# REAL FUNCTIONS WITH ALGEBRAIC ADDITION THEOREMS*

BY

J. F. RITT

It is a well known theorem of Weierstrass that if an analytic function $\phi(x)$ satisfies, for all values of $x$ and $y$, an algebraic relation

$$(1) \qquad G[\phi(x),\ \phi(y),\ \phi(x+y)] = 0,$$

the function $\phi(x)$ is either an algebraic function, or an algebraic function of $e^{\mu x}$ ($\mu$ constant), or an algebraic function of $\wp(x)$. Weierstrass never published a proof of this theorem. Proofs were given by Phragmen and by Koebe.[†]

If one restricts oneself to the real domain, it is natural to require less than analyticity of $\phi(x)$, for instance, mere continuity. In that case, $\phi(x)$ no longer has to be a function of one of the three types just mentioned. For example, if $\phi(x)$ is one algebraic function of $\wp(x)$ in an interval $(a, b)$, and another algebraic function of the same $\wp(x)$ in $(b, c)$, $\phi(x)$ will have an algebraic addition theorem.

Even if the polynomial $G$ in (1) is irreducible, $\phi(x)$ need not be analytic. For instance, the addition theorem of $\sin x$ is

$$(2) \qquad (w^2 - u^2 - v^2 + 2u^2v^2)^2 - 4u^2v^2(1 - u^2)(1 - v^2) = 0$$

where

$$u = \sin x, \quad v = \sin y, \quad w = \sin(x + y).$$

The first member of (2) is irreducible.[‡] Now, because (2) involves $u$, $v$ and $w$ only in even powers, (2) also furnishes an addition theorem for the continuous but non-analytic function $|\sin x|$.

In the present paper, we consider functions $\phi(x)$, continuous in an interval $(0, a>0)$, and possessing an addition theorem (1) with $G$ irreducible. We obtain, relative to $\phi(x)$, the three theorems which follow.

---

† Phragmen, Acta Mathematica, vol. 7 (1885), p. 33; Koebe, *Mathematische Abhandlungen H. A. Schwarz zu seinem fünfzigjährigen Doctorjubiläum gewidmet*, p. 192.

‡ This follows from the fact that to given values of $\sin x$ and $\sin y$, there correspond four values of $\sin(x+y)$.

**THEOREM I.** *Let $\phi(x)$, real and continuous in $(0, a)$ $(a>0)$, satisfy for $0 \leqq x \leqq a$, $0 \leqq y \leqq a - x$, a relation*

$$G[\phi(x), \phi(y), \phi(x + y)] = 0,$$

*in which $G$ is an irreducible polynomial in its three arguments. Then $\phi(x)$ is piecewise analytic in $(0, a)$, that is, $(0, a)$ can be divided into a finite number of intervals in the interior of each of which $\phi(x)$ coincides with an analytic function.*

Let the intervals in which $\phi(x)$ is analytic be

$$(x_0, x_1), \quad (x_1, x_2), \cdots, (x_{n-1}, x_n) \qquad\qquad (x_0 = 0, x_n = a).$$

Let $\phi_i(x)$ be used to represent $\phi(x)$ in $(x_{i-1}, x_i)$. With respect to $\phi_2, \cdots, \phi_n$, we prove, with the help of a method due to Koebe, the following two theorems:

**THEOREM II.** *For every $i \leqq n$, there exists a constant $\alpha_i$ such that*

(3) $$\phi_i(x) = \phi_1(x + \alpha_i).$$

**THEOREM III.** *If there is a point in $(x_{k-1}, x_k)$ which is the sum of a point in $(x_{i-1}, x_i)$ and a point in $(x_{j-1}, x_j)$, then*

$$\phi_i(x) = \phi_1(\epsilon x + d_1), \quad \phi_j(x) = \phi_1(\epsilon x + d_2), \quad \phi_k(x) = \phi(\epsilon x + d_1 + d_2)$$

