# CARLITZ EXTENSIONS

KEITH CONRAD

## 1. Introduction

The ring $\mathbf{Z}$ has many analogies with the ring $\mathbf{F}_p[T]$, where $\mathbf{F}_p$ is a field of prime size $p$. For example, for nonzero $m \in \mathbf{Z}$ and $M \in \mathbf{F}_p[T]$, the residue rings $\mathbf{Z}/(m)$ and $\mathbf{F}_p[T]/M$ are both finite. The unit groups $\mathbf{Z}^\times = \{\pm 1\}$ and $\mathbf{F}_p[T]^\times = \mathbf{F}_p^\times$ are both finite. Every nonzero integer can be made positive after multiplication by a suitable sign $\pm 1$, and every nonzero polynomial in $\mathbf{F}_p[T]$ can be made monic (leading coefficient 1) after multiplication by a suitable nonzero constant in $\mathbf{F}_p^\times$. What we will examine is a deeper analogy: the unit group $(\mathbf{F}_p[T]/M)^\times$ can be interpreted as the Galois group of a suitable extension of the field $\mathbf{F}_p(T)$ in a manner similar to $(\mathbf{Z}/(m))^\times$ being the Galois group of the $m$th cyclotomic extension $\mathbf{Q}(\mu_m)$ of $\mathbf{Q}$, where $\mu_m$ is the group of $m$th roots of unity.

The $m$th roots of unity are the roots of $X^m - 1 \in \mathbf{Z}[X]$, and they form an abelian group under multiplication. We will construct an analogous family of polynomials $[M](X) \in \mathbf{F}_p[T][X]$, parametrized by elements $M$ of $\mathbf{F}_p[T]$ rather than by positive integers, and the roots of each $[M](X)$ will form an $\mathbf{F}_p[T]$-module rather than an abelian group ($\mathbf{Z}$-module). In particular, adjoining the roots of $[M](X)$ to $\mathbf{F}_p(T)$ will yield a field extension of $\mathbf{F}_p(T)$ whose Galois group is isomorphic to $(\mathbf{F}_p[T]/M)^\times$ instead of $(\mathbf{Z}/(m))^\times$.

The polynomials $[M](X)$ and their roots were first introduced by Carlitz [2, 3] in the 1930s. Since Carlitz gave his papers unassuming names (look at the title of [3]), their relevance was not widely recognized until being rediscovered several decades later (*e.g.*, in work of Lubin–Tate in the 1960s and Drinfeld in the 1970s).

## 2. Carlitz polynomials

For each $M \in \mathbf{F}_p[T]$ we will define the *Carlitz polynomial* $[M](X)$ with coefficients in $\mathbf{F}_p[T]$. Our definition will proceed by recursion and linearity. Define $[1](X) := X$ and

$$[T](X) := X^p + TX.$$

For $n \geq 2$, define

$$[T^n](X) := [T]([T^{n-1}](X)).$$

**Example 2.1.** $[T^2](X) = [T]([T](X)) = (X^p + TX)^p + T(X^p + TX) = X^{p^2} + (T^p + T)X^p + T^2X$.

For a general polynomial $M = c_nT^n + \cdots + c_1T + c_0$ in $\mathbf{F}_p[T]$, define $[M](X)$ using $\mathbf{F}_p$-linearity:

$$[M](X) := c_n[T^n](X) + \cdots + c_1[T](X) + c_0X \in \mathbf{F}_p[T][X].$$

**Example 2.2.** For $c \in \mathbf{F}_p$, $[c](X) = cX$ and

$$[T^2 - T](X) = [T^2](X) - [T](X) = X^{p^2} + (T^p + T - 1)X^p + (T^2 - T)X.$$

**Remark 2.3.** The Carlitz polynomials $[M](X)$ have many notations in the literature: $\rho_M(X)$, $\phi_M(X)$, $C_M(X)$, $\omega_M(X)$ (Carlitz's original notation[1]), and $X^M$. The notation $[M](X)$ used here for Carlitz polynomials is taken from the setting of Lubin-Tate formal groups, where $[\ ](X)$ is a standard notation.

Our examples suggest general properties of $[M](X)$. For instance, in both $[T^2](X)$ and $[T^2+T](X)$ we only see $X$-terms with $p$-power exponents, the lowest degree $X$-terms are, respectively, $T^2X$ and $(T^2+T)X$, and both polynomials have $X$-degree $p^2$.

**Theorem 2.4.** *For nonzero $M \in \mathbf{F}_p[T]$, $[M](X)$ has $X$-degree $p^{\deg M}$. Moreover, as a polynomial in $X$, $[M](X)$ is a "$p$-polynomial":*

$$[M](X) = \sum_{j=0}^{\deg M} a_j(T)X^{p^j} = (\operatorname{lead} M)X^{p^{\deg M}} + \cdots + MX,$$

*where $a_j(T) \in \mathbf{F}_p[T]$ with $a_0(T) = M$ and $a_{\deg M}(T) = \operatorname{lead} M \in \mathbf{F}_p$ being the leading coefficient of $M$.*

*Proof.* This can be proved for $M = T^n$ by induction on $n$ and then for all $M$ by $\mathbf{F}_p$-linearity. □

The coefficients $a_j(T)$ will be examined closely in Section 6. They are analogues of binomial coefficients.

**Corollary 2.5.** *For $M \in \mathbf{F}_p[T]$, indeterminates $X$ and $Y$, and $c \in \mathbf{F}_p$,*

$$[M](X+Y) = [M](X) + [M](Y) \text{ and } [M](cX) = c[M](X).$$

*For $M_1$ and $M_2$ in $\mathbf{F}_p[T]$,*

$$[M_1+M_2](X) = [M_1](X) + [M_2](X) \text{ and } [M_1M_2](X) = [M_1]([M_2](X)).$$

*In particular, if $D|M$ in $\mathbf{F}_p[T]$ then $[D](X)|[M](X)$ in $\mathbf{F}_p[T][X]$.*

*Proof.* The basic polynomial $[T](X) = X^p + TX$ is $\mathbf{F}_p$-linear in $X$, and since other $[M](X)$ are defined by iteration and $\mathbf{F}_p$-linearity from $[T](X)$, we get $\mathbf{F}_p$-linearity of every $[M](X)$. That $[M](X)$ is not only $\mathbf{F}_p$-linear in $X$ but additive and multiplicative in $M$ can be proved by induction on the degree of $M$.

The last part is the analogue of $d|m$ implying $(X^d-1)|(X^m-1)$ in $\mathbf{Z}[X]$ and is left to the reader. (Hint: $[M](X)$ is divisible by $X$.) □

Each $X^m - 1$ is separable over $\mathbf{Q}$ since it has no root in common with its derivative $mX^{m-1}$, so there are $m$ different $m$th roots of unity in characteristic 0. The polynomial $[T](X) = X^p + TX$ is separable over $\mathbf{F}_p(T)$, since its $X$-derivative is $T$, which is a nonzero constant as a polynomial in $X$, so $([T](X), [T]'(X)) = 1$ in $\mathbf{F}_p(T)[X]$. A similar calculation shows

**Theorem 2.6.** *For nonzero $M$ in $\mathbf{F}_p[T]$, $[M](X)$ is separable in $\mathbf{F}_p(T)[X]$.*

*Proof.* Since $[M](X)$ is a $p$-polynomial in $X$ with lowest degree term $MX$, the $X$-derivative of $[M](X)$ is $M \in \mathbf{F}_p[T]$, which is a nonzero constant in $\mathbf{F}_p(T)[X]$. Therefore $[M](X)$ is relatively prime to its $X$-derivative, so it is separable as a polynomial in $X$. □

[1]Writing $[M](X)$ as $[M](X,T)$ to make its dependence on $T$ more visible, Carlitz's $\omega_M(X)$ is actually $[M](X,-T)$, *e.g.*, $\omega_T(X) = X^p - TX$ rather than $X^p + TX$.

**Corollary 2.7.** *For nonzero $M$ and $N$ in $\mathbf{F}_p[T]$, $[M](X)$ and $[N](X)$ have the same roots if and only if $M = cN$ for some $c \in \mathbf{F}_p^\times$.*

*Proof.* If $M = cN$ then $[M](X) = c[N](X)$, so $[M](X)$ and $[N](X)$ have the same roots. Conversely, assume $[M](X)$ and $[N](X)$ have the same roots. We will show $M|N$ and $N|M$, so $M$ and $N$ are equal up to a scaling factor in $\mathbf{F}_p^\times$.

Write $N = MQ + R$ where $R = 0$ or $\deg R < \deg M$. If $R \neq 0$ then for any root $\lambda$ of $[M](X)$ we have $[M](\lambda) = 0$ and $[N](\lambda) = 0$, so

$$0 = [MQ + R](\lambda) = [Q]([M](\lambda)) + [R](\lambda) = [Q](0) + [R](\lambda) = [R](\lambda).$$

Therefore the number of roots of $[R](X)$ is at least the number of roots of $[M](X)$. By Theorem 2.6, the number of roots of $[M](X)$ is $\deg([M](X)) = p^{\deg M}$, so $p^{\deg M} \leq \deg([R](X)) = p^{\deg R}$, so $\deg M \leq \deg R$. This contradicts the inequality $\deg R < \deg M$, so $R = 0$ and $M|N$. The argument that $N|M$ is similar, so we're done. □

The rest of this section concerns analogies between the $p$th power map for prime $p$ and the polynomial $[\pi](X)$ for (monic) irreducible $\pi$ in $\mathbf{F}_p[T]$.

Since $(X^m - 1)' = mX^{m-1}$, in $(\mathbf{Z}/(p))[X]$ the polynomial $X^m - 1$ is separable if $(m, p) = 1$ while $X^p - 1 \equiv (X - 1)^p \bmod p$. Analogously, what can be said about the reduction $[M](X) \bmod \pi$ in $(\mathbf{F}_p[T]/\pi)[X]$?

**Theorem 2.8.** *Let $\pi$ be monic irreducible in $\mathbf{F}_p[T]$ and set $\mathbf{F}_\pi = \mathbf{F}_p[T]/(\pi)$. For $M$ in $\mathbf{F}_p[T]$, let $\overline{[M]}(X) \in \mathbf{F}_\pi[X]$ be the result of reducing the coefficients of $[M](X)$ modulo $\pi$. If $(M, \pi) = 1$ then $\overline{[M]}(X)$ is separable in $\mathbf{F}_\pi[X]$, while $\overline{[\pi]}(X) = X^{p^{\deg \pi}}$.*

*Proof.* If $(M, \pi) = 1$, $\overline{[M]}'(X) = M \bmod \pi$ is a nonzero constant as a polynomial in $X$, so $\overline{[M]}(X)$ is separable over $\mathbf{F}_\pi$. On the other hand, $\overline{[\pi]}'(X) = \pi \bmod \pi$ is 0, so $\overline{[\pi]}(X)$ is inseparable in $\mathbf{F}_\pi[X]$. Since $[\pi](X)$ has degree $p^{\deg \pi}$ and is monic (because $\pi$ is), its reduction $\overline{[\pi]}(X)$ in $\mathbf{F}_\pi[X]$ is monic with degree $p^{\deg \pi}$. Therefore we can show $\overline{[\pi]}(X) = X^{p^{\deg \pi}}$ by showing the only root of $\overline{[\pi]}(X)$ in the algebraic closure $\overline{\mathbf{F}}_\pi$ is 0.

