Cliff Bott cliffbott@hotmail.com 27 Feb 2010 revised 6 August 2013

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

As stated by Weierstrass, the Lindemann-Weierstrass Theorem is as follows: if each of $z_1, \dots, z_n$ is algebraic and distinct, and $N_1, \dots, N_n$ are algebraic, then

$$N_1 e^{z_1} + \dots + N_n e^{z_n}$$

cannot be zero unless all of $N_1, \dots, N_n$ are zero.

In the final part of his 1885 paper, Weierstrass proved the theorem first for $N_1, \dots, N_n$ integers, then for the general case.

I. First we prove the special case where $N_1, \dots, N_n$ are integers. $z_1, \dots, z_n$ are roots of a single integral polynomial, all of whose roots are distinct and which has a positive leading coefficient that we call A (Lemma 1). Let r be the degree of this polynomial. We denote its additional roots as $z_{n+1}, \dots, z_r$.

We begin with the assumption

$$N_1 e^{z_1} + \dots + N_n e^{z_n} = 0 \quad \dots\dots \quad (1)$$

where not all of $N_1, \dots, N_n$ are zero, and show this leads to a contradiction. Let

$$P = \prod_{i=1}^{r!} \pi_i (N_1 e^{z_1} + \dots + N_r e^{z_r}) \quad \dots\dots \quad (2)$$

Here $N_{n+1}, \dots, N_r = 0$ and π runs through all permutations of the symbols $N_1, \dots, N_n$. If we take π_1 to be the identity permutation then the first term in the product is

$$N_1 e^{z_1} + \dots + N_r e^{z_r} = N_1 e^{z_1} + \dots + N_n e^{z_n} = 0$$

so our starting assumption implies that

$$P = 0 \quad \dots\dots \quad (3)$$

We expand the product (2) in the manner suggested by the following array:

$$\begin{array}{llll} N_1 & N_2 & \dots\dots & N_r \\ N_1^{(2)} & N_2^{(2)} & \dots\dots & N_r^{(2)} \\ \dots & \dots & \dots\dots & \dots \\ N_1^{(r!)} & N_2^{(r!)} & \dots\dots & N_r^{(r!)} \\ z_1 & z_2 & \dots\dots & z_r \end{array} \quad \dots\dots \quad (4)$$

in which each row, except for the last, is a distinct permutation of $N_1, \dots, N_r$. Thus

$$P = \sum N_1' N_2' \dots N_{r!}' e^{z_1' + z_2' + \dots + z_r'} \quad \dots\dots \quad (5)$$

where in forming a general term in the sum we have taken an element N_j' from each row j of (4) except the last, and included in the exponent the value z_j' taken from the same column as N_j'. We note that the value of P is not altered by altering the order of the columns.

There are $r^{r!}$ terms in the sum thus formed, although not all the exponents are necessarilyalgebraically, or numerically, distinct. If we collect terms with the same numeric exponent, we have

$$P = \sum_{\lambda=0}^{s} C_\lambda e^{\zeta_\lambda} \quad \dots\dots \quad (6)$$

where s + 1 is the number of numerically distinct exponents and

$$C_\lambda = \sum N_1' N_2' \dots N_{r!}' \quad \dots\dots \quad (7)$$

The sum is over all coefficients of the terms in (5) for which

$$z_1' + z_2' + \dots + z_r' = \zeta_\lambda \quad \dots\dots \quad (8)$$

and each term $N_1'N_2'....N_{r!}'$ must be formed with reference to each individual exponent $z_1' + \dots + z_r'$ in the manner described previously.

Now we show that there is at least one C_λ that is non-zero. We do this by identifying a λ for which the sum in (7) contains exactly one non-zero term.

Since we can re-order the columns in (4) without changing the value of P we can assume at the outset that they are ordered such that if z_j is to the left of z_k then either $Re(z_j) > Re(z_k)$ or, if these quantities are equal, then $Im(z_j) > Im(z_k)$.

