Mahler's Polynomials and the Roots of Unity

James Pedersen

1 Introduction

In their paper [4] "Mahler Polynomials and the Roots of Unity", Karl Dilcher and Larry Ericksen introduce the reader to the relationships between the partial sums

$$P_n(z) = \sum_{j=0}^{n-1} z^{2^j} = z + z^2 + \cdots + z^{2^{n-1}} \qquad (z \in \mathbb{C}, n \in \mathbb{N}) \tag{1}$$

of the function of the function $f(z) = \sum_{j=0}^{\infty} z^{2^j}$ (defined on the open unit disk), and the roots of unity. One of their paper's goals to is prove a result concerning the divisibility of $P_n(z)$ by cyclotomic polynomials (a result which they call Proposition 2.2) that they claim that Mahler only referred to implicitly. In addition, the paper contains many results concerning the distribution of zeros of $P_n(z)$. For instance, they show that for $n \geq 3$, all zeros of $P_n(z)$ must lie in some open disk centered at the origin with a radius depending on n [4, p.346], and also that for $n \geq 2$, all critical points of $P_n(z)$ must lie inside the unit circle [4, p.349]. In the process, the authors mention related conjectures, in particular, that there are infinitely many Wieferich primes, and that every $P_n(z)$ has at least one zero outside the unit circle. The paper concludes by connecting cyclotomic polynomials to Meresenne primes and by challenging the reader to generalize the paper's results to the polynomials $P_{q;n}(z) = \sum_{j=0}^{n-1} z^{q^n}$, where q is some arbitrary natural number, not necessarily 2.

This paper fills in the details of Proposition 2.2, and builds up the algebraic and number-theoretic machinery that is necessary to understand the proposition. Without further ado, into the Mathematics we go!

2 Definitions:

Definition 2.1. *If z is an nth root of unity and $z^n \neq 1$ for every natural $k < n$, we say that z is a **primitive nth root of unity** [7].*

Definition 2.2. *According to Wolfram Mathworld, for $n \in \mathbb{N}$ the **nth cyclotomic polynomial** is the polynomial:*

$$\phi_n(z) = \prod_{\substack{k=1, \\ \zeta_k \ primitive}}^{n} (z - \zeta_k) \tag{2}$$

where ζ_k denotes the kth primitive nth root of unity [6].

Definition 2.3. *If $m \in \mathbb{N}$ then $t(m)$, called the **order of m modulo 2**, is the smallest positive integer t such that $t^m \equiv 1$ mod m [4, p.339].*

Definition 2.4. *For a prime p, let $w(p)$ be the highest power of p that divides $2^{p-1} - 1$ [4, p.344].*

Remark. *Dilcher and Ericksen note that for every prime p, $w(p) \geq 1$ by Fermat's little theorem [4, p.344].*

Definition 2.5. *A **Wieferich prime** is a prime p satisfying $w(p) = 1$ [4, p.344].*

The following algebraic definitions are taken from Bruce Ikenaga's website [5] and from Patrick's Morandi's course notes [3], and are provided to aid the reader.

Remark. *F, in the following theorem statements, denotes any arbitrary field, finite or infinite.*

Definition 2.6. *A **ring** [5, Rings] is an abelian group R with binary operation + ("addition"), together with a second binary operation $\bullet$ ("multiplication"). The operations satisfy the following axioms*

1. *Multiplication is associative: For all $a, b, c \in \mathbb{R}$,*

$$(a \bullet b) \bullet c = a \bullet (b \bullet c). \tag{3}$$

2. *The Distributive Law holds. For all $a, b, c \in \mathbb{R}$,*

$$a \bullet (b + c) = a \bullet b + a \bullet c \text{ and } (a + b) \bullet c = a \bullet c + c \bullet c \tag{4}$$

Definition 2.7. *We say that a ring R has **multiplicative identity** (and call R a **ring with 1**, or a **ring with unity**) [5, Rings] if there is an element $1 \in \mathbb{R}$ such that $1 \neq 0$, and such that for all $a \in \mathbb{R}$,*

$$1 \bullet a = a \text{ and } a \bullet 1 = a \tag{5}$$

Definition 2.8. *If R is a ring, then $R[x]$ (called the **ring of polynomials in x with coefficients in** $\mathbb{R}$ [5, Polynomial-Rings], consists of all formal sums $\sum_{i=0}^{\infty} a_i x^i$, where $a_i \neq 0$ for all but finitely many values of i. Addition and multiplication of polynomials in this ring is defined in the usual way. It is easily verified that $R[x]$, as defined above, satisfies the ring axioms.*

Definition 2.9. *If $f(x) = \sum_{i=0}^{\infty} a_i x^i$ is a nonzero polynomial, the **degree** [3, p.53-54] of f, denoted by deg(f), is the largest $n \geq 0$ such that $a_n \neq 0$. If f is the zero polynomial (has all $a_n = 0$), we say that f has degree $-\infty$, and write $deg(f) = -\infty$. We formally define $-\infty + -\infty = -\infty$ and for every integer n, we formally define $-\infty + n = -\infty$. As stated by Morandi, the purpose of these conventions is to make the relationship between the degree of a product of polynomials and the degree of each as straight-forward as possible.*

Definition 2.10. *Let f and g be polynomials in F[x]. Then we say that* ***f divides g*** *[3, p.53-54] and write* $\boldsymbol{f \mid g}$ *if there is a polynomial* $h \in F[x]$ *with* $g = fh$.