Suppose there is a root $\gamma$ of $\overline{[\pi]}(X)$ in $\overline{\mathbf{F}}_\pi$ with $\gamma \neq 0$. We will get a contradiction. For any $M \in \mathbf{F}_p[T]$, $[M](\gamma)$ is a root of $\overline{[\pi]}(X)$ because $[\pi]([M](\gamma)) = [\pi M](\gamma) = [M]([\pi](\gamma)) = [M](0) = 0$. Therefore the number of roots of $\overline{[\pi]}(X)$ in $\overline{\mathbf{F}}_\pi$ is at least the number of different values of $[M](\gamma)$ as $M$ varies. To count this, consider the map $\mathbf{F}_p[T] \to \overline{\mathbf{F}}_\pi$ given by $M \mapsto [M](\gamma)$. By Corollary 2.5, this is additive with kernel

$$\{M \in \mathbf{F}_p[T] : [M](\gamma) = 0\}.$$

This kernel is not only a subgroup of $\mathbf{F}_p[T]$ but an ideal: if $[M](\gamma) = 0$ and $N \in \mathbf{F}_p[T]$ then $[NM](\gamma) = [N]([M](\gamma)) = [N](0) = 0$. This ideal is proper (since $[1](\gamma) = \gamma \neq 0$) and contains $\pi$. Since $(\pi)$ is a maximal ideal, the kernel is $(\pi)$, so the number of $[M](\gamma)$ as $M$ varies is $\#(\mathbf{F}_p[T]/\pi) = p^{\deg \pi} = \deg \overline{[\pi]}(X)$. Therefore $\overline{[\pi]}(X)$ has as many roots in $\overline{\mathbf{F}}_\pi$ as its degree, but it is inseparable: contradiction! So the only root of $\overline{[\pi]}(X)$ in $\overline{\mathbf{F}}_\pi$ is 0. □

**Corollary 2.9.** *For any irreducible $\pi \in \mathbf{F}_p[T]$, the coefficients of $[\pi](X)$ besides its leading term are multiples of $\pi$. In particular, $[\pi](X)/X$ is an Eisenstein polynomial with respect to $\pi$ with constant term $\pi$.*

*Proof.* For $c \in \mathbf{F}_p^\times$, $[c\pi](X) = c[\pi](X)$, so we may assume $\pi$ is monic. Then the leading term of $[\pi](X)$ in $\mathbf{F}_p[T][X]$ is $X^{p^{\deg \pi}}$ and by Theorem 2.8, $\overline{[\pi]}(X) = X^{p^{\deg \pi}}$ in $(\mathbf{F}_p[T]/\pi)[X]$.

Lifting this equation to $\mathbf{F}_p[T][X]$ shows all lower-degree coefficients of $[\pi](X)$ are multiples of $\pi$. Since the lowest degree term of $[\pi](X)$ is $\pi X$, $[\pi](X)/X$ has constant term $\pi$ and therefore is Eisenstein with respect to $\pi$. □

**Remark 2.10.** In many respects, $[M](X)$ is analogous not to $X^m-1$ but to $(1+X)^m-1$. For example, $(1+X)^m-1 = X^m+\cdots+mX$ has lowest degree term $mX$ and $[M](X)$ has lowest degree term $MX$. If we write $[m](X) = (1+X)^m - 1$ then $[m_1m_2](X) = [m_1]([m_2](X))$, which resembles part of Corollary 2.5 (but $[m](X)$ is *not* additive in $m$) and Corollary 2.9 resembles $[p](X)/X = ((1+X)^p - 1)/X$ being Eisenstein for prime $p$.

**Corollary 2.11.** *For any irreducible $\pi \in \mathbf{F}_p[T]$ and integer $k \geq 1$, the coefficients of $[\pi^k](X)$ besides its leading term are multiples of $\pi$.*

*Proof.* Induct on $k$ and use the identity $[\pi^k](X) = [\pi]([\pi^{k-1}](X))$. □

**Theorem 2.12.** *For any monic irreducible $\pi$ in $\mathbf{F}_p[T]$, $[\pi](A) \equiv A \bmod \pi$ for all $A \in \mathbf{F}_p[T]$.*

This is an analogue of $a^p \equiv a \bmod p$ for (positive) prime $p$ and all $a \in \mathbf{Z}$.

*Proof.* By Theorem 2.8, $\overline{[\pi]}(X) = X^{p^{\deg \pi}}$ in $(\mathbf{F}_p[T]/\pi)[X]$. Thus $[\pi](A) \equiv A^{p^{\deg \pi}} \bmod \pi$ for all $A \in \mathbf{F}_p[T]$. Since $\mathbf{F}_p[T]/\pi$ is a field of size $p^{\deg \pi}$, raising to this power on the field is the identity map, so $[\pi](A) \equiv A \bmod \pi$. □

Subtracting $A = [1](A)$ from both sides of the congruence in Theorem 2.12, we get

**Corollary 2.13.** *For any monic irreducible $\pi$ in $\mathbf{F}_p[T]$, $[\pi - 1](A) \equiv 0 \bmod \pi$ for all $A \in \mathbf{F}_p[T]$.*

This is an analogue of Fermat's little theorem: $a^{p-1} \equiv 1 \bmod p$ for (positive) prime $p$ and $a$ in $(\mathbf{Z}/(p))^\times$. That Fermat's theorem says the $(p-1)$-th power map is identically 1 on the multiplicative group $(\mathbf{Z}/(p))^\times$ while Corollary 2.13 says $[\pi - 1]$ acts as 0 on $\mathbf{F}_p[T]/\pi$ illustrates how analogues of $X^m - 1$ for Carlitz polynomials are additive rather than multiplicative.

Here is an analogue of $f(X^p) \equiv f(X)^p \bmod p$ for $f(X) \in \mathbf{Z}[X]$.

**Theorem 2.14.** *For monic irreducible $\pi$ in $\mathbf{F}_p[T]$ and $f(X) \in \mathbf{F}_p[T][X]$, $f([\pi](X)) \equiv f(X)^{p^{\deg \pi}} \bmod \pi$, where the congruence means coefficients of like powers of $X$ on both sides are equal in $\mathbf{F}_p[T]/\pi$.*

*Proof.* In $(\mathbf{F}_p[T]/\pi)[X]$, $\overline{[\pi]}(X) = X^{p^{\deg \pi}}$ (all the lower degree coefficients vanish modulo $\pi$), so $f([\pi](X)) \equiv f(X^{p^{\deg \pi}}) \bmod \pi$. In $\mathbf{F}_p[T]/\pi$ every element is its own $p^{\deg \pi}$th power, so $f(X)^{p^{\deg \pi}} \equiv f(X^{p^{\deg \pi}}) \bmod \pi$. □

If $\pi$ is not monic then the above results have a more awkward form. Letting $c$ be the leading coefficient of $\pi$, $[\pi](A) \equiv cA \bmod \pi$, $[\pi - c](A) \equiv 0 \bmod \pi$, and $f([\pi](X)) \equiv f(cX)^{p^{\deg \pi}} \bmod \pi$. Just remember that monic $\pi$ have nicer formulas.

Notationally, it is convenient to regard $p^{\deg M}$ as the analogue of the absolute value of an integer. Indeed, for nonzero $m \in \mathbf{Z}$ we have $|m| = \#(\mathbf{Z}/(m))$ and for nonzero $M \in \mathbf{F}_p[T]$ we have $p^{\deg M} = \#(\mathbf{F}_p[T]/M)$. Set

$$\mathrm{N}(M) = p^{\deg M} = \#(\mathbf{F}_p[T]/M).$$

With this notation, we set some formulas in $\mathbf{Z}[X]$ and $\mathbf{F}_p[T][X]$ side by side:

$$(1+X)^m - 1 = X^m + \cdots + mX, \quad [M](X) = X^{\mathrm{N}(M)} + \cdots + MX,$$

$$(1+X)^p - 1 \equiv X^p \bmod p, \quad [\pi](X) \equiv X^{\mathrm{N}(\pi)} \bmod \pi,$$
$$f(X^p) \equiv f(X)^p \bmod p, \quad f([\pi](X)) \equiv f(X)^{\mathrm{N}(\pi)} \bmod \pi.$$

Here $m$ and $p$ are positive while $M$ and $\pi$ are monic.

## 3. Carlitz torsion

Let $K$ be a field extension of $\mathbf{F}_p(T)$. Multiplication in $K$ lets us think about $K$ as an $\mathbf{F}_p(T)$-vector space, so in particular as an $\mathbf{F}_p[T]$-module. Using the Carlitz polynomials we can define a new and non-obvious $\mathbf{F}_p[T]$-module structure on $K$, as follows.

**Definition 3.1.** Let $K$ be a field extension of $\mathbf{F}_p(T)$. We make $K$ into an $\mathbf{F}_p[T]$-module by letting $\mathbf{F}_p[T]$ act on $K$ through the Carlitz polynomials:

$$M \cdot \alpha = [M](\alpha)$$

for $M \in \mathbf{F}_p[T]$ and $\alpha \in K$. This is called the *Carlitz action* of $\mathbf{F}_p[T]$ on $K$.

**Example 3.2.** In the Carlitz action, $T \cdot \alpha = [T](\alpha) = \alpha^p + T\alpha \in K$.

The verification that the Carlitz action really is an $\mathbf{F}_p[T]$-module structure on $K$ requires checking several identities:

$$1 \cdot \alpha = \alpha, \quad M \cdot (\alpha + \beta) = M \cdot \alpha + M \cdot \beta$$

and

$$(M_1 + M_2) \cdot \alpha = M_1 \cdot \alpha + M_2 \cdot \alpha, \quad M_1 \cdot (M_2 \cdot \alpha) = (M_1 M_2) \cdot \alpha.$$

The first identity follows from the definition $[1](X) = X$. The remaining identities follow from Corollary 2.5 by specializing $X$ and $Y$ to elements of $K$. For example, since $[M](X + Y) = [M](X) + [M](Y)$ in $\mathbf{F}_p[T][X, Y]$, upon specialization of $X$ and $Y$ to $\alpha$ and $\beta$ in $K$ we get $[M](\alpha + \beta) = [M](\alpha) + [M](\beta)$, which says $M \cdot (\alpha + \beta) = M \cdot \alpha + M \cdot \beta$.

In case of ambiguity, we write $C(K)$ rather than $K$ to stress that $K$ is being viewed as an $\mathbf{F}_p[T]$-module through the Carlitz polynomials. The case $K = \overline{\mathbf{F}_p(T)}$, an algebraic closure of $\mathbf{F}_p(T)$, will be of particular importance.

The $\mathbf{F}_p[T]$-module $C(\overline{\mathbf{F}_p(T)})$, which is called the *Carlitz module*, is analogous to the multiplicative group $\overline{\mathbf{Q}}^\times$ as a $\mathbf{Z}$-module: $m \in \mathbf{Z}$ acts on $\alpha \in \overline{\mathbf{Q}}^\times$ by $\alpha \mapsto \alpha^m$ and $M \in \mathbf{F}_p[T]$ acts on $\alpha \in \overline{\mathbf{F}_p(T)}$ by $\alpha \mapsto [M](\alpha)$. The torsion elements in the $\mathbf{Z}$-module $\overline{\mathbf{Q}}^\times$ are the $\alpha \in \overline{\mathbf{Q}}^\times$ satisfying $\alpha^m = 1$ for some $m$; these are the roots of unity and they generate abelian extensions of $\mathbf{Q}$. The torsion elements of $C(\overline{\mathbf{F}_p(T)})$ are the $\alpha \in \overline{\mathbf{F}_p(T)}$ satisfying $[M](\alpha) = 0$ for some $M$, and we will see in Section 4 that they generate abelian extensions of $\mathbf{F}_p(T)$.