We can form the first term in the sum (5) by taking the first non-zero element in each row of (4). There must be one, because by hypothesis not all N_j are zero. The exponent in this term, $z_1' + \dots + z_r'$, identifies a particular λ and so the value of $N_1'...N_{r!}'$ selected is the first term of the sum in (7). If there are no other non-zero elements in the sum (7), then C_λ is non-zero.

Otherwise, let $N_1''...N_{r!}''$ be the coefficient of another non-zero element in (5) and let the corresponding exponent be $z_1'' + \dots + z_r''$. None of the $N_1'', \dots, N_{r!}''$ is to the left of $N_1', \dots, N_{r!}'$ respectively, but at least one is to the right. Thus either $Re(z_1' + \dots + z_r') > Re(z_1'' + \dots + z_r'')$ or if $Re(z_1' + \dots + z_r') = Re(z_1'' + \dots + z_r'')$ then $Im(z_1' + \dots + z_r') > Im(z_1'' + \dots + z_r'')$. In either case $z_1' + \dots + z_r' \neq z_1'' + \dots + z_r''$. Thus the C_λ we first identified contains exactly one non-zero element.

Let us momentarily regard $z_1, .., z_r$ as variables and let $\xi_1, \dots, \xi_w$ ($w = r^r$) be the set of functions $z_{1'} + \dots + z_{r'}$. Any permutation of $z_1, \dots, z_r$ simply re-orders the ξ.

The ξ have degree 1 in $z_1, \dots, z_r$. Now consider Corollary 2 to Lemma 2 of Part 2, namely:

Let $z_1, \dots, z_r$ be complex variables and let $\xi_1, \dots, \xi_w$ be integral polynomials over the $z_1, \dots, z_r$ of degree ν and such that any permutation of $z_1, \dots, z_r$ simply reorders the $\xi_1, \dots, \xi_w$. Let ϕ be some integral function over the $\xi_1, \dots, \xi_w$ that is symmetric with respect to those variables and of degree μ. If $z_1, \dots, z_r$ take specific values that are distinct roots of an integral polynomial of degree r with leading coefficient A then $A^{\nu\mu}\phi(\xi_1, \dots, \xi_w)$ is an integer

Giving $z_1, \dots, z_r$ their specific values we see that

$$A^w \prod_{\mu=1}^{w} (\xi - \xi_\mu)$$

is an integral polynomial, since expanding the product gives terms in ξ whose coefficients are the basic symmetric functions of the ξ_μ of degree no greater than w.

Thus the ξ_μ are roots of an integral polynomial, but not all are necessarily distinct numerically once $z_1, \dots, z_n$ are given their numeric values. We have already denoted the numerically distinct ξ as $\zeta_0, \dots, \zeta_s$ and as a consequence of Lemma 1 below they are distinct roots of an integral polynomial with leading coefficient A^w. We recall the Lemma described in the previous post, namely:

Let $f(\zeta)$ be a polynomial of degree s+1 with integer coefficients and whose roots $\zeta_0, \dots, \zeta_s$ are all distinct. Then there exists a system of polynomials $g_0(\zeta), \dots, g_s(\zeta)$ of *degree not greater than s in ζ, and with integer coefficients, such that (i) each of the differences*

$$g_\nu(\zeta_0)e^{\zeta_\lambda} - g_\nu(\zeta_\lambda)e^{\zeta_0}$$

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

From part (i) we obtain

$$e^{-\zeta_0} g_\nu(\zeta_0) C_\lambda e^{\zeta_\lambda} = C_\lambda g_\nu(\zeta_\lambda) + \epsilon_\nu C_\lambda e^{-\zeta_0} \qquad (\nu = 0, \dots, s)$$

where the ε are expressions that can be made as small as we wish in absolute value. Thus

$$e^{-\zeta_0} g_\nu(\zeta_0) \sum_{\lambda=0}^{s} C_\lambda e^{\zeta_\lambda} = \sum_{\lambda=0}^{s} C_\lambda g_\nu(\zeta_\lambda) + \epsilon_\nu e^{-\zeta_0} \sum_{\lambda=0}^{s} C_\lambda \qquad (\nu = 0, \dots, s) \quad \dots\dots\dots\dots \quad (9)$$