*where $d_1$ and $d_2$ are constants, and where $\epsilon = \pm 1$ unless $\phi_1(x)$ is an algebraic function of $\wp(x)$ with $g_3 = 0$ or $g_2 = 0$.\* If $\phi_1(x)$ is algebraic in $\wp(x)$ with $g_3 = 0$, $\epsilon$ may be $\pm(-1)^{1/2}$, whereas, if $g_2 = 0$, $\epsilon$ may be a sixth root of unity.*

Given a $\phi_1(x)$ with an addition theorem, the determination of all functions $\phi(x)$ continuous in $(0, a)$ which have the same addition theorem as $\phi_1(x)$ depends on the determination of all possibilities for the constants $\epsilon$ and $d$ in Theorem III.

### PROOF OF THEOREM I

1. Let $\phi(x)$ satisfy the hypothesis of Theorem I. If $\phi(x)$ is a constant, Theorem I is certainly true for it. Suppose that $\phi(x)$ is not a constant. We shall show first that there exists an interval in which $\phi(x)$ assumes no value twice, and, after this, that there is an interval in which $\phi(x)$ is analytic. It will then be easy to complete the proof of our theorem.

2. We shall study the circumstances under which $\phi(x)$ assumes equal values at two points, $x$ and $x+y$. If, in $G$, we put $\phi(x+y) = \phi(x)$, we obtain

---

\* The quantities $\epsilon$, $d_1$, $d_2$ depend on $i, j, k$.

a polynomial $H[\phi(x),\ \phi(y)]$. Here $H$ is not identically zero. If it were, $G$ would be divisible by $\phi(x+y)-\phi(x)$. Then either $G$ would be reducible, or $G$ would be a constant times $\phi(x+y)-\phi(x)$. In the latter case, $\phi(x)$ would be a constant.

Because $H=0$ when $\phi(y)=\phi(0)$, $H$ is divisible by $\phi(y)-\phi(0)$. Let

$$(4) \qquad\qquad H = [\phi(y) - \phi(0)]^n K[\phi(x),\ \phi(y)],$$

where $K$ is a polynomial not divisible by $\phi(y)-\phi(0)$. Then the equation $K[\phi(x),\ \phi(0)]=0$ has only a finite number of solutions for $\phi(x)$.

We can thus find a $c$ in $(0,\ a)$ such that $K[\phi(c),\ \phi(0)]\neq 0$. Then if $x$ is close to $c$, and $y$ close to 0, $K[\phi(x),\ \phi(y)]$ is not 0.

Suppose that there does not exist an infinite sequence of $y$'s, decreasing towards 0, for each of which $\phi(y)=\phi(0)$. We see from (4), in that case, that for $x$ close to $c$ the equation $H=0$ will not be satisfied by a $y$ close to, and distinct from 0. We have thus an interval about $c$ in which $\phi(x)$ assumes no value more than once.

We shall show that, because $\phi(x)$ is not a constant, there cannot be a sequence of $y$'s decreasing towards 0 for each of which $\phi(y)=\phi(0)$. Suppose that there does exist such a sequence of positive values of $y$,

$$(5) \qquad\qquad \alpha_1,\ \alpha_2,\ \cdots,\ \alpha_n,\ \cdots$$

with $\phi(\alpha_n)=\phi(0)$ $(n=1,\ 2,\ \cdots)$, and $\lim \alpha_n=0$. As $G$ is irreducible, and not equal to a constant times $\phi(x)-\phi(0)$, it cannot vanish identically when $\phi(x)$ is replaced by $\phi(\alpha_n)=\phi(0)$. We consider the polynomial

$$(6) \qquad\qquad G[\phi(\alpha_n),\ \phi(y),\ \phi(x + y)].$$

Because $\phi(x)$ assumes an infinite number of values in $(0,\ a)$, there is an $h=\phi(\beta)\neq\phi(a)$ $(0\leq\beta<a)$ such that (6) is not divisible by $\phi(y)-h$. Then the equation