Definition 2.11. *If R and S are rings, a function* $f : \mathbb{R} \to S$ *is a* ***ring homomorphism*** *(or a* ***ring map****) [5, Ring homomorphisms and isomorphisms] if* $f(x+y) = f(x)+f(y)$ *and* $f(xy) = f(x)f(y)$ *for all* $x, y \in R$. *If* R, S *are rings with identity, then we require* $f(1) = 1$.

Definition 2.12. *For* $x \in \mathbb{Z}$, *I define* $\overline{x}$ *as* $\{y \in \mathbb{Z} : y \equiv x \bmod n\}$. *This is called the* ***equivalence class of a mod n****.*

Definition 2.13. *Let R be a ring. An* ***ideal*** *I of R [5, Ideals] is a subset of R such that*

1. *I is closed under addition: If* $a, b \in I$, *then* $a + b \in I$
2. *I contains the zero element of R:* $0 \in I$.
3. *I is closed under additive inverses: If* $a \in I$, *then* $-a \in I$.
4. *If* $a \in I$ *and* $r \in R$, *then* $ar \in I$ *and* $ra \in I$. *That is, S is closed under multiplication (on either side) by artibrary ring elements.*

Example. *Let* $R = \mathbb{Z}$. *For* $n \in \mathbb{Z}$, *let* $n\mathbb{Z} \equiv \{an : a \in \mathbb{Z}\}$, *the set of integer multiples of n. I claim that* $n\mathbb{Z}$ *is an ideal of* $\mathbb{Z}$. *Suppose* $x, y \in n\mathbb{Z}$. *Then there exists integers* a, b *such that* $x = an$, $y = bn$. *Then* $x + y = (a + b)n$, *so* $x + y \in \mathbb{Z}$, *so* $n\mathbb{Z}$ *is closed under addition [3, p.56]. Evidently,* $0 \in n\mathbb{Z}$, *and if* $a \in n\mathbb{Z}$ *then* $-a \in n\mathbb{Z}$. *Now suppose that* $x = na \in n\mathbb{Z}$ *and* $r \in R$. *Then* $xr = rx = r(na) = r(an) = (ra)n$, *and since* $ra \in \mathbb{Z}$, $xr, rx \in I$. *Thus I is an ideal of* $\mathbb{Z}$. *Note that if* $n > 0$,

$$n\mathbb{Z} = \{\dots, -2n, -n, 0, n, 2n, \dots\} = \overline{0}. \tag{6}$$

This important connection will be revisited later.

Example. *Suppose* $f \in F[x]$. *Now consider* $I = \{gf : g \in F[x]\}$, *the set of all multiples of f. Then* I, *denoted by* (f), *is an ideal of* $F[x]$, *following the exact calculation as in the previous example.*

Example. *If R is any commutative ring and* $a, b \in R$, *then* $I = \{ar + bs : r, s \in R\}$ *is an ideal of R [3, p.57]. Also, if* $a_1, \dots, a_n \in R$, *then* $I = \{a_1r_1 + \dots + a_nr_n : r_1, \dots, r_n \in R\}$ *is an ideal of R, by a similar calculation [3, p.57].*

Remark. *Following the remarks given in [3, p.59-60], given any ideal of a ring R, we may define an equivalence relation* $x \equiv y \bmod I$ *provided that* $x - y \in I$. *It is easily proven that this defines an equivalence relation. Now we construct the corresponding equivalence classes. For any* $r \in I$, *I define* $\overline{r} = \{s \in R : s \equiv r \bmod I\}$, *which I call the* ***equivalence class of r mod I****.*

Definition 2.14. *The following definitions are found at [3, p.58]. Let R be a ring and I an ideal of R. If* $a \in R$, *then the* ***coset a + I*** *is defined as* $a + I = \{a + x : x \in I\}$. *a is called a* ***coset representative*** *of* $a + I$. *If* $a + I$ *and* $b + I$ *are cosets of R, we define*

$(a+I)+(a+I)=(a+b)+I$ *and* $(a+I)(a+I)=ab+I$. *Furthermore, if* n *is a nonnegative integer and* $a+I$ *a coset, we define*

$$n(a+I)=\underbrace{(a+I)+\ldots+(a+I)}_{n\ times} \tag{7}$$

$$(a+I)^n=\underbrace{(a+I)\ldots(a+I)}_{n\ times} \tag{8}$$

Furthermore, I formally define $(a+I)^0=1+I=\overline{1}$ *and* $0(a+I)=0+I$. *It must be verified that coset arithmetic is well defined, that is, the resulting coset does not depend on the choices of coset representative for the starting cosets. This is done in [3, p.59-60].*

Now, the following lemma connects cosets to equivalence classes.

Lemma 2.1. *Let I be an ideal of the ring R. Then for any* $a\in R, \overline{a}=a+I$.

