proof of Lindemann-Weierstrass theorem and that $e$ and $\pi$ are transcendental

rm50

2013-03-21 22:58:05

This article provides a proof of the Lindemann-Weierstrass theorem, using a method similar to those used by Ferdinand von Lindemann and Karl Weierstrass. This material is taken from [?] and expanded for clarity.

Before attacking the general case, we first use the same methods to prove two earlier theorems, namely that both $e$ and $\pi$ are transcendental. These proofs introduce the methods to be used in the more general theorem. At the end, we present some trivial but important corollaries.

Both $e$ and $\pi$ were both known to be irrational in the 1700’s (Euler showed the former; Lambert the latter). But $e$ was not shown to be transcendental until 1873 (by Hermite, see [?] and [?]), and Lindemann showed $\pi$ to be transcendental as well in the late 1870’s. He also sketched a proof of the general theorem, which was fleshed out by Weierstrass and Hilbert among others in the late 1800’s.

The following construct is used in all three proofs. Suppose $f(x)$ is a real polynomial, and let

$$I(t) = \int_0^t e^{t-x} f(x) \, dx.$$

Integrating by parts, we get

$$I(t) = (-e^{t-x} f(x)) \bigg|_0^t + \int_0^t e^{t-x} f'(x) \, dx = e^t f(0) - f(t) + \int_0^t e^{t-x} f'(x) \, dx.$$

Continuing, and integrating by parts a total of $m = \deg f$ times, we get

$$I(t) = e^t \sum_{j=0}^m f^{(j)}(0) - \sum_{j=0}^m f^{(j)}(t)$$

where $f^{(j)}(x)$ is the $j$th derivative of $f$. 
If \( f(x) = \sum a_i x^i \), let \( F(x) = \sum |a_i| x^i \); i.e., the polynomial whose coefficients are the absolute values of those for \( f \). Then using trivial bounds on the integrand, we get

\[
|I(t)| \leq \int_0^t |e^{t-x} f(x)| dx \leq |t| e^{t|t|} F(|t|).
\]

We now proceed to prove the theorems. The proofs of all three are similar, although the proof for \( e \) is the easiest. The steps of the proofs are as follows:

- Assume the theorem is false, and write down an equation in exponential form that shows that the number in question is algebraic (for \( \pi \), we will use \( e^{i\pi} = -1 \) to write the equation in that form).
- Define a polynomial or set of polynomials \( f \), and an associated number \( J \) (or a sequence of numbers) that is a linear combination of the values of \( I \) at the exponents in question. The motivation for the choice of \( f \) used in each theorem is not given in Hermite’s proof. An excellent exposition of how these definitions are relevant to the problem is given in [?]. In essence, the ratio of the terms of \( I(t) \) in the proof that \( e \) is transcendental relates to a Padé approximation to \( e^t \); this approximation is better the larger deg \( f \) gets.
- Analyze \( J \) to show that it is integral and nonzero, and derive a lower bound on \( J \).
- Use equation (??) to derive a trivial upper bound on \( J \).
- Note that the upper bound is lower than the lower bound, disproving the original assumption.

The “magic” in the proof consists of finding the appropriate choice of \( f \), as well as in defining the transformation \( I \) to begin with. Once that is done, the work in the proof is in showing \( J \) integral (which is harder for the more general theorems) and in deriving the lower bound. But the outline of the proof remains the same across all three theorems.

**Theorem 1.** \( e \) is transcendental.

*Proof.* Suppose not, so that \( e \) is algebraic. Then \( e \) satisfies some integer polynomial \( \sum a_i x^i = 0 \) with \( a_0 \neq 0 \). This means that

\[
a_0 + a_1 e + a_2 e^2 + \cdots + a_n e^n = 0.
\]

Let \( p \) be a (sufficiently large) prime, define a polynomial of degree \( m = (n + 1)p - 1 \) by

\[
f(x) = x^{p-1}(x - 1)^p \cdots (x - n)^p
\]

and let

\[
J = a_0 I(0) + a_1 I(1) + \cdots + a_n I(n).
\]
We derive two sets of inconsistent bounds on $J$, thus showing that the original hypothesis is false and $e$ is transcendental.

