Weierstrass' proof of the Lindemann-Weierstrass theorem (Part 2 of 3)

In part 2 of his 1885 paper, Weierstrass reproduced Lindemann's proof that, if \( z \) is algebraic, \( e^z + 1 \) is not zero.

The proposition is true if \( z \) is rational, so we need only consider the case where \( z \) has an imaginary component.

If \( z_i \) is algebraic then it is the root of an integral polynomial of degree \( r \) that has distinct roots \( z_1, z_2, \ldots, z_r \) (Lemma 1). We can assume that \( r > 1 \) because the complex conjugate of \( z_i \) is also a root. Now

\[
e^{z_i} + 1
\]

is non-zero for every algebraic \( z_i \) if and only if

\[
(1 + e^{z_1})(1 + e^{z_2}) \ldots (1 + e^{z_r}) = P
\]

is non-zero for every algebraic \( z_i \). It is this latter proposition that we prove.

We can write

\[
P = \sum_{k=0}^{p-1} e^{\xi_k} \quad \text{................................................................. (1)}
\]

where

\[
\xi_k = \epsilon_1 z_1 + \epsilon_2 z_2 + \ldots + \epsilon_r z_r ; \quad \epsilon = 0 \text{ or } 1
\]

The number of algebraically distinct \( \xi_k \) is \( p = 2^r \).

Because the \( \xi \) have degree one in \( z_1, z_2, \ldots, z_r \) of degree 1 (see Lemma 2 for definition) then if \( A \) is the leading coefficient of the integral polynomial of which the \( z_k \) are roots, by Corollary 2 to that Lemma

\[
A^p \prod_{\mu=0}^{p-1} (\xi - \xi_\mu) \quad \text{................................................................. (2)}
\]

is an integral polynomial.

Let \( n + 1 \) be the number of numerically distinct \( \xi \). We denote these quantities as \( \zeta \). Because \( \zeta_k \) include \( z_1, z_2, \ldots, z_r, n + 1 \geq 3 \). Furthermore since the \( \zeta_k \) are roots of the polynomial (2), they are also (Lemma 1) distinct roots of an integral polynomial

\[
f(z) = a_0 z^{n+1} + a_1 z^n + \ldots + a_n z + a_{n+1} \quad \text{................................. (3)}
\]

with \( a_0 > 0 \) (\( a_{n+1} \) is zero in this case since \( \zeta \) can take the value zero).

By part (i) of the Lemma described in Part 1, there exists a system of polynomials \( g_0(\zeta), \ldots, g_n(\zeta) \) of degree not greater than \( n \) in \( \zeta \), and with integer coefficients, such that (i) each of the differences

\[
g_\nu(\zeta_\lambda) e^{\zeta_\lambda} - g_\nu(\zeta_\nu) e^{\zeta_\nu}
\]

(where \( \nu, \lambda \) can take any of the values \( 0, 1, \ldots, n \)) can be made arbitrarily small in
absolute value, and (ii) the determinant whose elements are \( g_\nu(\zeta_\lambda) \) is non-zero. Thus

\[ -\epsilon_{\nu,\lambda} \delta < g_\nu(\zeta_0)e^{\zeta_\lambda} - g_\nu(\zeta_\lambda)e^{\zeta_0} < \epsilon_{\nu,\lambda} \delta \quad ; \quad \delta, \epsilon > 0 \]

Now we sum these inequalities over all the values taken by the \( \zeta_\mu \), including multiple values, and letting \( \zeta_0 = 0 \)

\[ a_0^n g_\nu(0) \sum_{\mu=0}^{p-1} e^{\xi_\mu} - a_0^n \sum_{\mu=0}^{p-1} g_\nu(\xi_\mu) < \delta a_0^n \sum_{\mu=0}^{p-1} \epsilon_{\nu,\mu} (v = 0, \ldots, n) \]  

where \( \delta \) can be made arbitrarily small. Since a permutation of the \( z_k \) results in a permutation of the subscripts \( \mu \) the second sum on the LHS is a symmetric polynomial of the \( z_k \) and by Corollary 2 to Lemma 2 is an integer.