Proof.

$$\begin{aligned}\overline{a}&=\{x\in R:x\equiv a \bmod I\} &(9)\\ &=\{x\in R:x-a\in I\} &(10)\\ &=\{x\in R:\exists r\in I:x-a=r\} &(11)\\ &=\{x\in R:\exists r\in I:x=a+r\} &(12)\\ &=a+I &(13)\end{aligned}$$

□

Theorem 2.2. *Let I be an ideal of a ring R. Then R/I, the set of cosets of I, forms a ring under coset addition and multiplication. The additive identity, or zero, is the coset* $0+I=\overline{0}$. *The reader may easily verify that R/I is indeed a ring as asserted. R/I is called a* ***quotient ring*** *of R. Furthermore, if R is a ring with unity, then R/I is also, with identity element* $\overline{1}=1+I$, *and if R is a commutative ring (multiplication is commutative), then* R/I *is also a commutative ring. The verification of these steps may be found in [3, p.60].*

Remark. *Since cosets are equivalence classes,* R/I *is the same thing as the set of equivalence classes mod* I *of R.*

The following quotient ring is especially pertinent to Dichler and Eriksen's results concerning divisibility by cyclotomic factors. It is $\frac{Z[z]}{(z^k-1)}$, where k is a natural number and $Z[z]$ is the ring of polynomials over $\mathbb{C}$ with integer coefficients. Note that because $Z[z]$ is a ring with unity, $\frac{Z[z]}{(z^k-1)}$ is too. The zero of this ring is $\overline{z^k-1}$. $\frac{Z[z]}{(z^k-1)}$ is a ring with unity. It is important to note that with respect to the ideal (z^k-1), for any $p(z)\in\mathbb{Z}[z], \overline{p(z)}=p(z)+(z^k-1)$, by the previous lemma. This means that $\frac{Z[z]}{z^k-1}=\{\overline{p(z)}:p(z)\in\mathbb{Z}[z]\}$.

Now, in the context of $\mathbb{Z}[z]$ with ideal $\left(z^k - 1\right)$, $\exists r \in \mathbb{Z}[z]$ where $p(z) = q(z) + r(z)(z^k - 1)$ iff $p \equiv q \bmod z^k - 1$ iff $p - q \in \left(z^k - 1\right)$ iff $p + \left(z^k - 1\right) = q + \left(z^k - 1\right)$ iff $\overline{p} = \overline{q}$ (the third if-and-only-if follows from Lemma 5.14 in [3, p.59]). Now, I prove the following lemmas which are critical for the proof of Proposition 2.2 (Dilcher and Eriksen do not themselves prove them).

Lemma 2.3. *For non-negative integers a, b where $a \geq b$ and positive integer k, if $a \equiv b$ mod c then $z^a \equiv z^b$ mod $z^c - 1$.*

Proof. Suppose a, b, c are as given and $a \equiv b \bmod c$. Now, let $a - b = lk$, where $k \in \mathbb{Z}$. Note that since $a \geq b, a - b \geq 0$, and since $k > 0$, $l \geq 0$. Now, $a = b + lk \geq 0$. Therefore, $\overline{z^a} = \overline{z^{b+lk}}$. Now, we have that $z^k - 1 \equiv 0 \bmod z^k - 1$. Therefore, $\overline{z^k} - \overline{1} = \overline{z^k - 1} = \overline{0}, \overline{z^k - 1} = \overline{0}, \overline{z^k} = \overline{1}, z^k \equiv 1 \bmod z^k - 1$. Furthermore, we have that $\overline{z^a} = \overline{z^{kl+b}} = \overline{z^{kb} z^b} = \left(\overline{z^k}\right)^b \overline{z^b} = \left(\overline{1}\right)^b \overline{z^b} = \left(\overline{1}\right) \overline{z^b} = \overline{z^b}$, all by the ring axioms under the ring $\frac{\mathbb{Z}[z]}{z^k - 1}$. Therefore, $z^a \equiv z^b \bmod z^k - 1$, and this completes the proof. □

Lemma 2.4. *For non-negative integers a, b and positive integer k, if $a \equiv b$ mod c then $z^a \equiv z^b$ mod $z^c - 1$.*

Proof. If $a \geq b$, then by the previous lemma, $z^a \equiv z^b \bmod z^c - 1$. If $a < b$, then $b > a$, $b \geq a$, so $z^b \equiv z^a \bmod z^c - 1$. Because congruence is an equivalence relation, $z^a \equiv z^b \bmod z^c - 1$. This completes the proof. □

Now, we have built up the minimum algebric structure required to understand the proof of Proposition 2.2. Yet a familiarity with basic number-theoretic concepts is also required, thus the following number-theoretic definitions and examples are provided to aid the reader.

Definition 2.15. *According to Long, the **gcd** [1, p.33] of two integers not both zero is the largest positive integer that divides both. That is, for $n, m \in \mathbb{Z}$ not both zero, $gcd(n, m) = d$ means that $d|n$ and $d|m$ and that if $k \in \mathbb{N}, d < k$, then $\neg(k|n$ and $k|m)$.*

Definition 2.16 (Totient Function)**.** *According to [1, p.85], the Euler Totient Function is the function $\phi : \mathbb{N} \to \mathbb{N}$ where $\phi(n)$ counts the number of natural numbers k in the range $1 \leq k \leq n$ where $gcd(k, n) = 1$ (that is, the number of positive integers in that range that are relatively prime to n).*

Definition 2.17. *According to [1], any subset R of the integers is called a **reduced residue system** modulo $n \in \mathbb{N}$ iff*

1. *gcd(r, n) = 1 for every $r \in R$;*
2. *R contains $\phi(n)$ elements;*
3. *No two distinct elements of R are congruent modulo n.*

Remark. *Note that by definition, all reduced residue systems modulo n have $\phi(n)$ elements.*

Example. *If p is a prime, the set $S = \{1, \dots, p-1\}$ is a reduced residue system modulo p. All of the elements of S are relatively prime with p. $\phi(p) = p - 1 = \#S$. Now suppose $r_1, r_2 \in S$ satisfy $r_1 \equiv r_2 \bmod p$. Then $p|r_1 - r_2$. But $r_1 < p - 1$ and $r_2 \geq 0$, so $-r_2 \leq 0, r_1 - r_2 < p - 1$. This contradictions the claim that $p|r_1 - r_2$ unless $r_1 - r_2 = 0, r_1 = r_2$. Therefore r_1 and r_2 are not distinct. This shows that S is indeed a reduced resiude system modulo p.*

The following lemma is also critical to the proof of Preposition 2.2.