$$(7) \qquad\qquad G[\phi(\alpha_n),\ h,\ \phi(x + y)] = 0$$

is satisfied by only a finite number of values of $\phi(x+y)$. We shall suppose that $\beta$ is the greatest value of $x$ less than $a$ for which $\phi(x)=h$. Then (7), when solved for $\phi(x+y)$, will give all values which $\phi(x)$ can possibly assume at the points $\alpha_n+\beta$. As there are only a limited number of solutions of (7), there is an infinite set of points among the points $\alpha_n+\beta$ at which $\phi(x)$ assumes only a single value. Because $\phi(x)$ is continuous, this single value must be $\phi(\beta)$. This contradiction of the assumption that $\beta$ is a maximum proves that no sequence (5) exists. This, as we have seen, implies the

existence of a point $c$ in the neighborhood of which $\phi(x)$ assumes no value more than once.

3. Thus, in some interval about $c$, $\phi(x)$ is monotone. A well known theorem of Lebesgue tells us that $\phi(x)$ has a finite derivative almost everywhere in this interval.

This information secured, we proceed to show the existence of an interval in which $\phi(x)$ is analytic.

4. If the equation $G=0$ is of degree $n$ in $\phi(x+y)$, it determines $\phi(x+y)$ as an algebraic function of $n$ branches in $\phi(x)$ and $\phi(y)$. There must be a $\gamma$ in the interval about $c$ found above, and a $\delta$ close to 0, such that the $n$ branches of $\phi(x+y)$ are analytic, and distinct from one another, at $[\phi(\gamma), \phi(\delta)]$. This is because $\phi(x)$ is not constant in the neighborhoods of $c$ and 0. Considered as functions of $x$ and $y$, the $n$ branches will be continuous functions, distinct from one another, for $x$ near $\gamma$ and $y$ near $\delta$. Consequently $\phi(x+y)$ is identical with one of these functions for $x$ near $\gamma$ and $y$ near $\delta$.

We may evidently suppose $\gamma$ to have been so chosen that $\phi(x)$ has a finite derivative at $x=\gamma$. This understood, we write, for $x$ and $y$ close to $\gamma$ and $\delta$ respectively,

$$(8) \qquad \phi(x + y) = F[\phi(x), \phi(y)],$$

where $F$ is analytic (and algebraic) in its arguments. Let $y$ be given any value close to $\delta$, and let $x$ vary past $\gamma$. Then the second member of (8) becomes a function of $x$ with a finite derivative for $x=\gamma$. We infer from (8) that $\phi(x)$ has a finite derivative for every $x$ close to $\gamma+\delta$, or, what is the same, that $\phi(\gamma+y)$ has a finite derivative with respect to $y$ for $y$ close to $\delta$.

For $x=\gamma$, and $y$ near $\delta$, we have, from (8),

$$(9) \qquad \frac{d}{dx}\, \phi(x + y) = \frac{d}{dy}\, \phi(\gamma + y) = U[\phi(\gamma), \phi(y)]\phi'(\gamma),$$

where $U$ is the partial derivative of $F$ with respect to $\phi(x)$. Also

$$(10) \qquad \phi(\gamma + y) = F[\phi(\gamma), \phi(y)].$$

We suppose now that $\delta$ is taken so small that $\gamma+\delta$ lies in the interval about $c$ in which $\phi(x)$ is monotone. Then, as $y$ varies past $\delta$, $\phi(\gamma+y)$ will increase or decrease monotonically. Hence, there is a value $\eta$ of $y$, close to $\delta$, at which the derivative of $\phi(\gamma+y)$ with respect to $y$ is not 0. Referring to (10), we see that the derivative of $F[\phi(\gamma), \phi(y)]$ with respect to $\phi(y)$ cannot be 0 at $\phi(\eta)$. We can thus solve for $\phi(y)$ in (10), expressing it as an

algebraic function of $\phi(\gamma+y)$, analytic for $\phi(\gamma+y)$ close to $\phi(\gamma+\eta)$. If we substitute this expression for $\phi(y)$ into (9), we have

$$(11) \qquad \frac{d}{dy}\phi(\gamma+y) = V[\phi(\gamma+y)],$$

where $V$ is analytic for $\phi(\gamma+y)$ close to $\phi(\gamma+\eta)$.