Now if the numerical value of \( \zeta_\mu \) occurs \( N_\mu \) times in the set of \( \zeta_k \) then

\[ \sum_{\mu=0}^{p-1} a_0^n g_\nu(\xi_\mu) = a_0^n \sum_{\lambda=0}^{n} N_\lambda g_\nu(\zeta_\lambda) \]

The sum on the RHS cannot be zero for every \( v \), since none of the \( N_\lambda \) is zero, and the determinant whose elements are \( g_\nu(\zeta_\lambda) \) is not zero, by (ii) of the Lemma of Part 1. Thus for at least one \( v \), by making \( \delta \) sufficiently small, the RHS of (4) can be made arbitrarily close to some positive integer, which would not be possible if \( P \) were zero. Therefore by (1), the result is proved.

**Lemma 1:** If \( z_1, \ldots, z_n \) are distinct roots of an integral polynomial of degree > \( n \), then they are the roots of an integral polynomial of degree \( n \). One such polynomial has a minimum positive leading coefficient and this polynomial is unique.

**Proof:** By the Fundamental Theorem of Algebra, if \( f \) has degree \( m \) in \( z \), it has \( m \) roots. Consequently

\[ f(z) = A (z - z_1)^\mu_1 \ldots (z - z_n)^\mu_n \]

where \( A \) is an integer and \( \mu_1 + \ldots + \mu_n = m \). Say \( z_k \) is a root of multiplicity \( \geq 2 \) of the integral polynomial \( f(z) \). Then

\[ f(z) = (z - z_k)^2 g(z) \]

\[ f'(z) = 2(z - z_k)g(z) - (z - z_k)^2 g'(z) \]

so \( z_k \) is a root of \( f'(z) \). \( f'(z) \) is clearly an integral polynomial. We can repeat this process until all instances of multiple roots are eliminated, in a finite number of steps.

If \( z_1, \ldots, z_n \) are distinct roots of an integral polynomial of degree \( n \) and this polynomial has a negative leading coefficient, then multiplication by -1 produces a polynomial with a positive leading coefficient. Let \( az^n + a_{n-1}z^{n-1} + \ldots + a_0 \) be the polynomial of this type such that \( a \) is least. If \( az^n + b_{n-1}z^{n-1} + \ldots + b_0 \) is another such polynomial then the \( z_1, \ldots, z_n \) are also roots of \((a_{n-1} - b_{n-1})z^{n-1} + \ldots + (a_0 - b_0)\) in contradiction to the Fundamental Theorem of Algebra.
Lemma 2 (a version of the Fundamental Theorem of Symmetric Functions): if $f(z_1, \ldots, z_n)$ is a symmetric polynomial with integer coefficients and of degree $\mu$ it can be expressed as a polynomial $\varphi(s_1, \ldots, s_n)$ with integer coefficients and of degree $\mu$ in the basic symmetric functions $s_1, \ldots, s_n$.

Proof: Suppose $f(z_1, \ldots, z_n)$ is a symmetric polynomial in $z_1, \ldots, z_n$ with integer coefficients. (We call $f(z_1, \ldots, z_n)$ symmetric if $f(z_1, \ldots, z_n) = \pi(f(z_1, \ldots, z_n))$ where $\pi_i$ is any of the $n!$ permutations of the subscripts of the variables). If $\mu$ is the greatest value taken by $\mu_1 + \ldots + \mu_n$ then we call $\mu$ the degree of $f$. It is not hard to see that every $\pi_1(z_1^{\mu_1}z_2^{\mu_2} \ldots z_n^{\mu_n})$ must have the same coefficient.

We introduce a new indeterminate variable $Z$ and define

$$P(Z) = (Z-z_1)(Z-z_2)\ldots(Z-z_n)$$

then

$$P(Z) = Z^n - s_1 Z^{n-1} + s_2 Z^{n-2} - \ldots + (-1)^n s_n$$

We call the $s_k$ the basic symmetric functions and it can be easily shown that

$$s_k = \sum z_{i_1} z_{i_2} \ldots z_{i_k}$$

where the sum ranges over all possible sets of $k$ distinct integers $i_1, \ldots, i_k$ that can be chosen from 1, 2, ..., n. We note that $s_k$ has degree $k$.