Lemma 2.5. *If $\{w_1, w_2, \dots, w_m\}$ is a reduced residue system modulo n, where n is odd, then for any non-negative integer j, the set $\{2^j w_1, 2^j w_2, \dots, 2^j w_m\}$ is also a reduced residue system modulo n.*

Proof. Suppose j is an arbitrary non-negative integer. First, I show that $gcd(2^j w_k, n) = 1$ for every $k \in \{1, \dots, m\}$. So suppose $k \in \{1, \dots, m\}$. Since $\{w_1, w_2, \dots, w_m\}$ is a reduced residue system modulo n, $gcd(w_k, n) = 1$. Now evidently, $1|w_k$ and $1|n$. Now suppose that $l > 1, l|2^j w_k$ and $l|n$. However, n is odd, so l is odd (if l were even then n would be even), so $l \nmid 2^j$. However, $l|2^j w_k$. Therefore, $l|w_k$. But this is a contradiction because $l > 1$ divides both w_k and n, yet $gcd(w_k, n) = 1$. Therefore, $gcd(2^j w_k, n) = 1$ for every $k \in \{1, \dots, m\}$. Second, since $\{w_1, w_2, \dots, w_m\}$ contains elements, $\phi(n)$, $\{2^j w_1, 2^j w_2, \dots, 2^j w_m\}$ must also contain $\phi(n)$ elements. Finally, suppose that there exist distinct $k_1, k_2 \in \{1, \dots, m\}$ where $2^j w_{k1} \equiv 2^j w_{k2} \bmod n$. Then there exists an integer a such that $2^j w_{k1} = 2^j w_{k2} + an$. Thus $2^j(w_{k1} - w_{k2}) = an$. Thus $2^j|an$. But since n is odd, $2^j \nmid n$, therefore $2^j|a$. Thus, there exists an integer l satisfying $2^j l = a$. But then $2^j(w_{k1} - w_{k2}) = 2^j ln$, $w_{k1} - w_{k2} = ln$, implying that $w_{k1} \equiv w_{k2} \bmod n$, which is impossible unless $w_{k_1} = w_{k_2}$, since $\{w_1, \dots, w_m\}$ is a reduced residue system modulo n. Therefore, $2^j w_{k_1} = 2^j w_{k_2}$. Thus, all three reduced residue system conditions have been verified, so $\{2^j w_1, 2^j w_2, \dots, 2^j w_m\}$ is a indeed a reduced residue system modulo n. □

Now, I formulate a more ring-theoretic way of understanding reduced residue systems. I claim that if $\overline{a}$ is an equivalence class of integers modulo n, then either all elements of $\overline{a}$ are relatively prime to n, or all elements of $\overline{a}$ are not relatively prime to n. Suppose that it is false that all elements of $\overline{a}$ are not relatively prime to n. Then there exists $k_1 \in \overline{a}$ satisfying $gcd(k_1, n) = 1$. I claim therefore that for any $k_2 \in \overline{a}$, $gcd(k_2, n) = 1$. Suppose not, that there exists $k_2 \in \overline{a}$ satisfying $gcd(k_2, n) \neq 1$. I will derive a contradiction by showing that this implies that $gcd(k_1, n) \neq 1$. Since $gcd(k_2, n) \neq 1$, there exists a natural number $l > 1$ where $l|k_2$ and $l|n$. However, since $k_1, k_2 \in \overline{a}$, $k_1 \equiv a \bmod n$ and $k_2 \equiv a \bmod n$, therefore there exists integers v_1, v_2 satisfying $k_1 - a = v_1 n, k_2 - a = v_2 n$. Therefore $k_1 - k_2 = (v_1 - v_2)n$, so $k_1 = k_2 + (v_1 - v_2)n$. Now, since $l|k_2$ and $l|n$, we have that $l|k_1$, but since $l|n$ and $l > 1$ and $gcd(k_1, n) = 1$, we have the aforementioned contradiction.

Staring at the definition of reduced residue system for long enough, one sees that a reduced residue system modulo n is merely a set of $\phi(n)$ integers, each from a distinct equivalence class of integers modulo that is relatively prime to n (of which there are $\phi(n)$ of). Therefore, between any two reduced residue systems $\{w_1, \dots, w_m\}$ and

$\{x_1, \ldots, x_m\}$ modulo n, we may form a one-to-one correspondence linking elements belonging only to the same equivalence class of integers modulo n. For instance, suppose that $\{1, 2, 3, 4\}$ and $\{17, 21, 43, 19\}$ are two reduced residue systems modulo five. Then the aforementioned bijection associates 17 to 2, 19 to 4, 43 to 3, and 21 to 1, because $17 \equiv 2 \bmod 5$, $19 \equiv 4 \bmod 5$, $43 \equiv 3 \bmod 5$, and $21 \equiv 1 \bmod 5$.