**Definition 3.3.** Let $\Lambda_M = \{\lambda \in \overline{\mathbf{F}_p(T)} : [M](\lambda) = 0\}$.

**Example 3.4.** Take $M = T$. Since $[T](X) = X^p + TX = X(X^{p-1} + T)$,

$$\Lambda_T = \{\lambda \in \overline{\mathbf{F}_p(T)} : \lambda^p + T\lambda = 0\} = \{0\} \cup \{c\sqrt[p-1]{-T} : c \in \mathbf{F}_p^\times\},$$

where $\sqrt[p-1]{-T}$ denotes some $(p-1)$-th root of $-T$. Notice $\mathbf{F}_p(T, \Lambda_T)/\mathbf{F}_p(T)$ is a Kummer extension of degree $p - 1$.

We know by Theorem 2.6 that $[M](X)$ has $p^{\deg M}$ roots in $\overline{\mathbf{F}_p(T)}$, so $\#\Lambda_M = p^{\deg M}$. Since $[M](X)$ has constant term 0, $0 \in \Lambda_M$ for all $M$. This is analogous to 1 being in $\mu_m$ for all $m$. Since $[M](X)$ is $\mathbf{F}_p$-linear in $X$, $\Lambda_M$ is a (finite) $\mathbf{F}_p$-vector space. But $\Lambda_M$ is more than a vector space:

**Theorem 3.5.** *The set $\Lambda_M$ is a submodule of $C(\overline{\mathbf{F}_p(T)})$: if $\lambda \in \Lambda_M$ and $A \in \mathbf{F}_p[T]$ then $[A](\lambda) \in \Lambda_M$.*

*Proof.* For $A \in \mathbf{F}_p[T]$ and $\lambda \in \Lambda_M$, $[A](\lambda) \in \Lambda_M$ since, using the last identity in Corollary 2.5,

$$[M]([A](\lambda)) = [MA](\lambda) = [A]([M](\lambda)) = [A](0) = 0.$$

Thus $\Lambda_M$ is a submodule of $C(\overline{\mathbf{F}_p(T)})$. □

**Example 3.6.** In Example 3.4 we saw $\Lambda_T = \mathbf{F}_p\,\sqrt[p-1]{-T}$, which is a 1-dimensional $\mathbf{F}_p$-vector space. The Carlitz action of $A \in \mathbf{F}_p[T]$ on $\lambda \in \Lambda_T$ is through multiplication by the constant term of $A$: writing $A = TQ + A(0)$,

$$[A](\lambda) = [TQ + A(0)](\lambda) = [Q]([T](\lambda)) + [A(0)]\lambda = A(0)\lambda.$$

The group structure on $\mu_m$ makes it not only a $\mathbf{Z}$-module but a $\mathbf{Z}/(m)$-module since, for $\zeta \in \mu_m$, $\zeta^a = \zeta^b$ when $a \equiv b \bmod m$. (Conversely, if $\zeta^a = \zeta^b$ for all $\zeta \in \mu_m$ then $a \equiv b \bmod m$.) The group $\mu_m$ is cyclic, and if $\mu_m = \langle\zeta\rangle$ then $\zeta^a$ is a generator if and only if $(a, m) = 1$. Exactly the same properties apply to $\Lambda_M$:

**Theorem 3.7.** *For $A$ and $B$ in $\mathbf{F}_p[T]$ and $\lambda \in \Lambda_M$, if $A \equiv B \bmod M$ then $[A](\lambda) = [B](\lambda)$, so the Carlitz action on $\Lambda_M$ makes it an $\mathbf{F}_p[T]/M$-module. Conversely, if $[A](\lambda) = [B](\lambda)$ for all $\lambda \in \Lambda_M$ then $A \equiv B \bmod M$. There exists a $\lambda_0 \in \Lambda_M$ which is a Carlitz generator:*

$$\Lambda_M = \{[A](\lambda_0) : A \in \mathbf{F}_p[T]/M\},$$

*and the generators of $\Lambda_M$ are precisely the $[A](\lambda_0)$ where $(A, M) = 1$.*

*Proof.* Writing $A = B + MN$,

$$[A](\lambda) = [B + MN](\lambda) = [B](\lambda) + [N]([M](\lambda)) = [B](\lambda) + [N](0) = [B](\lambda).$$

To show that if $[A](\lambda) = [B](\lambda)$ for all $\lambda \in \Lambda_M$ then $A \equiv B \bmod M$, we can subtract to reduce ourselves to showing that if $[A](\lambda) = 0$ for all $\lambda \in \Lambda_M$ then $A \equiv 0 \bmod M$. Write $A = MQ + R$ where $R = 0$ or $\deg R < \deg M$. Then for all $\lambda \in \Lambda_M$,

$$0 = [A](\lambda) = [Q]([M](\lambda)) + [R](\lambda) = [Q](0) + [R](\lambda) = [R](\lambda).$$

If $R \neq 0$, the Carlitz polynomial $[R](X)$ has degree $p^{\deg R} < p^{\deg M} = \#\Lambda_M$, so $[R]$ has more roots than its degree. This is impossible, so $R = 0$ and $M|A$.

To prove $\Lambda_M$ has a generator as an $\mathbf{F}_p[T]$-module, we adapt the proof that $\mu_m$ is cyclic based on looking for elements of maximal possible order. For each $\lambda \in \Lambda_M$, its (Carlitz) annihilators

$$\mathrm{Ann}_C(\lambda) = \{A \in \mathbf{F}_p[T] : [A](\lambda) = 0\}$$

form an ideal in $\mathbf{F}_p[T]$ containing $M$, so $\mathrm{Ann}_C(\lambda) = (D)$ for a unique monic divisor $D$ of $M$. Call $D$ the *Carlitz order* of $\lambda$. Among all Carlitz orders of elements of $\Lambda_M$, let $L$ be a Carlitz order of largest degree. As with torsion abelian groups, all Carlitz orders divide $L$ (use division in $\mathbf{F}_p[T]$ rather than in $\mathbf{Z}$ to check this), so $L$ is the unique Carlitz order of largest degree and all elements of $\Lambda_M$ are roots of $[L](X)$. Therefore $\#\Lambda_M \leq \deg([L](X))$, so $\deg M \leq \deg L$. We also have the reverse inequality since $L|M$, so $L$ and $M$ are scalar multiples of each other. This means there is a $\lambda_0 \in \Lambda_M$ whose annihilator ideal is $(L) = (M)$, so the submodule it generates in $\Lambda_M$ has size

$$\#\{[A](\lambda_0) : A \in \mathbf{F}_p[T]\} = \#(\mathbf{F}_p[T]/M) = p^{\deg M} = \#\Lambda_M,$$

which shows $\lambda_0$ is a generator of $\Lambda_M$ and there is an $\mathbf{F}_p[T]$-module isomorphism $\mathbf{F}_p[T]/M \cong \Lambda_M$ given by $A \bmod M \mapsto [A](\lambda_0)$, with standard multiplication by $\mathbf{F}_p[T]$ on the left and the Carlitz action on the right. In particular, $[A](\lambda_0)$ generates $\Lambda_M$ using the Carlitz action if and only if $A \bmod M$ generates $\mathbf{F}_p[T]/M$ as an $\mathbf{F}_p[T]$-module in the usual way, and that occurs if and only if $(A, M) = 1$. □

To stress the similarities again, choosing a generator $\zeta$ of $\mu_m$ gives a group isomorphism $\mathbf{Z}/(m) \cong \mu_m$ by $a \bmod m \mapsto \zeta^a$, and in the same way choosing a generator $\lambda_0$ of $\Lambda_M$ leads to an $\mathbf{F}_p[T]$-module isomorphism $\mathbf{F}_p[T]/M \cong \Lambda_M$ by $A \bmod M \mapsto [A](\lambda_0)$, where $\mathbf{F}_p[T]/M$ is an $\mathbf{F}_p[T]$-module by standard multiplication.

The part of the proof of Theorem 3.7 dealing with the annihilator $\mathrm{Ann}_C(\lambda)$ looks similar to our discussion of the kernel in the proof of Theorem 2.8. This is not a coincidence, and the rest of this section (which we won't use later) explains a common viewpoint.

Given any $\mathbf{F}_p[T]$-algebra $\mathcal{A}$ (such as a field extension of $\mathbf{F}_p(T)$ or a quotient ring of $\mathbf{F}_p[T]$), it makes sense to evaluate polynomials in $\mathbf{F}_p[T][X]$ at elements of $\mathcal{A}$, so setting $M \cdot a = [M](a)$ makes $\mathcal{A}$ into an $\mathbf{F}_p[T]$-module in a new way. When $\mathcal{A}$ is considered with this $\mathbf{F}_p[T]$-module structure we denote it as $C(\mathcal{A})$. That is, $C(\mathcal{A})$ is $\mathcal{A}$ as an additive group but it has a new $\mathbf{F}_p[T]$-module action through the use of Carlitz polynomials acting on $\mathcal{A}$. (One might call $C(\mathcal{A})$ the "Carlitzification" of $\mathcal{A}$.) The proof of Theorem 2.8 was really treating $\overline{\mathbf{F}}_\pi$ as $C(\overline{\mathbf{F}}_\pi)$ without explicitly saying so (with $\overline{\mathbf{F}}_\pi$ being an $\mathbf{F}_p[T]$-algebra by virtue of it being an extension of the field $\mathbf{F}_\pi = \mathbf{F}_p[T]/\pi$).

Passing from $\mathbf{F}_p[T]$-algebras $\mathcal{A}$ to $\mathbf{F}_p[T]$-modules $C(\mathcal{A})$ behaves well on maps: if $f\colon \mathcal{A} \to \mathcal{B}$ is an $\mathbf{F}_p[T]$-algebra homomorphism, then $f([M](a)) = [M](f(a))$ so $f\colon C(\mathcal{A}) \to C(\mathcal{B})$ is an $\mathbf{F}_p[T]$-module homomorphism. Thus the Carlitz construction is really a functor, from $\mathbf{F}_p[T]$-algebras to $\mathbf{F}_p[T]$-modules, and is analogous to the "unit" functor taking each commutative ring (a $\mathbf{Z}$-algebra) to its unit group (a $\mathbf{Z}$-module).

**Example 3.8.** For irreducible $\pi$ in $\mathbf{F}_p[T]$, $C(\mathbf{F}_p[T]/\pi)$ is $\mathbf{F}_p[T]/\pi$ with the $\mathbf{F}_p[T]$-module structure given by $M \cdot (A \bmod \pi) = [M](A) \bmod \pi$. Viewing $\mathbf{F}_p[T]/\pi$ as an $\mathbf{F}_p[T]$-module by ordinary multiplication instead, $\pi$ acts as 0: $\pi f \equiv 0 \bmod \pi$. But in the $\mathbf{F}_p[T]$-module $C(\mathbf{F}_p[T]/\pi)$, $\pi$ acts as the *identity*, since $[\pi](A) \equiv A \bmod \pi$.

The group $(\mathbf{F}_p[T]/\pi)^\times$ is cyclic and also the module $C(\mathbf{F}_p[T]/\pi)$ is cyclic, but these mean different things, *e.g.* 1 is a generator of $C(\mathbf{F}_3[T]/(T^2+1))$ since

$$\{[M](1) \bmod T^2+1 : M \in \mathbf{F}_3[T]\} = \mathbf{F}_3[T]/(T^2+1).$$

Data is collected in Table 1. Of course, 1 certainly is not a generator of the group $(\mathbf{F}_3[T]/(T^2+1))^\times$.

The group $(\mathbf{F}_p[T]/\pi)^\times$ and the $\mathbf{F}_p[T]$-module $C(\mathbf{F}_p[T]/\pi)$ present us with two ways to extend results about the groups $(\mathbf{Z}/(p))^\times$ to the polynomial setting. For example, a classical conjecture of Artin predicts when an integer generates $(\mathbf{Z}/(p))^\times$ for infinitely many primes $p$, and this has two polynomial analogues: when does a polynomial in $\mathbf{F}_p[T]$ generate infinitely many of the groups $(\mathbf{F}_p[T]/\pi)^\times$ or infinitely many of the modules $C(\mathbf{F}_p[T]/\pi)$? Both questions have good answers; see [5] and [7, Chap. 10].