The existence theorems for differential equations inform us that, because $\phi(\gamma+y)$ satisfies (11), it is analytic for $y=\eta$. Thus $\phi(x)$ is analytic for $x$ close to $\gamma+\eta$.

5. We shall now complete the proof of Theorem I. Consider an interval $(\alpha, \beta)$, interior to $(0, a)$, in which $\phi(x)$ is monotone and analytic. The existence of such an interval was just proved. Let $\gamma$ and $\delta$ be any numbers in $(0, a)$ such that $\delta-\gamma=\beta-\alpha$. We shall show that $(\gamma, \delta)$ can be divided into a finite number of intervals in each of which $\phi(x)$ coincides with some analytic function. This will, of course, settle the proof of Theorem I.

Suppose first that $\gamma \leqq \alpha$. For $x$ in $(\gamma, \delta)$ we have

$$(12) \qquad G[\phi(x), \phi(\alpha-\gamma), \phi(u)] = 0,$$

where $u=x+\alpha-\gamma$. As $x$ ranges over $(\gamma, \delta)$, $u$ ranges over $(\alpha, \beta)$.

Because $G$ is irreducible, the first member of (12) is not identically zero. Hence, except for a finite number of values of $\phi(u)$, for which (12) may hold for every $\phi(x)$, (12) determines $\phi(x)$ as one of several algebraic functions of $\phi(u)$. Let $\phi(\alpha)=s$, $\phi(\beta)=t$. We shall study $\phi(x)$ as a function of $\phi(u)$ as $\phi(u)$ varies from $s$ to $t$, that is, since $\phi(u)$ is monotone in $(\alpha, \beta)$, as $u$ increases from $\alpha$ to $\beta$.

To fix our ideas, suppose that $s<t$. Each of the algebraic functions of $\phi(u)$ defined by (12) has only a finite number of singular points. These singular points divide $(s, t)$ into a finite number of intervals in which each branch of each algebraic function is analytic. Consider one of these intervals. In it, two branches of the same or of different functions can be equal at only a finite number of points. Hence we can divide $(s, t)$ into a finite number of intervals in each of which (12) defines a certain number of analytic functions of $\phi(u)$, no two functions being equal at any point interior to an interval.

Let $(s', t')$ be one of these intervals, and let $(\alpha', \beta')$ be the corresponding interval for $u$. As $u$ varies over $(\alpha', \beta')$, (12) gives a certain number of functions of $u$, analytic, and nowhere equal to one another, in the interior of $(\alpha', \beta')$. One of these functions must coincide with $\phi(x)$ for $x$ in $(\gamma+\alpha'-\alpha, \gamma+\beta'-\alpha)$.

Thus $\phi(x)$ is analytic for $x$ interior to $(\gamma+\alpha'-\alpha,\ \gamma+\beta'-\alpha)$, which is any one of a certain finite number of intervals into which our procedure has divided $(\gamma,\ \delta)$.

A similar proof holds when $\gamma>\alpha$. Theorem I is established.

### Proofs of Theorems II and III

6. Let the points $x_0=0,\ x_1>x_0,\ \cdots,\ x_n=a$ divide $(0,\ a)$ into $n$ intervals, in the interior of each of which $\phi(x)$ is analytic. Let the analytic function with which $\phi(x)$ coincides in $(x_{i-1},\ x_i)$ be represented by $\phi_i(x)$. Because $\phi_1(x)$ has the algebraic addition theorem

$$(13) \qquad G[\phi_1(x),\ \phi_1(y),\ \phi_1(x+y)] = 0,$$

$\phi_1(x)$ is an algebraic function, or an algebraic function of $e^{\mu x}$, or an algebraic function of $\wp(x)$. As each $\phi_i(x)$ is algebraically related to $\phi_1(x)$, each $\phi_i(x)$ is also of one of the above types, and, indeed, it is of the same type as $\phi_1(x)$.