Now, we have built up enough machinery to understand Dilcher and Eriksen's Proposition 2.2.

3 Proposition 2.2

Theorem 3.1 (Proposition 2.2)**.** *If $p \geq 3$ is a fixed prime, then for all $m \geq 1$ we have that:*

$$\phi_p(z^{2^{t(p)m}-1}) \text{ divides } P_{t(p)pm}(z). \tag{14}$$

Proof. Dichler and Ericksen proceed as follows: Suppose p and m are as given. Then, letting $t = t(p)$ (note that $t \in \mathbb{N}$), regrouping the terms, they note that:

$$\begin{aligned}
P_{tpm}(z) = \quad & \left(z^{2^0} + z^{2^1} + z^{2^2} + \ldots + z^{2^{tm-1}}\right) \\
& + \left(z^{2^{tm}} + z^{2^{tm+1}} + \ldots + z^{2^{2tm-1}}\right) \\
& + \left(z^{2^{2tm}} + z^{2^{2tm+1}} + \ldots + z^{2^{3tm-1}}\right) \qquad (15) \\
& + \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
& + \left(z^{2^{(p-1)tm}} + z^{2^{(p-1)tm+1}} + \ldots + z^{2^{tpm-1}}\right) \\
= \quad & \left(z^{2^0} + z^{2^{tm}} + z^{2^{2tm}} + \ldots + z^{2^{(p-1)tm}}\right) \\
& + \left(z^{2^1} + z^{2^{tm+1}} + z^{2^{2tm+1}} + \ldots + z^{2^{(p-1)tm+1}}\right) \\
& + \left(z^{2^2} + z^{2^{tm+2}} + z^{2^{2tm+2}} + \ldots + z^{2^{(p-1)tm+2}}\right) \qquad (16) \\
& + \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
& + \left(z^{2^{tm-1}} + z^{2^{2tm-1}} + z^{2^{3tm-1}} + \ldots + z^{2^{tpm-1}}\right) \\
= \quad & \sum_{j=0}^{tm-1} \left(z^{2^j} + z^{2^{tm+j}} + z^{2^{2tm+j}} + \ldots + z^{2^{(p-1)tm+j}}\right) \qquad (17) \\
= \quad & \sum_{j=0}^{tm-1} z^{2^j} \left(1 + z^{2^j(2^{tm}-1)} + z^{2^j(2^{2tm}-1)} + \ldots + z^{2^j(2^{(p-1)tm}-1)}\right) \qquad (18) \\
= \quad & \sum_{j=0}^{tm-1} z^{2^j} Q_j(z), \qquad (19)
\end{aligned}$$

where

$$Q_j(z) = 1 + z^{2^j(2^{tm}-1)} + z^{2^j(2^{2tm}-1)} + \ldots + z^{2^j(2^{(p-1)tm}-1)}. \tag{20}$$

Then they assert the fact that

$$\phi_p(z^{2^{tm}-1})\left(z^{2^{tm}-1}-1\right) = z^{p(2^{tm}-1)}-1, \tag{21}$$

which follows from the identity $\phi_p(z) = z^{p-1}+z^{p-2}+\ldots+z+1$ and the difference of powers formula $(z^p-1) = (z-1)\phi_p(z)$, after having substituted $z^{2^{tm}-1}$ in for z. As the authors then clearly state, we want to show that each $Q_j(z)$ is divisible by $\phi_p(z^{2^{tm}-1})$. Then, the authors write that "to do so, we reduce each $[Q_j(z)]$[1] modulo $[z^{p(2^{tm}-1)}-1]$. This can be achieved by reducing the exponents modulo $p(2^{tm}-1)$". What is meant by this statement is very unclear. Then the authors make a calculation regarding the aforementioned exponents, which is restated more clearly as follows. For a fixed $\nu \in \{1,\ldots,p-1\}$, let $x_{j,\nu} = 2^j(2^{\nu tm}-1)$ be the exponent of a single monomial of $Q_j(z)$. Now by the difference of powers formula,

$$x_{j,\nu} = 2^j(2^{\nu tm}-1) = 2^j(2^{tm}-1)\left(2^{(\nu-1)tm}+2^{(\nu-2)tm}+\ldots+2^{tm}+1\right) \tag{22}$$

Let $B = 2^{tm}-1$. Evidently, $x_{j,\nu}$ is divisible by B, and (abbreviating $x_{j,\nu}$ by x until further notice),

$$x/B = 2^j\left(2^{(\nu-1)tm}+2^{(\nu-2)tm}+\ldots+2^{tm}+1\right). \tag{23}$$

Now, because $2^t \equiv 1 \bmod p$ by the definition of t, all of the monomials in the parenthesis of the above equation are congruent to one modulo p, thus after adding and multiplying the congruences we have that

$$x/B \equiv 2^j\nu \pmod{\text{p}} \tag{24}$$

as asserted by Dilcher and Eriksen. Now, this gives us that for some $k \in \mathbb{Z}$, $x/B - 2^j\nu = kp$, $x - 2^jB\nu = k(pB)$, thus $x \equiv 2^jB\nu \bmod pB$. As $x = x_{j,\nu}$, we thus have that:

