and global pseudoroots. Basically, a local pseudoroot resembles a root of unity at one point, a global pseudoroot everywhere. The concept of a local pseudoroot is formulated just for convenience.

2.8. **Lemma.** If $f$ is a local pseudoroot at some point function $p$, then $|f(Z(p))| = 1$.

**Proof.** Suppose this is not the case. Denote $Z(p)$ by $z_0$. Clearly $f$ is a local pseudoroot at $p$ iff $1/f$ is. Thus it may be assumed without loss of generality that $|f(z_0)| < 1$. Choose $\varepsilon, \delta > 0$ such that $U_{\delta}(z_0) \subseteq G$ and that $|f(z)| < 1 - \varepsilon$ whenever $z \in U_{\delta}(z_0)$. Let $r_0$ be as in Lemma 2.5 with the particular values we have chosen for $\varepsilon$ and $z_0$. Let $r = \min(r_0, \delta)$. Now choose a point function $q$ such that $Z(q) \in U_r(z_0)$. Assume that $s$ is divisible by $q$ and $f$-closed around $p$. Lemma 2.5 now implies that $z_0$ is a limit point of the set of zeroes of $s$, which is impossible, since $s$ is not identically zero.
2.9. **Corollary.** If a function is a global pseudoroot, then it is a constant whose absolute value is 1.

**Proof.** An analytic function whose absolute value is constant is itself constant.

2.10. **Lemma.** If \( f \) is a root of unity, then \( f \) is a global pseudoroot.

**Proof.** Let \( f \) be a constant with value \( \alpha \) such that \( \alpha^n = 1 \), and let \( p \) and \( q \) be point functions with distinct zero points. Denote \( Z(p) \) and \( Z(q) \) by \( z_0 \) and \( z_1 \), respectively, and write \( w_k \) for \( z_0 + \alpha^k(z_1 - z_0) \). We choose

\[
s = \prod_{k=0}^{n-1} P_{w_k}.
\]

Obviously \( s \) is divisible by \( q \). Moreover, if \( t \) is a point function dividing \( s \), there must be some \( k \in \{0, \ldots, n-1\} \) such that \( Z(t) = w_k \). Now the linear function \( u = L_{z_0,w_k} \) is \( (z_0, w_k) \)-quasilinear, and a simple calculation shows that \( P_{w_{k+1}} \) is divisible by \( u - f \); the quotient is, in fact, the constant with value \( 1/\alpha^k(z_1 - z_0) \). Similarly, \( u - 1/f \) divides \( P_{w_{k-1}} \), and hence it divides \( s \). Thus \( s \) is \( f \)-closed and \( 1/f \)-closed around \( p \), and \( q \) divides \( s \).

2.11. **Definition.** A function \( f \) is a basic constant if there are global pseudoroots \( f_1, f_2, f_3, \) and \( f_4 \) such that \( f_3 \neq f_4 \) and \( f = (f_1 - f_2)/(f_3 - f_4) \).

2.12. **Theorem.** A function \( f \) is constant if and only if \( f - g \) is either zero or a unit for all basic constants \( g \).

**Proof.** Clearly every basic constant is a constant function, and the difference of two constants is either zero or a unit. Hence if \( f \) is a constant, then the second condition is also true. Assume then that \( f \) is not a constant function. Then its range \( f[G] \) is a nonempty open set of complex numbers. By Lemma 2.10, the possible values of global pseudoroots form a dense subset of the unit circle. Hence, the values of the basic constants form a dense subset of the whole complex plane. Therefore there is a basic constant \( g \) whose value equals that of \( f \) at some point \( z \in G \). Then \( f - g \) is neither zero nor a unit.

It is worth noting that the definition of the constants could be given easily once we had a formula defining a set of constants whose values are dense in the unit circle. In fact, many other infinite definable sets of constants could be used in a similar way—a superset of the integers, a dense subset of the real interval \([0, 1]\), or the like.

**Acknowledgments**

My warmest thanks are due to Mr. E. Saksman for reading several versions of the manuscript and suggesting numerous improvements. I also thank Professor M. Vuorinen for his kind help.


References


Department of Mathematics, P.O. Box 4 (Hallituskatu 15), FIN-00014 University of Helsinki, Finland

E-mail address: taneli.huuskonen@helsinki.fi