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 nonzero $M \in \mathbf{F}_p[T]$, 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 unit, and every nonzero polynomial in $\mathbf{F}_p[T]$ can be made monic (leading coefficient 1) after multiplication by a suitable unit. We will examine a deeper analogy: the group $(\mathbf{F}_p[T]/M)^\times$ can be interpreted as the Galois group of an extension of the field $\mathbf{F}_p(T)$ in a manner similar to the group $(\mathbf{Z}/(m))^\times$ being the Galois group of the mth cyclotomic extension $\mathbf{Q}(\mu_m)$ of $\mathbf{Q}$, where μ_m is the group of mth roots of unity.

For each $m \geq 1$, the mth 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 Galois extension of $\mathbf{F}_p(T)$ whose Galois group is isomorphic to $(\mathbf{F}_p[T]/M)^\times$.

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).

I thank Darij Grinberg for his extensive comments and corrections on the text below.

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

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

Example 2.1. For $n = 2$ and $n = 3$,

$$[T^2](X) = [T](X)^p + T[T](X) = (X^p + TX)^p + T(X^p + TX) = X^{p^2} + (T^p + T)X^p + T^2X$$

and

$$[T^3](X) = [T^2](X)^p + T[T^2](X) = X^{p^3} + (T^{p^2} + T^p + T)X^{p^2} + (T^{2p} + T^{p+1} + T^2)X^p + T^3X.$$

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

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

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 is taken from Lubin–Tate formal groups (see Remark 2.7).

Our examples suggest general properties of $[M](X)$. For instance, in $[T^2](X)$, $[T^3](X)$, and $[T^2 - T](X)$ we only see X appearing with p-power exponents and the lowest degree X-terms are, respectively, T^2X, T^3X, and $(T^2 - T)X$.

Definition 2.4. Let A be an integral domain of prime characteristic p. A *p-polynomial* over A is a polynomial in $A[X]$ that is an A-linear combination of X, X^p, X^{p^2}, and so on: $f(X) = a_0X + a_1X^p + a_2X^{p^2} + \cdots + a_dX^{p^d}$ for some $a_j \in A$.

Theorem 2.5. *For nonzero $M \in \mathbf{F}_p[T]$, $[M](X)$ has X-degree $p^{\deg M}$. Moreover, $[M](X)$ is a p-polynomial in X:*

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

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 8. They are analogues of binomial coefficients.

Corollary 2.6. *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 $D \mid M$ in $\mathbf{F}_p[T]$ then $[D](X) \mid [M](X)$ in $\mathbf{F}_p[T][X]$.

Proof. The basic polynomial $[T](X) = X^p + TX$ is a p-polynomial in X, and since other $[M](X)$ are defined by composition and $\mathbf{F}_p$-linearity from $[T](X)$, every $[M](X)$ is a p-polynomial in X. For a p-polynomial $f(X)$ we have $f(X+Y) = f(X)+f(Y)$ and $f(cX) = cf(X)$ for $c \in \mathbf{F}_p$.

That $M \mapsto [M](X)$ is additive in M and sends products to composites can be proved by induction on the degree of M.

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

The polynomials $[M](X)$ commute with each other under composition by Corollary 2.6: $[M_1]([M_2](X)) = [M_1M_2](X) = [M_2M_1](X) = [M_2]([M_1](X))$. This will be crucial later, since it will imply the roots of $[M](X)$ generate *abelian* Galois extensions of $\mathbf{F}_p(T)$.

[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 $X^p + TX$.

Remark 2.7. This is for readers who know Lubin–Tate theory. Over the power series ring $\mathbf{F}_p[[T]]$, $[T](X) = X^p + TX$ is a Frobenius polynomial for the uniformizer T. Since $[M](X)$ has lowest degree term MX and commutes with $[T](X)$, $[M](X)$ is the endomorphism attached to M of the Lubin–Tate formal group that has Frobenius polynomial $[T](X)$.

Corollary 2.8. *For $M(T) \in \mathbf{F}_p[T]$, the X-derivative of $[M](X)$ is M.*

For example, $[T](X) = X^p + TX$ has X-derivative T.