$$x_{j,\nu} \equiv 2^jB\nu \pmod{p(2^{tm}-1)} \tag{25}$$

Now, the authors state that because $\{\nu : \nu = 1,\ldots,p-1\}$ is a reduced residue system modulo p, $\{2^j\nu : \nu = 1,\ldots,p-1\}$ is also a reduced residue system modulo p. This follows from Lemma 2.5. However, the authors assert in the next line that "Hence $[Q_j(z)]$ (mod $z^{p(2^{tm}-1)}-1$) is the same for any j, with the terms (other than the initial "1") permuted". The authors throw a lot of details under the rug with this assertion. What they mean is that for any $j_1, j_2 \in \{0,\ldots,tm-1\}$, $Q_{j_1}(z) \equiv Q_{j_2}(z) \bmod z^{p(2^{tm}-1)}-1$. The authors do not prove this assertion, so what follows is my own proof of it. Suppose that j_1, j_2 are in the aforementioned range. Now $Q_{j_1}(z) = \sum_{\nu=0}^{p-1} z^{x_{j_1,\nu}}$, $Q_{j_2}(z) = \sum_{\nu=0}^{p-1} z^{x_{j_2,\nu}}$. I claim that there exists a bijection $\sigma : J_{p-1} \to J_{p-1}$ of ${J_{p-1}}$[2] so that for every $\nu \in \{1,\ldots,p-1\}$, $x_{j_1,\nu} \equiv x_{j_2,\sigma(\nu)} \bmod p(2^{tm}-1)$. If I can show this, then I will be done. Why is this? Suppose we had such a σ. Then for every $\nu \in \{1,\ldots,p-1\}$,

[1] I use square brackets to denote my insertion of text into quotes. I do not attach any mathematical meaning to the notation at any point within this proof.

[2] For a natural number n, I define J_n as $\{k \in \mathbb{N} : 1 \le k \le n\}$.

by Lemma 2.3 we have that $z^{x_{j_1,\nu}} \equiv z^{x_{j_1,\sigma(\nu)}} \bmod z^{p(2^{tm}-1)} - 1$. Since congruence is preserved under addition, adding up each congruence yields that:

$$\sum_{\nu=1}^{p-1} z^{x_{j_1,\nu}} = \sum_{\nu=1}^{p-1} z^{x_{j_2,\sigma(\nu)}} \pmod{z^{p(2^{tm}-1)} - 1} \tag{26}$$

Now let $y_\nu = x_{j_2,\nu}$. Then because σ is a permutation, we have that $\{y_\nu\} = \{y_{\sigma(\nu)}\}$, $\{x_{j_2,\nu}\} = \{x_{j_2,\sigma(\nu)}\}$, thus we have that $\sum_{\nu=1}^{p-1} z^{x_{j_2,\sigma(\nu)}} = \sum_{\nu=1}^{p-1} z^{x_{j_2,\nu}}$, and from this and the above equation, we obtain that

$$\sum_{\nu=1}^{p-1} z^{x_{j_1,\nu}} = \sum_{\nu=1}^{p-1} z^{x_{j_2,\nu}} \pmod{z^{p(2^{tm}-1)} - 1} \quad . \tag{27}$$

After adding one to both sides (as for every j, $x_{j,0} = 0$), we obtain exactly what we were trying to prove.

Now, I will exhibit such a σ. First, I prove the following lemma.

Lemma 3.2. *For every $l_1, l_2 \in \{0, \dots, tm-1\}$ and $X \in J_{p-1}$, there exists a unique $Y \in J_{p-1}$ such that $2^{l_1} BX \equiv 2^{l_2} BY \bmod p(2^{tm}-1)$.*

Proof. Suppose that l_1, l_2 are as described and $X \in J_{p-1}$. Now let $S_1 = \{2^{l_1}\nu : \nu = 1, \dots, p-1\}$ and $S_2 = \{2^{l_2}\nu : \nu = 1, \dots, p-1\}$. Both S_1 and S_2 are reduced residue systems modulo p by Lemma 2.5. Now, because every reduced residue system modulo p is a choice of exactly one element from each of the $p-1$ equivalence classes of integers modulo p, if we let α denote the equivalence class of integers modulo p that $2^{l_2}X$ belongs to, there exists exactly one element $\zeta = 2^{l_1}Y$ of S_2 in α. But since all elements of the same equivalence class are equivalent to one another, we have that $2^{l_1}X \equiv 2^{l_2}Y \bmod p$. Multiplying each part of the congruence by B, we have that $2^{l_1}BX \equiv 2^{l_2}YB \bmod p(2^{tm-1}-1)$. Thus we have found a $Y \in J_{p-1}$ with the desired property. Now suppose Y' is such that $2^{l_1}BX \equiv 2^{l_2}BY' \bmod p(2^{tm}-1)$. Then $2^{l_1}X \equiv 2^{l_2}Y' \bmod p$. However, $2^{l_2}Y' \in S_2$, but since there is only one element of S_2 belonging to the equivalence class α, we have that $2^{l_2}Y' = 2^{l_2}Y$, therefore $Y = Y'$, which is what we were trying to show. Thus, the element Y that we found is unique, which completes the proof of the lemma. □