In another direction, for any prime $p \neq 2$ the groups $(\mathbf{Z}/(p^k))^\times$ are cyclic for all $k \geq 1$, but the groups $(\mathbf{F}_p[T]/\pi^k)^\times$ are *not* cyclic for any $k \geq 4$. (It is cyclic for $k = 2$ if $\deg \pi = 1$ and for $k = 3$ if $p = 2$ and $\deg \pi = 1$, but not otherwise.) So it looks like analogy between $\mathbf{Z}$ and $\mathbf{F}_p[T]$ breaks down. But we can recover a good analogy using the Carlitz action: $C(\mathbf{F}_p[T]/\pi^k)$ is a cyclic $\mathbf{F}_p[T]$-module for every $k \geq 1$.

| $M$ | $[M](1) \bmod T^2+1$ |
|---|---|
| $0$ | $0$ |
| $1$ | $1$ |
| $2$ | $2$ |
| $T$ | $T+1$ |
| $T+1$ | $T+2$ |
| $T+2$ | $T$ |
| $2T$ | $2T+2$ |
| $2T+1$ | $2T$ |
| $2T+2$ | $2T+1$ |

TABLE 1. Carlitz action on 1 in $\mathbf{F}_3[T]/(T^2+1)$

## 4. Carlitz extensions

We now adjoin $\Lambda_M$ to $\mathbf{F}_p(T)$ to produce abelian extensions, just as $\mathbf{Q}(\mu_m)$ is an abelian extension of $\mathbf{Q}$. Throughout this section, we write $\mathbf{F}_p(T)$ as $F$, so $\mathbf{F}_p(T, \Lambda_M) = F(\Lambda_M)$. In the literature the fields $F(\Lambda_M)$ are called "cyclotomic function fields" (see [7, Chap. 12]) because they share many similar properties with the usual cyclotomic fields $\mathbf{Q}(\mu_m)$.

Since $[M](X)$ is separable in $F[X]$, adjoining its roots $\Lambda_M$ to $F$ gives a Galois extension of $F$. We only need to adjoin a generator of $\Lambda_M$ to $F$, since the other elements of $\Lambda_M$ are polynomials in the generator (with $\mathbf{F}_p[T]$-coefficients). Each element of $\mathrm{Gal}(F(\Lambda_M)/F)$ permutes the roots $\Lambda_M$ of $[M](X)$ and is determined as a field automorphism by its effect on these roots. Keeping in mind that each element of $\mathrm{Gal}(\mathbf{Q}(\mu_m)/\mathbf{Q})$ is determined by the unique exponent in $(\mathbf{Z}/(m))^\times$ by which they act on all the $m$th roots of unity, we anticipate that each element of $\mathrm{Gal}(F(\Lambda_M)/F)$ acts on $\Lambda_M$ by a Carlitz polynomial. To make this explicit, we use a generator of $\Lambda_M$.

Choose $\sigma \in \mathrm{Gal}(F(\Lambda_M)/F)$. Letting $\lambda_0$ be a generator of $\Lambda_M$,

$$\Lambda_M = \sigma(\Lambda_M) = \sigma(\{[N](\lambda_0) : N \in \mathbf{F}_p[T]\}) = \{[N](\sigma(\lambda_0)) : N \in \mathbf{F}_p[T]\},$$

so $\sigma(\lambda_0)$ is also a generator of $\Lambda_M$: we can write $\sigma(\lambda_0) = [A](\lambda_0)$ for some $A$ in $\mathbf{F}_p[T]$, well-defined modulo $M$, with $(A, M) = 1$ (Theorem 3.7). That $\sigma$ acts like $A$ on $\lambda_0$ propagates to all of $\Lambda_M$: any $\lambda \in \Lambda_M$ has the form $[N](\lambda_0)$ for some $N \in \mathbf{F}_p[T]$, so

$$\sigma(\lambda) = \sigma([N](\lambda_0)) = [N](\sigma(\lambda_0)) = [N]([A](\lambda_0)) = [A]([N](\lambda_0)) = [A](\lambda).$$

Thus $\sigma$ has the same effect by the Carlitz action on all the elements of $\Lambda_M$. Write $A$ as $A_\sigma$ to indicate its dependence on $\sigma$: to each $\sigma \in \mathrm{Gal}(F(\Lambda_M)/F)$ we get a unit $A_\sigma \in (\mathbf{F}_p[T]/M)^\times$ which describes through its Carlitz polynomial how $\sigma$ permutes the elements of $\Lambda_M$.

**Theorem 4.1.** *The map $\sigma \mapsto A_\sigma$ is an injective group homomorphism $\mathrm{Gal}(F(\Lambda_M)/F) \hookrightarrow (\mathbf{F}_p[T]/M)^\times$.*

*Proof.* For $\sigma$ and $\tau$ in $\mathrm{Gal}(F(\Lambda_M)/F)$ and any $\lambda \in \Lambda_M$, $(\sigma\tau)(\lambda)$ equals

$$\sigma(\tau(\lambda)) = \sigma([A_\tau](\lambda)) = [A_\tau](\sigma(\lambda)) = [A_\tau]([A_\sigma](\lambda)) = [A_\tau A_\sigma](\lambda).$$

Also $(\sigma\tau)(\lambda) = [A_{\sigma\tau}](\lambda)$, so $A_{\sigma\tau}$ and $A_\tau A_\sigma = A_\sigma A_\tau$ have the same Carlitz action on $\Lambda_M$. Therefore $A_{\sigma\tau} \equiv A_\sigma A_\tau \bmod M$ (Theorem 3.7), which shows we have a homomorphism from $\mathrm{Gal}(F(\Lambda_M)/F)$ to $(\mathbf{F}_p[T]/M)^\times$.

When $\sigma$ is in the kernel, $A_\sigma \equiv 1 \bmod M$, so for all $\lambda \in \Lambda_M$ we have $\sigma(\lambda) = [A_\sigma](\lambda) = [1](\lambda) = \lambda$. Therefore $\sigma$ is the identity on $\Lambda_M$, so $\sigma$ is the identity in $\mathrm{Gal}(F(\Lambda_M)/F)$. □

Since $(\mathbf{F}_p[T]/M)^\times$ is abelian, $\mathrm{Gal}(F(\Lambda_M)/F)$ is abelian, so Carlitz extensions of $F = \mathbf{F}_p(T)$ are abelian extensions.

**Theorem 4.2.** *The embedding* $\mathrm{Gal}(F(\Lambda_M)/F) \hookrightarrow (\mathbf{F}_p[T]/M)^\times$ *is an isomorphism.*

*Proof.* We will adapt the proof of the analogous result that $\mathrm{Gal}(\mathbf{Q}(\mu_m)/\mathbf{Q})$ is isomorphic to $(\mathbf{Z}/(m))^\times$, taken from [6, p. 278].

Pick a generator $\lambda_0$ of $\Lambda_M$. The generators of $\Lambda_M$ are all $[A](\lambda_0)$ for $(A, M) = 1$, so surjectivity of $\mathrm{Gal}(F(\Lambda_M)/F) \to (\mathbf{F}_p[T]/M)^\times$ is equivalent to showing all generators of $\Lambda_M$ are $F$-conjugate to $\lambda_0$. Let $\lambda_0$ have minimal polynomial $f(X) \in F[X]$. Its $F$-conjugates are the roots of $f(X)$, so we want to show

$$(A, M) = 1 \Longrightarrow f([A](\lambda_0)) = 0.$$

Since $[A](\lambda_0)$ only depends on $A \bmod M$, we can choose $A$ to be monic and then $A$ is a product of monic irreducibles each not dividing $M$. Since $A \mapsto [A](X)$ converts multiplication to composition, it suffices to show $f([\pi](\lambda_0)) = 0$ for any monic irreducible $\pi \in \mathbf{F}_p[T]$ not dividing $M$.

Pick such a $\pi$ and let $g(X)$ be the minimal polynomial of $[\pi](\lambda_0)$ in $F[X]$. We want to show $g(X) = f(X)$. Both $f(X)$ and $g(X)$ are in $\mathbf{F}_p[T][X]$, not just $F[X] = \mathbf{F}_p(T)[X]$. Indeed, they both divide $[M](X)$ (since $\lambda_0$ and $[\pi](\lambda_0)$ are in $\Lambda_M$) and any monic factor of $[M](X)$ in $F[X]$ is in $\mathbf{F}_p[T][X]$ (analogue of Gauss' lemma from $\mathbf{Z}[X]$).

Since $g([\pi](\lambda_0)) = 0$, $g([\pi](X))$ has $\lambda_0$ as a root, so $f(X) | g([\pi](X))$ in $F[X]$. Both $f(X)$ and $g([\pi](X))$ are monic $X$-polynomials in $\mathbf{F}_p[T][X]$ (because $\pi$ is monic!), so the divisibility in $F[X]$ in fact takes place in $\mathbf{F}_p[T][X]$. That is, $g([\pi](X)) = f(X)h(X)$ for some $h(X)$ in $\mathbf{F}_p[T][X]$. (The proof of this is the same as the proof that if $f(X)$ and $g(X)$ are in $\mathbf{Z}[X]$ and $f(X) | g(X)$ in $\mathbf{Q}[X]$ then $f(X) | g(X)$ in $\mathbf{Z}[X]$: there is unique division with remainder by *monic* polynomials in both $\mathbf{Z}[X]$ and $\mathbf{Q}[X]$, and likewise in both $\mathbf{F}_p[T][X]$ and $F[X]$.) Hence $g([\pi](X)) = f(X)h(X)$ for some $h(X)$ in $\mathbf{F}_p[T][X]$. Reduce modulo $\pi$ and use Theorem 2.14 to get

$$\overline{g}(X)^{p^{\deg \pi}} = \overline{f}(X)\overline{h}(X).$$

Thus $\overline{f}(X)$ and $\overline{g}(X)$ have a common factor in $(\mathbf{F}_p[T]/\pi)[X]$, namely any irreducible factor of $\overline{f}(X)$.

To show $g(X) = f(X)$, assume not. They are then distinct monic irreducible factors of $[M](X)$, so $[M](X) = f(X)g(X)k(X)$ for some $k(X) \in \mathbf{F}_p[T][X]$. Reducing this modulo $\pi$,

$$\overline{[M]}(X) = \overline{f}(X)\overline{g}(X)\overline{k}(X)$$

in $(\mathbf{F}_p[T]/\pi)[X]$. This is impossible: the right side has a multiple irreducible factor (any common irreducible factor of $\overline{f}(X)$ and $\overline{g}(X)$) but $\overline{[M]}(X)$ is separable in $(\mathbf{F}_p[T]/\pi)[X]$ (Theorem 2.8). So $g(X) = f(X)$, which shows $f([\pi](\lambda_0)) = 0$. □

**Example 4.3.** Let $M = T$. The isomorphism $\mathrm{Gal}(F(\Lambda_T)/F) \cong (\mathbf{F}_p[T]/T)^\times$ associates to each $\sigma$ the unique $A \bmod T \in (\mathbf{F}_p[T]/T)^\times$ where $\sigma(\lambda) = [A](\lambda)$ for all $\lambda \in \Lambda_T$. Since $[A](\lambda) = A(0)\lambda$ (Example 3.6) and $(\mathbf{F}_p[T]/T)^\times \cong \mathbf{F}_p^\times$ by identifying each nonzero congruence class mod $T$ with the constant in that congruence class, the isomorphism in Theorem 4.2 identifies $\mathrm{Gal}(F(\Lambda_T)/F)$ with $\mathbf{F}_p^\times$ through scaling: $\sigma_c(\lambda) = c\lambda$ for all $\lambda \in \Lambda_T$ as $c$ runs through $\mathbf{F}_p^\times$.