We are going to show that Theorems II and III hold for the functions $\phi_i(x)$. The method we shall use is one which appears in Koebe's paper, where it serves a different purpose.

7. Let $\xi$, $\eta$, and $\xi+\eta$ be three points situated respectively in $(x_{i-1},\ x_i)$, $(x_{j-1},\ x_j)$, $(x_{k-1},\ x_k)$, at each of which $\phi_1(x)$ has at least one branch which is analytic. Whether or not the branches are real is immaterial. It is proved by Koebe* that no matter which branches are chosen at $\xi$, $\eta$ and $\xi+\eta$, equation (13) will be satisfied for $x$ close to $\xi$ and $y$ close to $\eta$.

As $G$ is irreducible, and as the equations (13) and

$$G[\phi_i(x),\ \phi_j(y),\ \phi_k(x+y)] = 0$$

are satisfied for $x$ near $\xi$ and $y$ near $\eta$, it is possible to vary $\phi_1(x)$ in (13) continuously from $\phi_1(\xi)$ to $\phi_i(\xi)$ and $\phi_1(y)$ from $\phi_1(\eta)$ to $\phi_j(\eta)$, in such a way that $\phi_1(x+y)$ varies from $\phi_1(\xi+\eta)$ to $\phi_k(\xi+\eta)$. These variations will occur while $x$ varies from $\xi$ to a point $\xi_1$, $y$ from $\eta$ to a point $\eta_1$, and hence, by (13), as $x+y$ varies from $\xi+\eta$ to $\xi_1+\eta_1$. The points $\xi_1$, $\eta_1$ need not be real.

We have

$$\phi_1(\xi_1) = \phi_i(\xi), \quad \phi_1(\eta_1) = \phi_j(\eta), \quad \phi_1(\xi_1+\eta_1) = \phi_k(\xi+\eta).$$

Changing $\xi$ and $\eta$ slightly if necessary, we may suppose that $\phi_1(x)$ is analytic, with a non-vanishing derivative at $\xi_1$, $\eta_1$, $\xi_1+\eta_1$, and that the equation

$$(14) \qquad G[\phi_1(\xi_1),\ \phi_1(\eta_1),\ \phi_1(x+y)] = 0$$

for $\phi_1(x+y)$ has no equal roots.

* Loc. cit., §3.

We associate with every $x$ close to $\xi$ that $x_1$ close to $\xi_1$ for which $\phi_1(x_1) = \phi_i(x)$. With every $y$ close to $\eta$ we associate that $y_1$ close to $\eta_1$ for which $\phi_1(y_1) = \phi_j(y)$. Thus, because (14) has no equal roots, every $x+y$ close to $\xi + \eta$ is associated with an $x_1 + y_1$ close to $\xi_1 + \eta_1$ such that $\phi_1(x_1 + y_1) = \phi_k(x+y)$. Let

$$x_1 = u(x), \quad y_1 = v(y), \quad x_1 + y_1 = w(x + y).$$

Here $u$, $v$ and $w$ are analytic. Then

(15)
$$w(x + y) = u(x) + v(y).$$

It is easy to prove, differentiating (15), that

$$u(x) = cx + d_1, \quad v(y) = cy + d_2,$$

where the coefficients are constants. Hence

(16) $\quad \phi_i(x) = \phi_1(cx + d_1), \quad \phi_j(x) = \phi_1(cx + d_2), \quad \phi_k(x) = \phi_1(cx + d_1 + d_2).$

Consider any $k \leq n$. Let a $\xi$ be chosen in $(x_0, x_1)$, and then let an $\eta$ be taken in such a way that $\xi + \eta$ lies in $(x_{k-1}, x_k)$. Then, in (16), $i = 1$, and the first and third relations give

(17)
$$\phi_k(x) = \phi_1\left(x + \frac{d_2}{c}\right).$$

This proves Theorem II.

8. We are going to prove, in this section and in §9, that $c$ in (16) is a root of unity, which is $\pm 1$ unless the functions $\phi_i(x)$ are algebraic functions of $\wp(x)$ with $g_3 = 0$ or $g_2 = 0$. In the exceptional cases, $c$ may be $\pm i$ if $g_3 = 0$, whereas, if $g_2 = 0$, $c$ may be a sixth root of unity.* Theorem III will be established as soon as these facts are proved.