First, apply equation (??) to $J$:

$$J = \sum_{k=0}^{n} a_k I(k)$$

$$= \sum_{k=0}^{n} a_k \left( e^k \sum_{j=0}^{m} f^{(j)}(0) - \sum_{j=0}^{m} f^{(j)}(k) \right)$$

$$= \sum_{k=0}^{n} a_k e^k \sum_{j=0}^{m} f^{(j)}(0) - \sum_{k=0}^{n} a_k \sum_{j=0}^{m} f^{(j)}(k)$$

$$= -\sum_{j=0}^{m} \sum_{k=0}^{n} a_k f^{(j)}(k)$$

where the last equality follows because of the assumed linear dependence of the powers of $e$.

Consider the values of $f^{(j)}(k)$. If $j < p - 1$, then none of the factors in $f$ vanish by differentiation, so that $f^{(j)}(k) = 0$ for all $k$. If $j = p - 1$, then only the initial factor $x^{p-1}$ can vanish in any term in $f^{(j)}(x)$, and so $f^{(j)}(k) = 0$ if $k > 0$. In the case where $j = p - 1, k = 0$, the only term in the derivative that contributes a nonzero value is the term where $x^{p-1}$ is differentiated each time; that term gives

$$f^{(p-1)}(0) = (p-1)!(n!)^p.$$

Finally, if $j \geq p$, then the terms in $f^{(j)}(k)$ that are nonzero are those terms in which $(x - k)^p$ has been differentiated away; in these cases, the terms have a leading coefficient that is thus a multiple of $p!$.

Now assume $p > n$; then $f^{(p-1)}(0)$ is a multiple of $(p - 1)!$ but not of $p!$. Putting together the above computations, we get

$$J = Np! + a_0 M (p - 1)! = (p - 1)! (Np + a_0 M)$$

where $M, N$ are integers and $p \nmid M$. Assume also $p > a_0$; then $Np + a_0 M$ must be nonzero and thus $|J| \geq (p - 1)!$.

On the other hand, obviously $F(k) \leq (2n)^m$ and thus

$$|J| \leq |a_1| eF(1) + \cdots + |a_n| ne^n F(n) \leq an e^n (2n)^{(n+1)p - 1} = \frac{ane^n}{2n} ((2n)^{n+1})^p \leq c^p$$

where $a = \max(|a_1|, \ldots, |a_n|)$ and $c$ is some constant that does not depend on $p$.

But for $p$ large enough, $(p - 1)! > c^p$ no matter what $c$ is, so these two bounds on $J$ are contradictory.

**Theorem 2.** $\pi$ is transcendental.
Proof. Again suppose not. Then $i\pi$ is also algebraic. Suppose the minimal polynomial, $f$, of $i\pi$ has degree $d$, say

$$f(x) = \sum_{i=0}^{d} a_i x^i$$

with $a_0 \neq 0$, and let $\theta_1 = i\pi, \theta_2, \ldots, \theta_d$ be the conjugates of $i\pi$ (the other roots of $f$). Then since $e^{i\pi} = -1$, we have

$$(1 + e^{\theta_1})(1 + e^{\theta_2}) \cdots (1 + e^{\theta_d}) = 0$$

Note that since $\theta_i$ is algebraic for each $i$, then $a_d \theta_i$ is an algebraic integer.

Each term in this product can be written as a power of $e$, where the exponent is of the form

$$\beta_{\epsilon_1, \ldots, \epsilon_d} = \epsilon_1 \theta_1 + \epsilon_2 \theta_2 + \cdots + \epsilon_d \theta_d$$

and each $\epsilon_i$ is either 0 or 1. Denote by $\alpha_1, \ldots, \alpha_n$ those exponents that are nonzero. Note that at least one exponent is zero, and thus $n < 2^d$. We then have

$$(2^d - n) + e^{\alpha_1} + \cdots + e^{\alpha_n} = 0$$

We will show that if we define $f$ by

$$f(x) = a_d^{np} x^{p-1} (x - \alpha_1)^p \cdots (x - \alpha_n)^p$$

with $p$ a (sufficiently large) prime, then

$$J = I(\alpha_1) + \cdots + I(\alpha_n)$$

satisfies the same incompatible bounds as in the previous theorem.