Since $c = \sigma_c(\lambda)/\lambda$ for any $\lambda \in \Lambda_T - \{0\}$, and $F(\Lambda_T)/F$ is a Kummer extension (Example 3.4), notice that the identification of the Galois group with $\mathbf{F}_p^\times$ is exactly how Kummer theory would apply in this situation too.

The Carlitz construction leads to abelian extensions not only of $\mathbf{F}_p(T)$ but of any characteristic $p$ field $K$ which is not algebraic over $\mathbf{F}_p$: denote an element of $K$ transcendental over $\mathbf{F}_p$ as $T$, so $K \supset \mathbf{F}_p(T)$. Using this $T$ we obtain the polynomials $[M](X) \in \mathbf{F}_p(T)[X] \subset K[X]$. Then $[M](X)$ is separable in $K[X]$ and $K(\Lambda_M)/K$ is a Galois extension with the effect of the Galois group on $\Lambda_M$ leading to an embedding $\mathrm{Gal}(K(\Lambda_M)/K) \hookrightarrow (\mathbf{F}_p[T]/M)^\times$, so the Galois group is abelian. This embedding need not be onto (depends on $K$ and the choice of $T$ in $K$).

## 5. More cyclotomic and Carlitz analogies

The roots of the polynomials $X^m - 1$ and $[M](X)$ have similar features (*e.g.*, cyclic group of size $m$, cyclic $\mathbf{F}_p[T]$-module of size $\mathrm{N}(M)$), but it is the isomorphisms of Galois groups, $\mathrm{Gal}(\mathbf{Q}(\mu_m)/\mathbf{Q}) \cong (\mathbf{Z}/(m))^\times$ and $\mathrm{Gal}(F(\Lambda_M)/F) \cong (\mathbf{F}_p[T]/M)^\times$, which are the first strong indication that the analogies we are seeing could be substantial. We explore some further analogies in this section.

By Theorem 4.2, $[F(\Lambda_M) : F] = \#(\mathbf{F}_p[T]/M)^\times$ for any $M \neq 0$, just as $[\mathbf{Q}(\mu_m) : \mathbf{Q}] = \#(\mathbf{Z}/(m))^\times$ for $m \in \mathbf{Z}^+$. The size of $(\mathbf{Z}/(m))^\times$ is denoted $\varphi(m)$ and similarly the size of $(\mathbf{F}_p[T]/M)^\times$ is denoted $\varphi(M)$. Their values are given by similar formulas:

$$\varphi(m) = m\prod_{p|m}\left(1 - \frac{1}{p}\right), \quad \varphi(M) = \mathrm{N}(M)\prod_{\pi|M}\left(1 - \frac{1}{\mathrm{N}(\pi)}\right),$$

with the product running over (positive) prime factors of $m$ and (monic) irreducible factors of $M$. In particular, from these formulas one can check that

$$\varphi(ab) = \frac{\varphi(a)\varphi(b)(a,b)}{\varphi((a,b))}, \quad \varphi(AB) = \frac{\varphi(A)\varphi(B)\,\mathrm{N}((A,B))}{\varphi((A,B))}. \tag{5.1}$$

Let's put the two formulas in (5.1) to work towards analogous goals. Classically, two cyclotomic fields $\mathbf{Q}(\mu_m)$ and $\mathbf{Q}(\mu_n)$ with $m \leq n$ are equal if and only if $m = n$ or $m$ is odd and $n = 2m$ (*e.g.*, $\mathbf{Q}(\mu_3) = \mathbf{Q}(\mu_6)$, or even more simply $\mathbf{Q}(\mu_1) = \mathbf{Q}(\mu_2)$). We can ask similarly when $F(\Lambda_M) = F(\Lambda_N)$. First we will recall the proof of the classical result for cyclotomic fields and then just translate the argument over to the Carlitz setting.

**Theorem 5.1.** *For $m \leq n$, $\mathbf{Q}(\mu_m) = \mathbf{Q}(\mu_n)$ if and only if $m = n$ or $m$ is odd and $n = 2m$.*

*Proof.* Our argument is adapted from [1, p. 158]. When $m$ is odd, $-\zeta_m$ has order $2m$, so $\mu_{2m} \subset \mathbf{Q}(\mu_m)$. Therefore $\mathbf{Q}(\mu_{2m}) \subset \mathbf{Q}(\mu_m)$. The reverse inclusion is clear since $\mu_m \subset \mu_{2m}$, so $\mathbf{Q}(\mu_m) = \mathbf{Q}(\mu_{2m})$.

Conversely, assume $\mathbf{Q}(\mu_m) = \mathbf{Q}(\mu_n)$ with $m < n$. To show $m$ is odd and $n = 2m$, we count the number of roots of unity in $\mathbf{Q}(\mu_m)$. (The answer need not be $m$, *e.g.*, $\mathbf{Q} = \mathbf{Q}(\mu_1)$ contains 2 roots of unity rather than only 1.) If $\mu_r \subset \mathbf{Q}(\mu_m)$, then $\mathbf{Q}(\mu_r) \subset \mathbf{Q}(\mu_m)$, so taking degrees over $\mathbf{Q}$ shows $\varphi(r) \leq \varphi(m)$. As $r \to \infty$, $\varphi(r) \to \infty$ (albeit erratically) so there is a largest $r$ satisfying $\mu_r \subset \mathbf{Q}(\mu_m)$. Then $\mu_m \subset \mu_r$ (because $\mu_m\mu_r = \mu_{[m,r]}$), so $m|r$ and $\mathbf{Q}(\mu_r) = \mathbf{Q}(\mu_m)$. Write $r = ms$. Then by (5.1),

$$\varphi(r) = \varphi(ms) = \varphi(m)\varphi(s)\frac{(m,s)}{\varphi((m,s))} \geq \varphi(m)\varphi(s).$$

Since $\mathbf{Q}(\mu_m) = \mathbf{Q}(\mu_r)$, computing degrees over $\mathbf{Q}$ shows $\varphi(m) = \varphi(r) \geq \varphi(m)\varphi(s)$, so $1 \geq \varphi(s)$. Thus $\varphi(s) = 1$, so $s = 1$ or 2, so $r = m$ or $r = 2m$. This shows the only option for $n$ when $n > m$ is $n = 2m$. If $m$ is even then $\varphi(2m) = 2\varphi(m) > \varphi(m)$, so $r \neq 2m$. Thus $m$ is odd and $n = 2m$. □

A close look at the proof shows that the discrepancy with 2 comes from the fact that $\varphi(s) = 1$ (that is, $(\mathbf{Z}/(s))^\times$ is trivial) only for $s = 1$ and $s = 2$. Since $\varphi(S) = 1$ (that is, $(\mathbf{F}_p[T]/S)^\times$ is trivial) only if $S \in \mathbf{F}_p^\times$, we are naturally led to the next result.

**Theorem 5.2.** *For nonzero $M$ and $N$ in $\mathbf{F}_p[T]$, $F(\Lambda_M) = F(\Lambda_N)$ if and only if $N = cM$ for some $c \in \mathbf{F}_p^\times$.*

*Proof.* If $N = cM$ with $c \in \mathbf{F}_p^\times$ then $[N](X) = c[M](X)$, so $\Lambda_N = \Lambda_M$ and therefore $F(\Lambda_M) = F(\Lambda_N)$.

Now we show, conversely, that if $F(\Lambda_M) = F(\Lambda_N)$ then $M$ and $N$ are equal up to a scalar multiple from $\mathbf{F}_p^\times$.

We will compute the largest Carlitz torsion in $F(\Lambda_M)$. If $\Lambda_R \subset F(\Lambda_M)$ then $F(\Lambda_R) \subset F(\Lambda_M)$, so taking degrees over $F$ shows $\varphi(R) \leq \varphi(M)$. As $\mathrm{N}(R) \to \infty$, $\varphi(R) \to \infty$, so there is an $R$ with $\Lambda_R \subset F(\Lambda_M)$ and $\mathrm{N}(R)$ as large as possible. Then $\Lambda_M \subset \Lambda_R$ (because $\Lambda_M + \Lambda_R = \Lambda_{[M,R]}$), so $M|R$ and $F(\Lambda_R) = F(\Lambda_M)$. Write $R = MS$. Then by (5.1),

$$\varphi(R) = \varphi(MS) = \varphi(M)\varphi(S)\frac{\mathrm{N}((M,S))}{\varphi((M,S))} \geq \varphi(M)\varphi(S).$$

Since $F(\Lambda_M) = F(\Lambda_R)$, computing degrees over $F$ shows $\varphi(M) = \varphi(R) \geq \varphi(M)\varphi(S)$, so $1 \geq \varphi(S)$. Thus $\varphi(S) = 1$, so $S \in \mathbf{F}_p^\times$, which means $R$ is a scalar multiple of $M$. Thus $\Lambda_R = \Lambda_M$, so all the Carlitz torsion in $F(\Lambda_M)$ is inside $\Lambda_M$. Hence if $F(\Lambda_M) = F(\Lambda_N)$ we get $\Lambda_M = \Lambda_N$, so by Corollary 2.7, $M = cN$ for some $c \in \mathbf{F}_p^\times$. □

For $m \in \mathbf{Z}^+$, the roots of unity in $\mathbf{C}$ of exact order $m$ share the same minimal polyomial over $\mathbf{Q}$, the $m$th cyclotomic polynomial:

$$\Phi_m(X) = \prod_{(a,m)=1} (X - \zeta^a) = \prod_{\substack{\zeta^m=1 \\ \zeta^d \neq 1}} (X - \zeta),$$

where $\zeta$ has order $m$ and in the second product $d$ runs over proper divisors of $m$. For example, if $p$ is prime then $\Phi_p(X) = (X^p - 1)/(X - 1)$: every $p$th root of unity has order $p$ except for 1. The polynomial $\Phi_p(X + 1) = ((X + 1)^p - 1)/X$ is Eisenstein with respect to $p$. By comparing degrees, roots, and leading coefficients, $\Phi_{p^k}(X) = \Phi_p(X^{p^{k-1}})$. Each $\Phi_{p^k}(X + 1)$ is Eisenstein with respect to $p$.

For monic $M$ in $\mathbf{F}_p[T]$, all generators of $\Lambda_M$ have the same minimal polynomial over $F = \mathbf{F}_p(T)$, which is an analogue of the cyclotomic polynomials:

$$\Phi_M(X) = \prod_{(A,M)=1} (X - [A](\lambda_0)) = \prod_{\substack{[M](\lambda)=0 \\ [D](\lambda)\neq 0}} (X - \lambda),$$

where $\lambda_0$ is a chosen generator of $\Lambda_M$ and the second product is taken over roots $\lambda$ of $[M](X)$ which are not roots of $[D](X)$ for any monic proper divisor $D$ of $M$. (Such $\lambda$ are the generators of $\Lambda_M$.)