Proof. The derivative of a p-polynomial $a_0X + a_1X^p + a_2X^{p^2} + \cdots + a_dX^{p^d}$ is a_0 since $(X^{p^j})' = 0$ in characteristic p when $j \geq 1$, and $[M](X)$ has X-coefficient M. $\square$

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 mth 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 2.9. *For nonzero M in $\mathbf{F}_p[T]$, $[M](X)$ is separable in $\mathbf{F}_p(T)[X]$.*

Proof. The X-derivative of $[M](X)$ is $M \in \mathbf{F}_p[T]$ (Corollary 2.8), and M is a nonzero "constant" in $\mathbf{F}_p(T)[X]$. Therefore $[M](X)$ is relatively prime to its X-derivative, so $[M](X)$ is separable as a polynomial in X. $\square$

Corollary 2.10. *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 $M \mid N$ and $N \mid M$, 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 every root λ 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 2.9, 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 $M \mid N$. The argument that $N \mid M$ is similar, so we're done. $\square$

The rest of this section concerns analogies between the pth power map for prime p and the polynomial $[\pi](X)$ for (monic) irreducible π 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 2.11. *Let π 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 π. If $(M, \pi) = 1$ then $\overline{[M]}(X)$ is separable in $\mathbf{F}_\pi[X]$, while $\overline{[\pi]}(X) = X^{p^{\deg \pi}}$.*

Proof. By Corollary 2.8, $[M]'(X) = M$ and $[\pi]'(X) = \pi$. 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, $\overline{[\pi]}'(X) = \pi \bmod \pi$, and this 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 π 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 γ of $\overline{[\pi]}(X)$ in $\overline{\mathbf{F}_\pi}$ with $\gamma \neq 0$. We will get a contradiction. For all $M \in \mathbf{F}_p[T]$, $[M](\gamma)$ 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 2.6, 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 π. Since (π) is a maximal ideal, the kernel is (π), 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. $\square$

Corollary 2.12. *For every irreducible $\pi \in \mathbf{F}_p[T]$, the coefficients of $[\pi](X)$ besides its leading term are multiples of π. In particular, $[\pi](X)/X$ is an Eisenstein polynomial with respect to π with constant term π.*

Proof. For $c \in \mathbf{F}_p^\times$, $[c\pi](X) = c[\pi](X)$, so we may assume π is monic. Then the leading term of $[\pi](X)$ in $\mathbf{F}_p[T][X]$ is $X^{p^{\deg \pi}}$ and by Theorem 2.11, $\overline{[\pi]}(X) = X^{p^{\deg \pi}}$ 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 π. Since the lowest degree term of $[\pi](X)$ is πX, $[\pi](X)/X$ has constant term π and therefore is Eisenstein with respect to π. $\square$

Remark 2.13. 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 $[m_1 m_2](X) = [m_1]([m_2](X))$, which resembles part of Corollary 2.6 (but $[m](X)$ is *not* additive in m) and Corollary 2.12 resembles $[p](X)/X = ((1+X)^p - 1)/X$ being Eisenstein for prime p.

Corollary 2.14. *For every irreducible $\pi \in \mathbf{F}_p[T]$ and integer $k \geq 0$, the coefficients of $[\pi^k](X)$ besides its leading term are multiples of π.*

Proof. It's true for $k = 0$ and 1. For higher k use the identity $[\pi^k](X) = [\pi]([\pi^{k-1}](X))$. $\square$

Theorem 2.15. *For every monic irreducible π 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 2.11, $\overline{[\pi]}(X) = X^{p^{\deg \pi}}$ 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 $[\pi](A) \equiv A \bmod \pi$. $\square$

Subtracting $A = [1](A)$ from both sides of the congruence in Theorem 2.15, we get

Corollary 2.16. *For every monic irreducible π 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 $(\mathbf{Z}/(p))^\times$. However, Corollary 2.16 is true for *all* A, not just A that are relatively prime to π. That Fermat's little theorem is about a in the multiplicative group $(\mathbf{Z}/(p))^\times$ while Corollary 2.16 is about A in $\mathbf{F}_p[T]/\pi$ 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 2.17. *For monic irreducible π 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 π), 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 π is not monic then the above results have a more awkward form. Letting c be the leading coefficient of π, $[\pi](A) \equiv cA \bmod \pi$, $[\pi - c](A) \equiv 0 \bmod \pi$, and $f([\pi](X)) \equiv f(cX)^{p^{\deg \pi}} \bmod \pi$. Just remember that monic π 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 π are monic.