Let $m = \deg f = (n+1)p - 1$. As before, we see that

$$J = \sum_{k=1}^{n} I(\alpha_k)$$

$$= \sum_{k=1}^{n} \left( e^{\alpha_k} \sum_{j=0}^{m} f^{(j)}(0) - \sum_{j=0}^{m} f^{(j)}(\alpha_k) \right)$$

$$= \sum_{k=1}^{n} e^{\alpha_k} \sum_{j=0}^{m} f^{(j)}(0) - \sum_{k=1}^{n} \sum_{j=0}^{m} f^{(j)}(\alpha_k)$$

$$= (n - 2^d) \sum_{j=0}^{m} f^{(j)}(0) - \sum_{j=0}^{m} \sum_{k=1}^{n} f^{(j)}(\alpha_k)$$

where the last equality follows from equation (3).

The remainder of the proof is quite similar to the above proof for $e$, except that we must first show that the sum over $k$ above is an integer; this was clear in the previous theorem.
Consider the inner sum over \( k \). It is clear that this is a symmetric polynomial with integer coefficients in \( a_d \alpha_1, \ldots, a_d \alpha_n \). The \( a_d \alpha_i \) are algebraic integers since the \( a_d \theta_i \) are. By the fundamental theorem of symmetric polynomials, it follows that the inner sum is in fact a polynomial in the elementary symmetric functions on the \( a_d \alpha_i \). Since the \( a_d \theta_i \) are algebraic integers, this is a symmetric polynomial with integer coefficients in the \( a_d \alpha_i \). Again applying the fundamental theorem of symmetric polynomials, we see that the sum over \( k \) must be a polynomial in the elementary symmetric functions of the \( a_d \theta_i \). But these elementary symmetric functions are simply the coefficients of the minimal polynomial of \( a_d i \pi \), which are integers. Thus the inner sum is an integer.

By arguments identical to those in the previous theorem, we have that 
\[
 f^{(j)}(0) \neq 0 \quad \text{when} \quad j < p
\]
and that
\[
 f^{(p-1)}(0) = (p-1)!(-a_d)^{np}(\alpha_1 \ldots \alpha_n)^p
\]
is an integral multiple of \( (p-1)! \) but is not divisible by \( p \) if \( p \) is chosen to exceed \( a_d \alpha_1 \ldots \alpha_n \). Thus if also \( p > n - 2^d \), we have that \( J \) is nonzero and divisible by \( (p-1)! \) and thus \( |J| \geq (p-1)! \). But again, similar to the proof in the previous theorem, we have that
\[
 |J| \leq |\alpha_1| e^{\alpha_1} F(|\alpha_1|) + \cdots + |\alpha_n| e^{\alpha_n} F(|\alpha_n|) \leq \epsilon^p
\]
and again these estimates are contradictory. \( \square \)

The Lindemann-Weierstrass theorem generalizes both these two statements and their proofs.

**Theorem 3.** If \( \alpha_1, \ldots, \alpha_n \) are algebraic and distinct, and if \( \beta_1, \ldots, \beta_n \) are algebraic and non-zero, then

\[
 \beta_1 e^{\alpha_1} + \cdots + \beta_n e^{\alpha_n} \neq 0
\]

Note that the facts that \( e \) and \( \pi \) are transcendental follow trivially from this theorem. For example, if \( e \) were algebraic, then \( e \) is the root of a polynomial \( \sum \beta_i x^i \) where \( \beta_i \in \mathbb{Q} \), in contradiction to the theorem.