**Example 5.3.** If $\pi$ is irreducible then $\Phi_\pi(X) = [\pi](X)/X$ since $[\pi](X)/X$ is Eisenstein with respect to $\pi$ (Corollary 2.9) and thus is irreducible over $\mathbf{F}_p(T) = F$. Comparing degrees, roots, and leading coefficients, for any $k \geq 1$ we have $\Phi_{\pi^k}(X) = \Phi_\pi([\pi^{k-1}](X))$, so the constant term of $\Phi_{\pi^k}(X)$ is $\Phi_\pi([\pi^{k-1}](0)) = \Phi_\pi(0) = \pi$. Since $\Phi_\pi(X)$ has all non-leading $X$-coefficients divisible by $\pi$ and $[\pi^{k-1}](X)$ also has all non-leading $X$-coefficients divisible by $\pi$ (Corollary 2.11), $\Phi_{\pi^k}(X)$ has all non-leading $X$-coefficients divisible by $\pi$. Therefore $\Phi_{\pi^k}(X)$ is Eisenstein with respect to $\pi$ for any $k$.

**Remark 5.4.** It was noted in Remark 2.10 that $[M](X)$ more closely resembles $(1+X)^m - 1$ than $X^m - 1$. Since $[M](X) = \prod_{D|M} \Phi_D(X)$, where the product is taken over the monic divisors $D$ of $M$, we might anticipate that $\Phi_M(X)$ more closely resembles $\Phi_m(X+1)$ than $\Phi_m(X)$, and this does appear to be true. For instance, $\Phi_{\pi^k}(X)$ is Eisenstein with respect to $\pi$ while $\Phi_{p^k}(X+1)$ – not $\Phi_{p^k}(X)$ – is Eisenstein with respect to $p$. If $m$ is not a power of a prime then $\Phi_m(1) = 1$. If $M$ is monic and not a power of an irreducible, the analogous equation reads $\Phi_M(0) = 1$.

The Kronecker-Weber theorem says every finite abelian extension of $\mathbf{Q}$ lies in a cyclotomic extension $\mathbf{Q}(\mu_m)$. There is an analogue of the Kronecker-Weber theorem for $\mathbf{F}_p(T)$, due to Hayes [4]. It says every finite abelian extension of $\mathbf{F}_p(T)$ lies in some $\mathbf{F}_{p^d}(T, \Lambda_M, \Lambda_{1/T^n})$ for some $d \geq 1$, $n \geq 1$, and $M \in \mathbf{F}_p[T]$, where $\Lambda_{1/T^n}$ is the set of roots of the Carlitz polynomial $[1/T^n](X)$ built with $1/T$ in place of $T$.

**Example 5.5.** Using $1/T$ as the generator over $\mathbf{F}_p$ for $\mathbf{F}_p(T) = \mathbf{F}_p(1/T)$, the polynomial $[1/T](X) = X^p + (1/T)X = X(X^{p-1} + 1/T)$ has roots which generate the same extension of $\mathbf{F}_p(T)$ as $[T](X)$. But for $[1/T^2](X)$ we get something new:

$$[1/T^2](X) = [1/T]([1/T](X)) = X^{p^2} + ((1/T)^p + (1/T))X^p + (1/T^2)X,$$

and the extension $\mathbf{F}_p(T, \Lambda_{1/T^2})/\mathbf{F}_p(T)$ turns out to have a property (wild ramification at $\infty$) which is not satisfied by subfields of $\mathbf{F}_{p^d}(T, \Lambda_M)$, so it is not inside such a field.

Table 2 summarizes some of the analogous features we have seen with cyclotomic extensions of $\mathbf{Q}$ and Carlitz extensions of $\mathbf{F}_p(T)$.

| Cyclotomic | Carlitz |
|---|---|
| $\#\mu_m = m$ | $\#\Lambda_M = \mathrm{N}(M)$ |
| $d\|m \Leftrightarrow \mu_d \subset \mu_m$ | $D\|M \Leftrightarrow \Lambda_D \subset \Lambda_M$ |
| $\zeta \in \mu_m, a \in \mathbf{Z} \Rightarrow \zeta^a \in \mu_m$ | $\lambda \in \Lambda_M, A \in \mathbf{F}_p[T] \Rightarrow [A](\lambda) \in \Lambda_M$ |
| $a \equiv b \bmod m \Rightarrow \zeta^a = \zeta^b$ | $A \equiv B \bmod M \Rightarrow [A](\lambda) = [B](\lambda)$ |
| $\zeta^a = \zeta^b (\text{all } \zeta) \Rightarrow a \equiv b \bmod m$ | $[A](\lambda) = [B](\lambda) (\text{all } \lambda) \Rightarrow A \equiv B \bmod M$ |
| $\mathrm{Gal}(\mathbf{Q}(\mu_m)/\mathbf{Q}) \cong (\mathbf{Z}/(m))^\times$ | $\mathrm{Gal}(\mathbf{F}_p(T, \Lambda_M)/\mathbf{F}_p(T)) \cong (\mathbf{F}_p[T]/M)^\times$ |
| $X^m - 1 = \prod_{d\|m} \Phi_d(X)$ | $[M](X) = \prod_{D\|M} \Phi_D(X)$ |
| Kronecker-Weber theorem | Carlitz-Hayes theorem |

TABLE 2. Analogies between $\mu_m$ and $\Lambda_M$

## 6. The Carlitz Exponential

In this section, we describe how the Carlitz polynomials were first discovered by Carlitz, using an exponential function in characteristic $p$. The idea is to find a characteristic $p$ analogue of the complex-analytic description of roots of unity as $e^{2\pi i a/b}$.

The exponential series $e^X = \sum_{n\geq 0} X^n/n!$, as a function on $\mathbf{C}$, is a homomorphism $\mathbf{C} \to \mathbf{C}^\times$ with (discrete) kernel $2\pi i\mathbf{Z}$. There is an infinite product decomposition for $e^z - 1$ over its roots $2\pi i\mathbf{Z}$:

$$e^z - 1 = e^{z/2} z \prod_{n\geq 1} \left(1 + \frac{z^2}{4\pi^2 n^2}\right) = e^{z/2} z \prod_{n\neq 0} \left(1 - \frac{z}{2\pi i n}\right).$$

The extra $e^{z/2}$ in the formula reflects the fact that knowing the zeros (and their multiplicities) of a complex entire function only determines it up to multiplication by $e^{h(z)}$ for some entire function $h(z)$.

Let's create an analogous infinite product in characteristic $p$ using $\mathbf{F}_p[T]$ in place of $\mathbf{Z}$. Since the characteristic $p$ analogue of $\pi$ (better, $2\pi i$) is not obvious, we will work heuristically at first. Once we find what we are looking for, a precise theorem will be stated.

In a field extension of $\mathbf{F}_p(T)$, pick a nonzero element $\xi$ and think of $\xi\mathbf{F}_p[T]$ as an analogue of $2\pi i\mathbf{Z}$ A power series having simple roots at $\xi\mathbf{F}_p[T]$ is

$$f(X) := X \prod_{\substack{\alpha \in \xi\mathbf{F}_p[T] \\ \alpha \neq 0}} \left(1 - \frac{X}{\alpha}\right), \tag{6.1}$$

and our field extension of $\mathbf{F}_p(T)$ will need some kind of completeness in order for this product to make sense, since the coefficients of the product when it is multiplied out are infinite series.

Because the zero set $\xi\mathbf{F}_p[T]$ of $f(X)$ is closed under scaling by $\mathbf{F}_p$, $f(cX) = cf(X)$ for any $c \in \mathbf{F}_p$. It can also be shown[2] that $f(X)$ is additive. Therefore $f(TX)$ has the roots

$$\frac{1}{T}\xi\mathbf{F}_p[T] = \bigcup_{c\in\mathbf{F}_p} \left(\frac{c}{T}\xi + \xi\mathbf{F}_p[T]\right)$$

and all roots are simple. Because $f(X)$ is additive and vanishes on $\xi\mathbf{F}_p[T]$, on any coset $c\xi/T + \xi\mathbf{F}_p[T]$ the common value of $f(X)$ is $f(c\xi/T)$, so another function besides $f(TX)$ with the roots $(1/T)\xi\mathbf{F}_p[T]$, all of multiplicity 1, is

$$\prod_{c\in\mathbf{F}_p} (f(X) - f(c\xi/T)) = \prod_{c\in\mathbf{F}_p} (f(X) - cf(\xi/T)) = f(X)^p - f(\xi/T)^{p-1} f(X).$$

It is natural to compare this with $f(TX)$, and it would be a very special situation (that is, require a special choice of $\xi$) for the two functions to match:

$$f(TX) = f(X)^p - f(\xi/T)^{p-1} f(X).$$

Let's assume this happens. Then comparing the coefficient of $X$ in the series expansion of both sides forces $T = -f(\xi/T)^{p-1}$, so

$$f(TX) = f(X)^p + Tf(X). \tag{6.2}$$

[2]For any $d$, the coefficients of $X^j$ in $f(X)$ for $j \leq d$ are the limit of the $j$th coefficient of $f_d(X) = X\prod_{\deg A \leq d}(1 - X/\xi A)$ as $d \to \infty$. This product is a polynomial whose roots are simple and are a finite additive group. Any such polynomial is an additive polynomial by Appendix A, so $f_d(X + Y) = f_d(X) + f_d(Y)$. Now let $d \to \infty$.

The condition $T = -f(\xi/T)^{p-1}$ nearly determines $\xi$. By the product defining $f(X)$ in (6.1), we get

$$-T = \left(\frac{\xi}{T}\prod_{A\neq 0}\left(1 - \frac{1}{TA}\right)\right)^{p-1},$$

where $A$ runs over nonzero polynomials in $\mathbf{F}_p[T]$. (The infinite product converges in $\mathbf{F}_p((1/T))$ because in the $1/T$-adic topology, $|1/TA| < 1$ for any nonzero polynomial $A$ and there are only finitely many $A$ with a given degree. An infinite product $\prod_{n\geq 1}(1+\alpha_n)$ in a complete non-archimedean valued field converges when $|alpha_n| \to 0$.) Therefore

$$\xi^{p-1} = \frac{-T^p}{\prod_{A\neq 0}(1 - 1/TA)^{p-1}}. \tag{6.3}$$

This product converges in $\mathbf{F}_p((1/T))$ and $\xi$ is algebraic over $\mathbf{F}_p((1/T))$. We can use this equation to define $\xi$ (at last). Since $\xi$ appears in (6.3) through $\xi^{p-1}$, the equation only determines $\xi$ up to scaling by a $(p-1)$-th root of unity, namely an element of $\mathbf{F}_p^\times$. This ambiguity doesn't affect the meaning of $\xi\mathbf{F}_p[T]$, which is what shows up in the definition of $f(X)$.

Since $f(X)$ is an additive series, $f(X) = \sum_{j\geq 0} a_j X^{p^j}$ (Appendix A) with $a_j$ to be determined now that we have pinned down a choice of $\xi$ (so that (6.2) is satisfied). The product defining $f(X)$ has first term $X$, so we need $a_0 = 1$. Substituting the series for $f(X)$ into (6.2) gives the recursion $a_j T^{p^j} = a_j T + a_{j-1}^p$ for $j \geq 1$, so $a_j(T^{p^j} - T) = a_{j-1}^p$. Since $a_0 = 1$, we get

$$a_1 = \frac{1}{T^p - T}, \quad a_2 = \frac{1}{(T^{p^2} - T)(T^p - T)^p},$$

and in general $a_j$ is the reciprocal of a polynomial. Let $D_j = 1/a_j$, so $D_0 = 1$ and $D_j = (T^{p^j} - T)D_{j-1}^p$ for $j \geq 1$.

**Definition 6.1.** The *Carlitz exponential* is the power series

$$e_C(X) := \sum_{j\geq 0} \frac{X^{p^j}}{D_j} \in \mathbf{F}_p(T)[[X]],$$

where $D_0 = 1$ and $D_j = (T^{p^j} - T)D_{j-1}^p$ for $j \geq 1$.

**Remark 6.2.** It can be shown for $j \geq 1$ that

$$D_j := \prod_{\substack{h\in\mathbf{F}_p[T] \text{ monic} \\ \deg h = j}} h.$$

In particular, $\deg(D_j) = jp^j$.