Therefore, we have that for every $\nu \in J_{p-1}$ there exists a unique $\sigma(\nu) \in J_{p-1}$ such that $2^{j_1}B\nu \equiv 2^{j_2}B\sigma(\nu) \bmod p(2^{tm}-1)$. Rephrasing this, we have that for every $\nu \in J_{p-1}$, $2^{j_1}B\nu \equiv 2^{j_2}B\sigma(\nu) \bmod p(2^{tm}-1)$. Equivalently, for every $\nu \in J_{p-1}$, $x_{j_1,\nu} \equiv x_{j_2,\sigma(\nu)} \bmod p(2^{tm}-1)$. I now claim that σ is bijective. First, I show that σ is onto. Suppose that $y \in J_{p-1}$. Then by Lemma 3.2, there exists a $\nu \in J_{p-1}$[3] such that $2^{j_1}B\nu \equiv 2^{j_2}By \bmod p(2^{tm}-1)$. But by uniqueness of $\sigma(\nu)$, $y = \sigma(\nu)$. This shows that σ is onto. Now I claim that σ is one-to-one. So suppose that $\sigma(\nu_1) = \sigma(\nu_2)$. I show that $\nu_1 = \nu_2$. Now, we have that $2^{j_1}B\nu_1 \equiv 2^{j_2}B\sigma(\nu_1) \bmod p(2^{tm}-1)$ and $2^{j_1}B\nu_2 \equiv 2^{j_2}B\sigma(\nu_2) \bmod p(2^{tm}-1)$. But $2^{j_2}B\sigma(\nu_1) = 2^{j_2}B\sigma(\nu_2)$. Thus, we have that $2^{j_2}B\sigma(\upsilon_1) \equiv 2^{j_1}B\nu_1 \bmod$

[3] Using the first sentence in the above paragraph with $l_1 = j_2, l_2 = j_1$.

$p(2^{tm}-1)$ and $2^{j_2}B\sigma(v_1) \equiv 2^{j_1}Bv_2 \mod p(2^{tm}-1)$. Now if we use the first line of the previous paragraph with $l_1 = j_2, l_2 = j_1, X = \sigma(v_1)$, by the uniqueness of the corresponding Y, we have that $v_1 = v_2$, which is what we were trying to show. So we have found a permutation σ with the desired property, which is what we were trying to show.

The rest of the proof is rather straightforward. Since we now know that for every $j_1, j_2 \in \{0, \dots, tm-1\}$, $Q_{j_1}(z) \equiv Q_{j_2}(z) \mod z^{p(2^{tm}-1)} - 1$, we have that for every $j \in \{0, \dots, tm-1\}$, $Q_0(z) \equiv Q_j(z) \mod z^{p(2^{tm}-1)} - 1$. Now, as stated by Dilcher and Eriksen, if we can show that:

$$1 + z^{2^{tm}-1} + z^{2^{2tm}-1} + \dots + z^{2^{(p-1)tm}-1} \equiv \phi_p(z^{2^{tm}-1}) \mod z^{p(2^{tm}-1)} - 1 \quad , \tag{28}$$

we will be done. This is because $Q_0(z) = 1 + z^{2^{tm}-1} + z^{2^{2tm}-1} + \dots + z^{2^{(p-1)tm}-1}$, and if $Q_0(z) \equiv \phi_p(z^{2^{tm}-1}) \mod z^{p(2^{tm}-1)} - 1$, then for an arbitrary $j \in \{0, \dots, tm-1\}$, $Q_j(z) \equiv \phi_p(z^{2^{tm}-1}) \mod z^{p(2^{tm}-1)} - 1$, so

$$Q_j(z) - \phi_p(z^{2^{tm}-1}) = K(z)\left(z^{p(2^{tm}-1)} - 1\right) = K(z)\phi_p(z^{2^{tm}-1})\left(z^{2^{tm}-1} - 1\right), \tag{29}$$

by Equation 21. This immediately shows that $\phi_p(z^{2^{tm}-1})$ divides $Q_j(z)$, but since this is true for arbitrary j in the desired range, we have that $\phi_p(z^{2^{tm}-1})$ divides $P_{t(p)pm}(z)$, which is what the Proposition asks to be shown.

Then, Dilcher and Eriksen prove Equation 28. As they state, they show that for $v = 1, \dots, p-1$,

$$2^{vtm} - 1 \equiv v(2^{tm}-1) \mod p(2^{tm}-1) \quad . \tag{30}$$

This completes the proof because if it is true, then for $v = 1, \dots, p-1$, $z^{2^{vtm}-1} \equiv z^{v(2^{tm}-1)} \mod z^{p(2^{tm}-1)} - 1$ by my Lemma 2.5, therefore since $\phi_p(z^{2^{tm}-1}) = 1 + z^{2^{tm}-1} + z^{2(2^{tm}-1)} + \dots + z^{(p-1)(2^{tm}-1)}$, after adding up each congruence and adding one we have that:

$$1 + z^{2^{tm}-1} + z^{2^{2tm}-1} + \dots + z^{(2^{(p-1)tm}-1)} \equiv \phi_p(z^{2^{tm}-1}) \mod z^{p(2^{tm})-1} - 1, \tag{31}$$

which is what we were trying to show. Finally, Equation 30 is true because, as the authors state, it is equivalent upon rearrangement to the equation

$$(2^{tm}-1)[2^{(v-1)tm} + 2^{(v-2)tm} + \dots + 2^{tm} + 1 - v] \equiv 0 \mod p(2^{tm}-1), \tag{32}$$

which is true since the expression in square brackets vanishes modulo p. With this, the proof is complete. □