**Proof.** The proof follows the same general lines as above, but there are additional complexities introduced by the arbitrary \( \alpha_i \). In the proof of the transcendality of \( \pi \), we were able to use facts about the relationship of the exponents in the proof; no such relationship is available to us in this more general setting.

Again start by supposing

\[
 \beta_1 e^{\alpha_1} + \cdots + \beta_n e^{\alpha_n} = 0 \quad (4)
\]

where the \( \alpha_i, \beta_i \) are as given.
Claim we can assume, without loss of generality, that $\beta_i \in \mathbb{Z}$. For if not, take all the expressions formed by substituting for one or more of the $\beta_i$ one of its conjugates, and multiply those by the equation above. The result is a new expression of the same form (with different $\alpha_i$), but where the coefficients are rational numbers. Clear denominators, proving the claim.

Next, claim we can assume that the $\alpha_i$ are a complete set of conjugates, and that if $\alpha_i, \alpha_j$ are conjugates, then $\beta_i = \beta_j$. To see this, choose an irreducible integral polynomial having $\alpha_1, \ldots, \alpha_n$ as roots; let $\alpha_{n+1}, \ldots, \alpha_N$ be the remaining roots, and define $\beta_{n+1} = \ldots \beta_N = 0$. Then clearly we have

$$\prod_{\sigma \in S_N} (\beta_1 e^{\alpha_{\sigma(1)}} + \cdots \beta_N e^{\alpha_{\sigma(N)}}) = 0$$

(Note the similarity with the proof for $\pi$). There are $N!$ factors in this product, so expanding the product, it is a sum of terms of the form

$$e^{h_1 \alpha_1 + \cdots + h_N \alpha_N}$$

with integral coefficients, and $h_1 + \cdots + h_N = N!$. Clearly the set of all such exponents forms a complete set of conjugates. By symmetry considerations, we see that the coefficients of two conjugate terms are equal. Also, the product is not identically zero. To see this, consider the term in the product formed by multiplying together, from each factor, the nonzero terms with the largest exponents in the lexicographic order on $\mathbb{C}$. Since the $\alpha_i$ are unique (because the polynomial is irreducible), there is only one term with this largest exponent,

and it has a nonzero coefficient by construction.

Finally, order the terms so that the conjugates of a particular $\alpha_i$ appear together. That is, for the remainder of the proof we may assume that

$$\beta_1 e^{\alpha_1} + \cdots + \beta_n e^{\alpha_n} = 0$$

with the $\beta_i \in \mathbb{Z}$, and that there are integers $0 = n_0 < n_1 < \cdots < n_r = n$ chosen so that, for each $0 \leq t < n$ we have

$$\alpha_{n_t+1}, \ldots, \alpha_{n_t+1} \text{ form a complete set of conjugates}$$

$$\beta_{n_t+1} = \beta_{n_t+2} = \cdots \beta_{n_t+1}$$

Now, since $\alpha_i, \beta_i$ are algebraic, we can choose $l$ such that $l \alpha_i, l \beta_i$ are algebraic integers. Let

$$f_i(x) = \frac{\mu_p((x - \alpha_1) \cdots (x - \alpha_n))^p}{(x - \alpha_i)}, \quad 1 \leq i \leq n$$

where again $p$ is a (large) prime. We will develop contradictory estimates for $|J_1 \ldots J_n|$, where

$$J_i = \beta_1 I_i(\alpha_1) + \cdots + \beta_n I_i(\alpha_n), \quad 1 \leq i \leq n$$

and $I_i$ is the integral associated with $f_i$, as above (see equation (??)).
Using equations (??) and (??), we see that

\[
J_i = \sum_{k=1}^{n} \beta_k I_i(\alpha_k)
\]

\[
= \sum_{k=1}^{n} \left( \beta_k e^{\alpha_k} \sum_{j=0}^{np-1} f_i^{(j)}(0) \right) - \sum_{k=1}^{n} \left( \beta_k \sum_{j=0}^{np-1} f_i^{(j)}(\alpha_k) \right)
\]

\[
= \left( \sum_{j=0}^{np-1} f_i^{(j)}(0) \right) \left( \sum_{k=1}^{n} \beta_k e^{\alpha_k} \right) - \sum_{k=1}^{n} \left( \beta_k \sum_{j=0}^{np-1} f_i^{(j)}(\alpha_k) \right)
\]

\[
= - \sum_{j=0}^{np-1} \sum_{k=1}^{n} \left( \beta_k f_i^{(j)}(\alpha_k) \right).
\]