**Theorem 6.3** (Carlitz)**.** *There is an infinite product decomposition*

$$e_C(X) = X\prod_{A\neq 0}\left(1 - \frac{X}{A\xi}\right),$$

*with the product running over nonzero $A$ in $\mathbf{F}_p[T]$ and with*

$$\xi := \frac{-T^{1/(p-1)}T}{\prod_{A\neq 0}(1 - 1/TA)}.$$

**Remark 6.4.** Carlitz gave another expression for $\xi$:

$$\xi = (T - T^p)^{1/(p-1)} \prod_{j \geq 1} \left(1 - \frac{T^{p^j} - T}{T^{p^{j+1}} - T}\right).$$

The Carlitz exponential satisfies $e_C(X+Y) = e_C(X) + e_C(Y)$ rather than $e^{X+Y} = e^X e^Y$. Instead of $(e^X)' = e^X$ we have $e_C'(X) = 1$. Actually, the equation for $e_C(X)$ which is as important for it as the differential equation is for $e^X$ is not $e_C'(X) = 1$ but rather (6.2): $e_C(TX) = e_C(X)^p + Te_C(X)$.

The parameter $\xi$ doesn't appear in the coefficients of the Carlitz exponential series in Definition 6.1, just like $\pi$ doesn't appear in the definition of the usual exponential series. The value $\xi$ is a characteristic $p$ analogue of $2\pi i$. It is not in $\mathbf{F}_p((1/T))$, just as $2\pi i \notin \mathbf{R}$. Wade [8] proved $\xi$ is transcendental over $\mathbf{F}_p(T)$, which is analogous to $2\pi i$ being transcendental over $\mathbf{Q}$.

As a function on $\mathbf{F}_p((1/T))$, the formal power series for $e_C(X)$ is an "entire function": it converges everywhere. Indeed, for any $x \in \mathbf{F}_p((1/T))$, the $1/T$-adic absolute value of the general term in the series $e_C(x)$ is $|x|^{p^j}/|D_j| = |x|^{p^j}/(1/p)^{-jp^j} = (p^j|x|)^{p^j}$ by Remark 6.2. This tends to 0 as $j \to \infty$, so the series $e_C(x)$ converges. Setting $X = 1$, for instance, we get the $1/T$-adic power series

$$e_C(1) = 1 + \frac{1}{T^p - T} + \frac{1}{(T^{p^2} - T)(T^p - T)^p} + \cdots = 1 + \frac{1}{T^p} + \frac{1}{T^{2p-1}} + \cdots.$$

The homomorphism $e_C \colon \mathbf{F}_p((1/T)) \to \mathbf{F}_p((1/T))$ is injective, just like the homomorphism $\exp \colon \mathbf{R} \to \mathbf{R}_{>0}$. We need to enlarge the domain of $e_C(X)$ beyond $\mathbf{F}_p((1/T))$ to find its full kernel $\xi\mathbf{F}_p[T]$.

We now explore the relation between the Carlitz exponential and Carlitz polynomials. The property $e_C(TX) = e_C(X)^p + Te_C(X)$ says, in terms of the Carlitz polynomial $[T](X)$, that $e_C(TX) = [T](e_C(X))$. The Carlitz exponential series converts plain multiplication by $T$ into the Carlitz action by $T$. Since $e_C(X)$ is $\mathbf{F}_p$-linear in $X$, it follows for any $M \in \mathbf{F}_p[T]$ that

$$e_C(MX) = [M](e_C(X)). \tag{6.4}$$

In other words, $e_C(MX)$ is a polynomial in $e_C(X)$, and that polynomial is precisely the Carlitz polynomial $[M](X)$. If we had not known about the Carlitz polynomials, they would be forced upon us when we express $e_C(MX)$ in terms of $e_C(X)$. The analogue of (6.4) for $e^X$ is the much simpler $e^{mX} = (e^X)^m$, or equivalently $e^{mX} - 1 = [m](e^X - 1)$ with $[m](X) = (1+X)^m - 1$. What gives (6.4) analytic content is the next result, which is the analogue for Carlitz torsion of the complex-analytic parametrization of $m$th roots of unity: $\mu_m = \{e^{2\pi i a/m} : a \in \mathbf{Z}\}$.

**Theorem 6.5.** *For nonzero $M$ in $\mathbf{F}_p[T]$, $\Lambda_M = \{e_C((A/M)\xi) : A \in \mathbf{F}_p[T]\}$.*

*Proof.* For any $A$ in $\mathbf{F}_p[T]$,

$$[M](e_C((A/M)\xi)) = e_C(M(A/M)\xi) = e_C(A\xi) = 0,$$

so $e_C((A/M)\xi) \in \Lambda_M$. To show these Carlitz exponential values fill up $\Lambda_M$, we count how many values there are. If $e_C((A/M)\xi) = e_C((B/M)\xi)$ then subtracting shows $e_C(((A-B)/M)\xi) = 0$, so $(A-B)/M \in \mathbf{F}_p[T]$ by Theorem 6.3. Thus $A \equiv B \bmod M$, so the number of values for $e_C((A/M)\xi)$ as $A$ varies is $\#(\mathbf{F}_p[T]/M) = \#\Lambda_M$. □

**Corollary 6.6.** *As an* $\mathbf{F}_p[T]$*-module,* $\bigcup_M \Lambda_M \cong \mathbf{F}_p(T)/\mathbf{F}_p[T]$.

*Proof.* The map $\mathbf{F}_p(T) \to \bigcup_M \Lambda_M$ given by $A/B \mapsto e_C((A/B)\xi)$ is $\mathbf{F}_p[T]$-linear, surjective, and its kernel is $\mathbf{F}_p[T]$. □

Corollary 6.6 is analogous to the group of all roots of unity in $\overline{\mathbf{Q}}^\times$ being isomorphic to $\mathbf{Q}/\mathbf{Z}$ using $z \mapsto e^{2\pi i z}$.

The Carlitz exponential helps us describe the coefficients of $[M](X)$. Finding these coefficients is analogous to finding the coefficients of $[m](X) = (1+X)^m - 1 = \sum_{j=1}^m \binom{m}{j} X^j$ from scratch as if we did not know what binomial coefficients were. In fact, we will show how to find the formula for binomial coefficients first, and then translate the steps into the Carlitz setting.

We start off by writing

$$[m](X) = (1+X)^m - 1 = \sum_{j=1}^m c_{j,m} X^j, \tag{6.5}$$

where $c_{m,m} = 1$. (One doesn't need to know the binomial theorem to see that $(1+X)^m - 1$ has leading term $X^m$ and constant term 0.) Our goal is to show $c_{j,m}$ is given by a universal polynomial formula in $m$. Because $[m](e^X - 1) = e^{mX} - 1$, replacing $X$ with $\log(1+X) = X + \cdots$ gives

$$[m](X) = e^{m\log(1+X)} - 1 = \sum_{j \geq 1} \frac{m^j (\log(1+X))^j}{j!}. \tag{6.6}$$

The right side is a series in $\mathbf{Q}[[X]]$ and it makes sense since $(\log(1+X))^j = X^j + \cdots$. Replacing $m$ with an indeterminate $Y$,

$$e^{Y\log(1+X)} - 1 = \sum_{j \geq 1} \frac{Y^j (\log(1+X))^j}{j!} = \sum_{j \geq 1} P_j(Y) X^j$$

for some $P_j(Y) \in \mathbf{Q}[Y]$. Paying attention to how terms with $X^j$ arise, $\deg P_j(Y) = j$. Since $e^0 - 1 = 0$, $P_j(0) = 0$ for all $j$. Now setting $Y = m$,

$$[m](X) = \sum_{j \geq 1} P_j(m) X^j.$$

Comparing this with (6.5), we observe that

$$c_{j,m} = P_j(m) \text{ for } 1 \leq j \leq m, \text{ and } P_j(m) = 0 \text{ for } j > m.$$

The first part tells us $c_{j,m}$ is some universal polynomial in $m$, and the second part actually tells us what the polynomial is: since $P_j(Y)$ vanishes at positive integers less than $j$ and at 0, $P_j(Y)$ is divisible by $Y(Y-1)\cdots(Y-(j-1))$, which has degree $j$. Since the degree of $P_j(Y)$ is $j$ and $P_j(j) = c_{j,j} = 1$, we must have

$$P_j(Y) = \frac{Y(Y-1)\cdots(Y-(j-1))}{j(j-1)\cdots(j-(j-1))} = \frac{Y(Y-1)\cdots(Y-(j-1))}{j!}.$$

Therefore $c_{j,m} = P_j(m)$ is our friend the binomial coefficient $\binom{m}{j}$.

Now we turn to characteristic $p$, and carry out an analogous procedure. Write

$$[M](X) = \sum_{j=0}^{\deg M} a_{j,M}(T) X^{p^j}, \quad a_{j,M}(T) \in \mathbf{F}_p[T]. \tag{6.7}$$

(In Theorem 2.4 we wrote the coefficients as $a_j(T)$ rather than as $a_{j,M}(T)$, but the coefficient do depend on $M$ and now we need to keep track of that information.) We know by Theorem 2.4 that $a_{\deg M,M} = \operatorname{lead} M$, so $a_{\deg M,M} = 1$ for monic $M$. Since $[M](e_C(X)) = e_C(MX)$, we want to replace $X$ with the composition inverse of $e_C(X)$ to mimic (6.6). This inverse of $e_C(X)$ is the *Carlitz logarithm*, denoted $\log_C(X)$. Since $e_C(X) = X + \cdots$, $\log_C(X) = X + \cdots$. Since $e_C(X)$ is additive, $\log_C(X)$ is additive, so it is a series with terms $X^{p^j}$. (In particular, $\log_C'(X) = 1$.) The logarithmic equivalent of (6.4) is $\log_C([M](X)) = M \log_C(X)$. Explicitly, Carlitz found the formula

$$\log_C(X) = \sum_{j\geq 0} \frac{X^{p^j}}{L_j},$$

where $L_0 = 1$ and $L_j = (T^{p^j} - T)(T^{p^{j-1}} - T)\cdots(T^p - T)$ for $j \geq 1$. But we don't need to know this explicit formula for $\log_C(X)$, just as we never needed to know explicit coefficients of $\log(1 + X)$ above when using that series.

Replacing $X$ with $\log_C(X)$ in the equation $[M](e_C(X)) = e_C(MX)$, we get

$$[M](X) = e_C(M \log_C(X)) = \sum_{j\geq 0} \frac{M^{p^j}(\log_C(X))^{p^j}}{D_j}. \tag{6.8}$$

This series is in $\mathbf{F}_p(T)[[X]]$. Replacing $M \in \mathbf{F}_p[T]$ with a variable $Y$,

$$e_C(Y \log_C(X)) = \sum_{j\geq 0} \frac{Y^{p^j}(\log_C(X))^{p^j}}{D_j} = \sum_{j\geq 0} E_j(Y)X^{p^j},$$

where $E_j(Y) \in \mathbf{F}_p(T)[Y]$. (The series for $e_C(Y \log_C(X))$ involves only $p$-power terms in $X$ since that is all that occurs in $\log_C(X)$, which itself is being raised to $p$-powers when $e_C(Y \log_C X)$ is expanded out.) Because $(\log_C(X))^{p^j}$ begins with the term $X^{p^j}$, $E_j(Y)$ has degree $p^j$ in $Y$. Since $e_C(0) = 0$, $E_j(0) = 0$ for all $j$.