We can express $f$ as a sum of terms in

$$c(\mu_1, \ldots, \mu_n) = \sum \pi_1(z_1^{\mu_1}z_2^{\mu_2} \ldots z_n^{\mu_n})$$

ordered such that terms of higher degree precede terms of lower degree and, where terms have equal degree, $\mu_1 \geq \mu_2 \geq \ldots \geq \mu_n$. Suppose

$$c(v_1, \ldots, v_n) = \sum \pi_1(z_1^{v_1}z_2^{v_2} \ldots z_n^{v_n})$$

(1)

is the term of highest degree. Now consider the polynomial

$$s_1^{v_1-v_2} s_2^{v_2-v_3} \ldots s_{n-1}^{v_{n-1}-v_n} s_n^{v_n}$$

(2)

If we expand this as a polynomial in $z_1, \ldots, z_n$ we see that it contains a term

$$(z_1^{v_1-v_2})(z_1 z_2)^{v_2-v_3} \ldots (z_1 z_2 \ldots z_{n-1})^{v_{n-1}-v_n}(z_1 z_2 \ldots z_n)^{v_n} = z_1^{v_1} z_2^{v_2} \ldots z_n^{v_n}$$

exactly once. Since (2) is symmetrical, it must contain the expression (1) exactly once. Therefore if we subtract (2), multiplied by an appropriate integer, from $f(z_1, \ldots, z_n)$, we obtain a polynomial whose leading term has lower degree than (1). Since both $f$ and (2) are symmetric functions, this polynomial must be also. If we repeat this process we must obtain after a finite number of steps a polynomial whose term of highest degree is that for which $\mu_1 = \mu_2 = \ldots = \mu_n = 0$. Thus we have an expression of the form

$$f - f_1 - \ldots - f_{j-1} = f_j = \text{a constant}$$

where the $f_k$ are polynomials in the basic symmetric functions. This proves the result.
Corollary 1: if \( f(z_1, \ldots, z_n) \) is a symmetric polynomial with integer coefficients and of degree \( \mu \) and if these variables take specific values that are distinct roots of an integral polynomial of degree \( n \) with leading coefficient \( a \) then \( a^\nu f(z_1, \ldots, z_n) \) is an integer.

Proof: if \( z_1, \ldots, z_n \) are distinct roots of an integral polynomial of degree \( n \) with leading coefficient \( a \) then the polynomial must be

\[
a(z - z_1)(z - z_n) = az^n - a_s(z_1, \ldots, z_n)z^{n-1} + \ldots + (-1)^n a_s(z_1, \ldots, z_n)
\]

Therefore \( a_s \) is an integer.

By the Lemma, \( f \) can be expressed as a polynomial in the basic symmetric functions \( s_j \) whose term of highest degree is a sum of terms of the form \( k s_1^{\nu_1} \ldots s_n^{\nu_n} \) where \( k \) is some integer and \( \nu_1 + \ldots + \nu_n = \mu \). If we multiply this sum by \( a^{\nu_1} \ldots a^{\nu_n} = a^\mu \) we obtain

\[
k(a_s)^{\nu_1} \ldots (a_s)^{\nu_n}
\]

which is an integer. Clearly multiplying the terms of lesser degree by the same factor also yields integers.

Corollary 2: Let \( z_1, \ldots, z_n \) be complex variables and let \( \zeta_1, \ldots, \zeta_r \) be integral polynomials over the \( z_1, \ldots, z_n \) of degree \( v \) and such that any permutation of \( z_1, \ldots, z_n \) simply reorders the \( \zeta_1, \ldots, \zeta_r \). Let \( \phi \) be some integral function over the \( \zeta_1, \ldots, \zeta_r \) that is symmetric with respect to those variables and of degree \( \mu \). If \( z_1, \ldots, z_n \) take specific values that are distinct roots of an integral polynomial of degree \( n \) with leading coefficient \( a \) then \( a^\nu \phi(\zeta_1, \ldots, \zeta_r) \) is an integer.

Proof: A re-ordering of \( z_1, \ldots, z_n \) re-orders the \( \zeta \) and leaves \( \phi \) unchanged. Therefore \( \phi \) is a symmetric function of \( z_1, \ldots, z_n \). Now let \( \zeta_1^{\sigma_1} \zeta_2^{\sigma_2} \ldots \zeta_r^{\sigma_r} \) be a term of highest degree in \( \phi \). This term has degree \( v(\sigma_1 + \ldots + \sigma_r) = v\mu \) in \( z_1, \ldots, z_n \). The result follows by Corollary 1.