The sum on the left hand side of this expression is by (6) and (3) equal to zero for all ν. Now the expressions

$$\sum_{\lambda=0}^{s} C_\lambda g_\nu(\zeta_\lambda)$$

can be regarded as the product of a matrix and a vector thus:

$$\begin{bmatrix} g_0(\zeta_0) & \cdots & g_0(\zeta_s) \\ \cdots & \cdots & \cdots \\ g_s(\zeta_0) & \cdots & g_s(\zeta_s) \end{bmatrix} \begin{bmatrix} C_0 \\ \cdots \\ C_s \end{bmatrix}$$

Since the determinant of the matrix is non-zero by part (ii) of the Lemma, and at least one of the C_λ is non-zero, the resulting vector has at least one non-zero element. Thus for at least one ν, the sum on the right hand side of (9) is non-zero.

We can expand the sum as follows:

$$\sum_{\lambda=0}^{s} C_\lambda g_\nu(\zeta_\lambda) = \sum N_1' N_2' \dots N_{r!}' g_\nu(z_1' + z_2' + \dots + z_r') = \sum_{i=1}^{w} N_1' N_2' \dots N_{r!}' g_\nu(\xi_i)$$

We show that this is a symmetric function of $z_1, \dots, z_r$. It suffices to show that the transposition of any z_i and z_j leaves the sum

$$\sum N_1' N_2' \dots N_{r!}' (z_1' + z_2' + \dots + z_r')^m = \sum_{i=1}^{w} N_1' N_2' \dots N_{r!}' (\xi_i)^m$$

unchanged, where m is a positive integer. Terms in the sum that contain neither z_i nor z_j are unchanged by the transposition. We refer to the remaining terms as comprising the set S. The sum is unchanged by the transposition if the transposition maps S onto itself. S must map to itself, since terms containing only x_i are mapped to terms containing only x_j and vice-versa, and terms containing both are mapped to terms containing both. The mapping is onto because every term created by the transposition has a unique precursor in S.

The polynomials g have degree s or less, so since the ξ have degree 1 in the z, if the leading coefficient of g_ν is a_0 then by the Corollary quoted above

$$a_0^s \sum_{\lambda=0}^{s} C_\lambda g_\nu(\zeta_\lambda)$$

is an integer (and we have already shown it is non-zero). By taking

$$|\epsilon_\nu| < a_0^{-s} e^{-\zeta_0} \sum_{\lambda=0}^{s} C_\lambda$$

we have generated the desired contradiction since if (9) is multiplied by a_0^s the left hand side is zero but the right hand side is non-zero.

II. Now suppose the $N_1, \dots, N_n$ in (1) are any algebraic numbers whatsoever so long as they are not all zero. As in part I we assume this expression is zero and show this leads to a contradiction.

By Lemma 2 the N_k may all be expressed as

$$N_k = G_k(\eta)$$

where the G_k are integral polynomials and η is algebraic. We assume that the G_k are the integral polynomials of least degree such that N_k can be expressed in this form. We let $\eta^{(2)}, \dots, \eta^{(r)}$ be the other roots of the polynomial of least degree of which η is a root and we denote as A the leading coefficient of this polynomial. (We can make the polynomial unique as described in Lemma 1 of Part 2). If $\eta^{(k)}$ is any given root, then $G_j(\eta^{(k)})$ is zero only if η is zero (Lemma 3). Thus at least one of the $G_j(\eta^{(k)})$ is non-zero.

We form the product:

$$P' = \prod^{r} G_1(\eta^{(k)}) e^{z_1} + \dots + G_n(\eta^{(k)}) e^{z_n} \quad \dots\dots \quad (10)$$

By hypothesis P' = 0.