From (3) and (16) we find, replacing $x$ by $x - \alpha_j$, that

(18)
$$\phi_1(x) = \phi_1(cx - c\alpha_j + d_2).$$

If, supposing that $c \neq 1$, we put $\beta = (-c\alpha_j + d_2)/(1-c)$, (18) becomes

(19)
$$\phi_1(x) = \phi_1[c(x - \beta) + \beta].$$

This shows that $c$ is a root of unity, for, if it were not, $\phi_1(x)$ would assume values infinitely often in the neighborhood of $\beta$. For roots of unity of any order, there are *algebraic* functions which satisfy relations like (19). But if $\phi_1(x)$ is *periodic*, $c = \pm 1$ except in the two cases mentioned above.†

---

* Similar details occur in Koebe's paper. What is new here is the proof that $c = \pm 1$ when $\phi_1(x)$ is algebraic.

† Koebe, loc. cit., or Ritt, these Transactions, vol. 23 (1922), p. 20.

9. We shall show that when the functions $\phi_i(x)$ are algebraic, $c$ in (16) must be $\pm 1$, except if $\beta$ is 0. In the latter case, by (19), (16) becomes

$$\phi_i(x) = \phi_1\left(x + \frac{d_1}{c}\right), \quad \phi_j(x) = \phi_1\left(x + \frac{d_2}{c}\right), \quad \phi_k(x) = \phi_1\left(x + \frac{d_1 + d_2}{c}\right).$$

Thus, when the functions $\phi_i(x)$ are algebraic, $\pm 1$ will always serve as values of $c$ in (16).

We begin by observing that, when $\phi_1(x)$ is algebraic, every $\alpha_i$ in (3) must be real. Suppose that an $\alpha_i$ is not real. Then $\phi_1(x) = \phi_i(x - \alpha_i)$ assumes real values on the straight segment joining $x_{i-1} - \alpha_i$ to $x_i - \alpha_i$. This segment is parallel to the real axis. But $\phi_1(x)$ assumes real values in $(0, x_1)$. Hence it assumes conjugate values at conjugate points. Thus, given any point on the above mentioned segment, $\phi_1(x)$ assumes the same real value at that point and at its conjugate. The difference between the two conjugate points stays equal, as the first point varies over its segment, to twice the imaginary part of $\alpha_i$. This constant difference must be a period of $\phi_1(x)$. But $\phi_1(x)$, being algebraic, has no periods. This proves that $\alpha_i$ is real.

From (16) and (3), we find that

$$\phi_k(x) = \phi_i\left(x + \frac{d_2}{c}\right) = \phi_1\left(x + \alpha_i + \frac{d_2}{c}\right).$$

Thus we may use $\alpha_i + d_2/c$ for $\alpha_k$, so that $d_2$ is a real multiple of $c$.

Suppose that $c \neq 1$. Referring to the value of $\beta$, we see that, if $c$ is not $-1$, $\beta$ will not be real unless $\beta = 0$. Suppose that $c \neq -1$ and that $\beta \neq 0$. Let $\bar{\beta}$ represent the conjugate of $\beta$. Because $\phi_1(x)$ assumes conjugate values at conjugate points, there exists, corresponding to (19), a relation

(20)                         $\phi_1(x) = \phi_1[c(x - \bar{\beta}) + \bar{\beta}].$

Equating the second members of (19) and (20) we see that

$$(c - 1)(\beta - \bar{\beta}),$$

which does not vanish, is a period of $\phi_1(x)$. Thus, when $c$ is not $-1$, we have $\beta = 0$. This shows that $\phi_i(x)$, $\phi_j(x)$ and $\phi_k(x)$ satisfy a relation (16) with $c = \pm 1$, and completes the proof of Theorem III.

COLUMBIA UNIVERSITY,
   NEW YORK, N. Y.