4 Conclusion:

After this result is proven, the authors prove a number of subsequent results. Some of those results utilize Proposition 2.2 and some do not. The ones that do not mainly concern the distribution of zeros of $P_n(z)$. For instance, Dilcher and Eriksen also prove that for all $n \geq 3$, all zeros of $P_n(z)$ lie in the disk: $\left\{ |z| < 1 + \frac{\log(n-1)}{2^{n-2}} \right\}$. However, these resulsts are not the main focus of this paper. Some additional results proven in [4] concerning cyclotomic factors of $P_n(z)$ are listed below:

Corollary 4.0.1. *[4, p.343]: Let n be such that $n = t(p)pm$ for some prime $p \geq 3$ and integer $m \geq 1$. Then $P_n(z)$ is divisible by all $\Phi_d(z)$ with $d|p(2^{tm} - 1)$ and $p^{u(p)+1}|d$, where $u(p)$ is the highest power of p dividing $2^{t(p)m} - 1$.*

Corollary 4.0.2. *[4, p.344]: If $p \geq 3$ is a prime and $t = t(p)$, $w = w(p)$, then for all $m \geq 1$,*

$$\prod_{p^k|m} \Phi_{p^{k+w+1}}(z) \ \ |P_{tpm}(z). \tag{33}$$

One might wonder how exactly Proposition 2.2 is used implicitly by Mahler. In [2, p.208], Mahler writes the following:

> Denote by $Z_k(z)$ the kth cyclotomic polynomial. It is of degree $\phi(k)$, has rational integral coefficients and highest coefficient 1, has as its roots all the primitive kth roots of unity, and is irreducible over the rational field Q.
>
> By way of example, Fuchs's theorem implies that
>
> $P_6(z)$ is divisible by $Z_9(z)$,
> $P_{12}(z)$ is divisible by $Z_9(z)Z_{45}(z)$,
> $P_{18}(z)$ is divisible by $Z_9(z)Z_{27}(z)Z_{189}(z)$,
> $P_{20}(z)$ is divisible by $Z_{25}(z)Z_{75}(z)$,
> $P_{21}(z)$ is divisible by $Z_{49}(z)$,
>
> etc. I am indebted to D. H. Lehmer for a large table of such factors of which these five cases are the first examples.
>
> We see from these factorizations that $P_6(z)$ has 3 pairs of complex conjugate cyclotomic roots on U, $P_{12}(z)$ has 15 pairs, $P_{18}(z)$ has 66 pairs, $P_{20}(z)$ has 30 pairs, and $P_{21}(z)$ has 21 pairs of such zeros.

In [2, p.208], Mahler has computer-generated evidence that certain $P_n(z)$ are divisible by certain cyclotomic factors. Although Mahler is indeed invoking Proposition 2.2 implicitly by stating that $P_6(z)$, $P_{12}(z)$, etc. are divisible by certain cyclotomic factors, he has made no mistakes, that is, he is not using any results that have not yet been proven. As shown above, Mahler uses known cyclotomic factors of some $P_n(z)$ in order to understand how many pairs of complex-conjugate cyclotomic roots each such $P_n(z)$ has. Proposition 2.2 and Corollary 4.0.2 are especially useful in that they allow one to find cyclotomic factors of a $P_n(z)$ where n is so large that any computer-algebra system will be unable to do so.

5 Acknowledgements

I would like to heartily thank both Will Dana, for introducing me to the abstract algebra needed to make sense out of Proposition 2.2, and my father Paul Pedersen, for sending me notes about [4] and talking to me about the proof of Proposition 2.2.

References

[1] Calvin T. Long. *Elementary Introduction to Number Theory*. 2nd. D.C. Health and Company, 1972.

[2] K. Mahler. “On the Zeros of a Special Sequence of Polynomials”. In: *Mathematics of Computation* 39.159 (July 1982), pp. 207–212. DOI: 10.2307/2007631. URL: http://www.ams.org/journals/mcom/1982-39-159/S0025-5718-1982-0658225-3/S0025-5718-1982-0658225-3.pdf.

[3] David R. Finston and Patrick J. Morandi. “Chapter 5, Quotient Rings and Field Extensions”. In: *An Introduction to Abstract Algebra via Applications*. Department of Mathematical Sciences, New Mexico State University, 2007, pp. 53–61. URL: http://sierra.nmsu.edu/morandi/OldWebPages/Math331Spring2003/.

[4] Karl Dilcher and Larry Ericksen. “The Polynomials of Mahler and Roots of Unity”. In: *The American Mathematical Monthly* 122.4 (2015), pp. 338–353. ISSN: 00029890, 19300972. URL: http://www.jstor.org/stable/10.4169/amer.math.monthly.122.04.338.

[5] Bruce Ikenaga. *Abstract Algebra Course Notes*. URL: http://sites.millersville.edu/bikenaga/abstract-algebra-1/abstract-algebra-1-notes.html.

[6] Eric W. Weisstein. *Cyclotomic Polynomial*. URL: http://mathworld.wolfram.com/CyclotomicPolynomial.html.

[7] Eric W. Weisstein. *Primitive Root of Unity*. URL: http://mathworld.wolfram.com/PrimitiveRootofUnity.html.