We expand the product in the manner suggested by the following array:

$$\begin{array}{lllll} G_1(\eta^{(1)}) & G_2(\eta^{(1)}) & \dots\dots & G_n(\eta^{(1)}) & \\ G_1(\eta^{(2)}) & G_2(\eta^{(2)}) & \dots\dots & G_n(\eta^{(2)}) & \\ \dots & \dots & \dots\dots & \dots & \dots\dots \quad (11) \\ G_1(\eta^{(r)}) & G_2(\eta^{(r)}) & \dots\dots & G_n(\eta^{(r)}) & \\ z_1 & z_2 & \dots\dots & z_r & \end{array}$$

The value of P' is not changed by altering the order of the first r rows because this operation simply alters the order of the factors in (10). It is not changed by altering the order of the columns since this simply changes the order of the terms in the sum following the product sign in (10). Thus

$$P' = \sum^{n^r \text{ terms}} G_1'(\eta^{(1)}) G_2'(\eta^{(2)}) \dots G_n'(\eta^{(r)}) e^{z_1' + \dots + z_n'} = \sum \phi(\eta^{(1)}, \dots, \eta^{(r)}) e^{\xi_k} \quad \dots\dots \quad (12)$$

where we denote as $\xi_1, \dots, \xi_w$ the set of functions $\nu_1 z_1 + \dots + \nu_n z_n$ where for the moment the z_k are regarded as variables, the ν are non-negative integers and $\nu_1 + \dots + \nu_n = n$.

The coefficients of the e^{ξ} are symmetric functions of the $\eta^{(k)}$ because if we interchange rows i and j, ξ does not change and the terms that sum to the coefficient of e^{ξ} are simply changed in order, although the operation has interchanged $\eta^{(i)}$ and $\eta^{(j)}$.

Suppose the G_k of largest degree has degree d. Then the coefficients of the e^{ξ} have degree at most nd and so if we multiply P' by A^{nd} we obtain an expression in which the coefficients of the e^{ξ} are all integers. This can be seen by substituting η for z in Corollary 2 quoted in I above.

Now if, as in Part I, we give the z_k their numerical values and denote the numerically distinct ξ as ζ we obtain from (12)

$$A^{nd} P' = \sum_{\lambda=0}^{s} C_\lambda e^{\zeta_\lambda}$$

where the C_λ are integers. Because each of the rows in (11) has at least one non-zero element, and because P' is not altered by changing the order of the columns in (11), we can show in the same manner as in I that the C_λ are not all zero and hence that P' is not zero.

Lemma 1: If $z_1, \ldots, z_n$ are distinct algebraic numbers then they are roots of a single integral polynomial, all of whose roots are distinct.

Proof: If $z_1, \ldots, z_n$ are roots of integral polynomials $P_1(z), \ldots, P_n(z)$ then they are the roots of the single integral polynomial $P_1(z) \ldots P_n(z)$. If this is the case, then by Lemma 1 of Part 2, they are the roots of an integral polynomial with distinct roots.

Lemma 2: if $z_1, \ldots, z_n$ are algebraic, they can be expressed as $z_1 = G_1(\eta), \ldots, z_n = G_n(\eta)$ where the G are polynomials with integer coefficients and η is algebraic.

Proof: the Lemma is a special case of a general theorem in abstract algebra which states that if F is some field and $\alpha, \beta, \ldots$ are finite in number and algebraic over F (that is, are roots of a polynomial whose coefficients are in F) then if $F(\alpha, \beta, \ldots)$ is the field formed by adjoining α, $\beta, \ldots$ to F, $F(\alpha, \beta, \ldots)$ is a simple extension of F (it can be formed by adjoining to F a single element η that is algebraic over F. In the case of the Lemma, the field F is the rational numbers. A straightforward proof can be found in Stewart and Tall "Algebraic Number Theory and Fermat's Last Theorem", Third Edition, as Theorem 2.2

Lemma 3: Let η be the root of an integral polynomial and let η' another root of the integral polynomial of least degree, uniquely defined as in Lemma 1 of Part 2, of which η is a root. Let $G(z)$ be an integral polynomial. Then if $G(\eta) = 0$, $G(\eta') = 0$.

Proof: the Lemma follows by applying to $G(z)$ the process described in Lemma 1 of Part 2. If G does not have a factor $(z - \eta')$, then nor does the polynomial of least degree of which η is a root.