Now setting $Y = M \in \mathbf{F}_p[T]$ in (6.8), we get

$$[M](X) = e_C(M \log_C(X)) = \sum_{j\geq 0} E_j(M)X^{p^j},$$

so a comparison with (6.7) gives

$$a_{j,M}(T) = E_j(M) \text{ for } 0 \leq j \leq \deg M, \text{ and } E_j(M) = 0 \text{ for } j > \deg M.$$

The second part tells us $E_j(X)$ is divisible by

$$\prod_{\deg h<j} (X - h),$$

where $h = 0$ is included in the product. This product has degree $p^j = \deg E_j(X)$, so it differs from $E_j(X)$ by a factor in $\mathbf{F}_p(T)$. Since $E_j(T^j) = a_{j,T^j}(T) = 1$ (because $T^j$ is monic), we obtain (using Remark 6.2)

$$E_j(X) = \frac{\prod_{\deg h<j}(X - h)}{\prod_{\deg h<j}(T^j - h)} = \frac{\prod_{\deg h<j}(X - h)}{\prod_{\substack{h \text{ monic} \\ \deg h=j}} h} = \frac{\prod_{\deg h<j}(X - h)}{D_j}. \tag{6.9}$$

Therefore when $j \leq \deg M$,

$$a_{j,M} = E_j(M) = \frac{\prod_{\deg h < j}(M-h)}{D_j},$$

which gives a universal polynomial formula for $a_{j,M}$ in terms of $M$. The formula is also valid for $j > \deg M$ since the formula is then 0. This is the analogue of the binomial coefficient formula and suggests that, on account of the degrees involved, $E_j(X)$ is an analogue of $\binom{X}{p^j}$ and $D_j$ is an analogue of $(p^j)!$.

**Example 6.7.** Since $E_1(X) = \prod_{c\in\mathbf{F}_p}(X-c)/D_1 = (X^p - X)/(T^p - T)$, we have $a_{1,M} = E_1(M) = (M^p - M)/(T^p - T)$.

## 7. Larger constant fields

We have carried out the construction of Carlitz extensions over $\mathbf{F}_p(T)$, but everything extends to $\mathbf{F}_q(T)$ as the base field for any finite field $\mathbf{F}_q$. Now we set $[T](X) = X^q + TX$ rather than $X^p + TX$, and define $[T^n](X)$ by iteration and $[M](X)$ for any $M \in \mathbf{F}_q[T]$ by $\mathbf{F}_q$-linearity (not $\mathbf{F}_p$-linearity). These are Carlitz polynomials adapted to $\mathbf{F}_q$. Now $[M](X)$ is a $q$-polynomial (only involving terms $X, X^q, X^{q^2}$, and so on) and its roots $\Lambda_M$ form an $\mathbf{F}_q[T]$-module of size $\mathrm{N}(M) := q^{\deg M}$ (new definition of the norm, adapted to the larger constant field). We get a functor $\mathcal{A} \rightsquigarrow C(\mathcal{A})$ from $\mathbf{F}_q[T]$-algebras to $\mathbf{F}_q[T]$-modules by letting $C(\mathcal{A})$ be $\mathcal{A}$ as an $\mathbf{F}_q$-vector space with $\mathbf{F}_q[T]$ acting on it through the Carlitz polynomials rather than through the original $\mathbf{F}_q[T]$-algebra structure on $\mathcal{A}$. The particular case $C(\overline{\mathbf{F}_q(T)})$, which is the field $\overline{\mathbf{F}_q(T)}$ equipped with the action of the Carlitz polynomials $[M](X)$ as defined above, is called the Carlitz module (over $\mathbf{F}_q$). All properties of Carlitz torsion from before extend: just replace $p$ by $q$ everywhere. In particular, for $M \in \mathbf{F}_q[T]$ the roots of $[M](X)$ generate an abelian extension of $\mathbf{F}_q(T)$ with Galois group isomorphic to $(\mathbf{F}_q[T]/M)^\times$.

The Carlitz exponential $e_C(X)$ for $\mathbf{F}_q[T]$, rather than $\mathbf{F}_p[T]$, is defined with denominators $D_j = (T^{q^j} - T)D_{j-1}^q$ and $D_0 = 1$. The zeros of the new $e_C(X)$ are $\mathbf{F}_q[T]$-multiples of a new transcendental $\xi$ which is given by the same formula as the old $\xi$ with $p$ replaced by $q$.

## Appendix A. Additive Polynomials

A polynomial $f(X)$ is called *additive* if $f(X+Y) = f(X) + f(Y)$. For a field $F$, we say $f(X) \in F[X]$ is *$F$-linear* if it is additive and $f(cX) = cf(X)$ for all $c \in F$. We will classify the additive and $F$-linear polynomials.

Before we see what additive polynomials look like in general, we give a result that shows how they can be constructed using finite additive subgroups of $F$. (Such subgroups are nontrivial only in positive characteristic.)

**Theorem A.1.** *Let $F$ have characteristic $p$ and let $V \subset F$ be a finite additive subgroup. The product*

$$\prod_{v\in V}(X - v) = X^{\#V} + \cdots$$

*is an additive polynomial.*

*Proof.* Call the product $f(X)$. For indeterminates $X$ and $Y$, let $g(Y) = f(X+Y) - f(X) - f(Y)$ in $F[X][Y] = F[X,Y]$. We want to show $g(Y) = 0$ in $F[X,Y]$. The leading $Y$-terms

in $f(X+Y)$ and $f(Y)$ match, so $\deg_Y(g) < \#V$. Therefore we can show $g(Y) = 0$ by showing $g(Y)$ has $\#V$ roots. For each $w \in V$,

$$g(w) = f(X+w) - f(X) \in F[X].$$

We will show this is 0 in $F[X]$. The leading $X$-terms of $f(X+w)$ and $f(X)$ match, so $g(w)$ is a polynomial whose $X$-degree is less than $\#V$. Since $f(X+w) - f(X)$ vanishes when we set $X$ to be any $u \in V$ (since $u+w$ and $u$ are both roots of $f$), $g(w) = 0$ in $F[X]$. Since $g(w) = 0$ for each $w \in V$ and the $Y$-degree of $g$ is less than $\#V$, $g = 0$ in $F[X,Y]$. □

**Theorem A.2.** *If $F$ has characteristic 0 then $f(X) \in F[X]$ is additive if and only if it has the form $f(X) = aX$. If $F$ has characteristic $p$ then $f(X) \in F[X]$ is additive if and only if it is of the form $f(X) = a_0X + a_1X^p + a_2X^{p^2} + \cdots + a_mX^{p^m}$ for some $m$.*

*Proof.* The indicated examples are additive.

To prove the converse, let $f(X)$ be additive. We apply differentiation with respect to $Y$ to the identity $f(X+Y) = f(X)+f(Y)$ and then set $Y = 0$. The result is $f'(X) = f'(0) \in F$. Putting $f(X) = \sum_{i=0}^{d} c_iX^i$, we get $\sum_{i=1}^{d} ic_iX^{i-1} = c_1$, so $ic_i = 0$ for $i > 1$. If $F$ has characteristic 0 then $c_i = 0$ for $i > 1$, so $f(X) = c_0 + c_1X$. Since $f(0) = 0$, $c_0 = 0$ so $f(X) = c_1X$. If $F$ has characteristic $p$ then $c_i = 0$ when $i$ is not divisible by $p$ (with $i > 1$), so $f(X) = c_1X + g(X^p)$ for some $g$. Write $c_1$ as $a_0$, so $f(X) = a_0X + g(X^p)$. If $g(X) = 0$ then we are done. Suppose $g(X) \neq 0$. Since $f(0) = 0$, also $g(0) = 0$, so $g(X)$ is a multiple of $X$ and $f(X) \equiv a_0X \bmod X^p$. We have

$$\begin{aligned} g(X^p + Y^p) &= g((X+Y)^p) \\ &= f(X+Y) - c_1(X+Y) \\ &= f(X) + f(Y) - c_1X - c_1Y \\ &= g(X^p) + g(Y^p). \end{aligned}$$

This implies $g(U+V) = g(U) + g(V)$, so $g$ is additive. Therefore $g(X) = a_1X + h(X^p)$ for some $a_1 \in F$, so

$$f(X) = a_0X + a_1X^p + h(X^{p^2}).$$

If $h(X) = 0$ then we are done. If $h(X) \neq 0$ then $h(X)$ is divisible by $X$, so $f(X) \equiv a_0X + a_1X^p \bmod X^{p^2}$. By a similar argument from before, $h(X)$ is additive and this lets us pull out an $a_2X^{p^2}$ term. Repeating this argument enough times, we eventually see $f(X)$ has the desired form since it is a polynomial. □

**Corollary A.3.** *If $F$ is infinite, a polynomial $f(X) \in F[X]$ is $F$-linear if and only if $f(X) = aX$. If $F$ is finite with size $q$ then $f(X) \in F[X]$ is $F$-linear if and only if it is of the form $f(X) = b_0X + b_1X^q + b_2X^{q^2} + \cdots + b_nX^{q^n}$ for some $n$.*

The difference between the $F$-linear and additive polynomials in the case of finite $F$ is that the exponents are $q$-powers rather than simply $p$-powers. For instance, $X + X^p$ is additive in characteristic $p$ and is $\mathbf{F}_p$-linear but is not $\mathbf{F}_{p^2}$-linear.

*Proof.* The indicated examples in the theorem are $F$-linear.

To prove the converse, first suppose $F$ has characteristic 0. Then additivity alone already forces $f(X) = aX$. When $F$ has characteristic $p$, write $f(X) = a_0X + a_1X^p + a_2X^{p^2} + \cdots + a_mX^{p^m}$ for some $m$. The $F$-linearity says $f(cX) = cf(X)$ for all $c \in F$, so $a_ic^{p^i} = ca_i$, which means $c^{p^i} = c$ for all $c \in F$ and any $i$ where $a_i \neq 0$. For $i > 0$, the equation $c^{p^i} = c$

has finitely many roots, so when $F$ is infinite with characteristic $p$ we are forced to have $a_i = 0$ for $i > 0$, so $f(X) = a_0 X$.

Now suppose $F$ is finite with characteristic $p$ and size $q$. Then the equation $c^{p^i} = c$ is satisfied for all $c \in F$ if and only if $X^{p^i} - X$ vanishes on $F$, which is equivalent to $(X^q - X)|(X^{p^i} - X)$ in $F[X]$. Since $q$ is a power of $p$, such a divisibility relation holds only when $p^i$ is a power of $q$ (proof left as an exercise), which means the only terms in $f(X)$ with nonzero coefficients are those where the exponent of $X$ is a $q$-power. This makes $f(X)$ of the desired form. □

Theorem A.2 and Corollary A.3 and their proofs carry over from polynomials to power series: the additive and $F$-linear power series are the same as the corresponding polynomials except there need not be a final term in the series. Checking the details is left as an exercise.

## References


[1] Z. I. Borevich and I. R. Shafarevich, "Number Theory," Academic Press, New York, 1966.
[2] L. Carlitz, *On certain functions connected with polynomials in a Galois field*, Duke Math J. **1** (1935), 137–168.
[3] L. Carlitz, *A class of polynomials*, Trans. Amer. Math. Soc. **43** (1938), 167–182.
[4] D. R. Hayes, *Explicit class field theory for rational function fields*, Bull. Amer. Math. Soc. **189** (1974), 77–91.
[5] C-N. Hsu, *On Artin's conjecture for the Carlitz module*, Compositio Math. **106** (1997), 247–266.
[6] S. Lang, "Algebra," 3rd revised ed., Springer-Verlag, New York, 2002.
[7] M. Rosen, "Number Theory in Function Fields," Springer-Verlag, New York, 2002.
[8] L. I. Wade, *Certain quantities transcendental over $GF(p^n, x)$*, Duke Math J. **8** (1941), 701–720.