Arguing similarly to the foregoing proofs, we see that \( f_i^{(j)}(\alpha_k) \) is an algebraic integer divisible by \( p! \) unless \( j = p - 1 \) and \( k = i \). In this particular case, we have that

\[
f_i^{p-1}(\alpha_i) = l^{np}(p-1)! \prod_{k=1}^{n} (\alpha_i - \alpha_k)^p
\]

and so again, if \( p \) is large enough, this is divisible by \( (p-1)! \) but not by \( p! \). Thus \( J_i \) is a nonzero algebraic integer divisible by \( (p-1)! \) but not by \( p! \). As before, we can prove that \( J_i \neq 0 \).

\( J_i \) can be written as follows:

\[
J_i = - \sum_{j=0}^{np-1} \sum_{t=0}^{r-1} \beta_{n+1} f_i^{(j)}(\alpha_{n+1} + \cdots + \alpha_{n+1})
\]

Note that by construction, \( f_i(x) \) can be written as a polynomial whose coefficients are polynomials in \( \alpha_i \), and the integral coefficients of \( those \) polynomials are integers independent of \( i \). Thus, noting that the \( \alpha_i \) form a complete set of conjugates and using the fundamental theorem on symmetric polynomials as in the previous proof, we see that the product of the \( J_i \) is in fact a rational number. But it is an algebraic integer, hence an integer. Thus \( J_1 \cdot \ldots \cdot J_n \in \mathbb{Z} \), and it is divisible by \( ((p-1)!)^n \). Thus \( |J_1 \cdot \ldots \cdot J_n| \geq (p-1)! \). But the same estimate as in the previous proofs shows that for each \( i \),

\[
|J_i| \leq \sum_{k=1}^{n} |\beta_k| |I_i(\alpha_k)| \leq \sum_{k=1}^{n} |\beta_k\alpha_k| e^{e^{\alpha_k}|F_i(\alpha_k)|}
\]

which as before is \( \leq c^p \) for some sufficiently large \( c \). These estimates are again in contradiction, proving the theorem.
Note that Theorems ?? and ?? are trivial corollaries of Theorem ??, as one would expect. (To prove that π is transcendental, note that if it were algebraic, then $e^{i\pi} = -1$ would be transcendental).

Here are some other more or less trivial corollaries.

**Corollary 4.** If $\alpha \neq 0$ is algebraic, then $e^{i\alpha}$ is transcendental.

*Proof.* If it were algebraic, say $e^{i\alpha} = \beta$, then we have

$$e^{i\alpha} - \beta e^0 = 0$$

in contradiction to the above theorem since $\alpha \neq 0$.

**Corollary 5.** If $\alpha \neq 0$ is algebraic, then cos $\alpha$ and sin $\alpha$ are both transcendental.

*Proof.* Recall that $\cos \alpha + i \sin \alpha = e^{i\alpha}$, which is transcendental. If either $\cos \alpha$ or $\sin \alpha$ were algebraic, then the other would be as well (and thus their sum would be) since $\cos^2 \alpha + \sin^2 \alpha = 1$. Therefore, both $\cos \alpha$ and $\sin \alpha$ are transcendental.

**Corollary 6.** If $\alpha > 0$ is algebraic with $\alpha \neq 1$, then $\ln \alpha$ is transcendental.

*Proof.* If $\beta = \ln \alpha$, then $e^\beta = \alpha$. By Corollary ??, since $\alpha$ is algebraic, $\beta$ cannot be.

**References**