The following theorem is an analogue of $X^m - 1$ having no roots in $\mathbf{Z}$ other than 1 if m is odd and other than ± 1 if m is even, except it's more tedious to prove.

Theorem 2.18. *For nonzero M and A in $\mathbf{F}_p[T]$, $[M](A) \neq 0$ unless perhaps $p = 2$ and $\deg A \leq 1$ because in other cases we can compute the degree of $[M](A)$:*

$$\deg([M](A)) = \begin{cases} p^{\deg M} \deg A, & \text{if } p \neq 2, \deg A \geq 1, \\ p^{\deg M} \deg A, & \text{if } p = 2, \deg A \geq 2, \\ p^{\deg M - 1}, & \text{if } p \neq 2, \deg A = 0, \deg M \geq 2, \\ \deg M, & \text{if } p \neq 2, \deg A = 0, \deg M = 0 \text{ or } 1. \end{cases}$$

If $p = 2$ and $\deg A \leq 1$ then $[M](A)$ is sometimes 0, *e.g.* $[T](T) = 0$ and $[T^2 + T](1) = 0$, so $[M](T) = 0$ if $T \mid M$ and $[M](1) = 0$ if $(T^2 + T) \mid M$.

Proof. Writing $M = \sum_{j=0}^{\deg M} c_j T^j$, we have $[M](A) = \sum_{j=0}^{\deg M} c_j [T^j](A)$, so to prove that $\deg([M](A)) = p^{\deg M} \deg A$ in the first and second cases it suffices to show $\deg([T^j](A)) = p^j \deg A$ for all $j \geq 0$.

When $j = 0$, $[T^0](A) = [1](A) = A$, which has degree $\deg A$.

When $j = 1$, $[T](A) = A^p + TA$, which has degree $p \deg A$ if $\deg(A^p) > \deg(TA)$. That inequality is the same as $(p-1)\deg A > 1$, and this holds when $p \neq 2$ and $\deg A \geq 1$, or when $p = 2$ and $\deg A \geq 2$.

Now assume by induction that $\deg([T^j](A)) = p^j \deg A$ for some $j \geq 1$. To show $\deg([T^{j+1}](A)) = p^{j+1} \deg A$, write $[T^{j+1}](A) = [T]([T^j](A)) = [T^j](A)^p + T[T^j](A)$. The first term has degree $p^{j+1} \deg A$ and the second term has degree $1 + p^j \deg A$. The inequality $p^{j+1} \deg A > 1 + p^j \deg A$ is equivalent to $p^j(p-1)\deg A > 1$, and that's true for nonconstant A if $p \neq 2$, and for $\deg A \geq 2$ (or even $\deg A \geq 1$) if $p = 2$.

To show $\deg([M](a)) = p^{\deg M - 1}$ when $p \neq 2$, $a \in \mathbf{F}_p^\times$, and $\deg M \geq 2$, first we compute $[T](a) = a^p + Ta$ and $[T^2](a) = a^{p^2} + (T^p + T)a^p + T^2 a$, so $\deg([M](a)) = p$ when $\deg M = 2$. Assuming $\deg([T^i](a)) = p^{i-1}$ for some $i \geq 2$, write $[T^{i+1}](a) = [T]([T^i](a)) = [T^i](a)^p + T[T^i](a)$ and we see by induction that the first term has larger degree than the second, so $\deg([T^{i+1}](a)) = p^i$.

Lastly, if $p \neq 2$, $a \in \mathbf{F}_p^\times$, and $M(T) = cT + d$ with $c, d \in \mathbf{F}_p$ (not both 0) then $[M](a) = caT + (ca^p + da)$, which has the same degree as M.

□

3. The Carlitz module and Carlitz torsion

Let K be a field extension of $\mathbf{F}_p(T)$. We can view K as an $\mathbf{F}_p(T)$-vector space in the usual way, so it is also an $\mathbf{F}_p[T]$-module by multiplication. Using the Carlitz polynomials we can define a *different* $\mathbf{F}_p[T]$-module structure on K, as follows.

Definition 3.1. Let K be a field extension of $\mathbf{F}_p(T)$. We make the additive group of K into an $\mathbf{F}_p[T]$-module by letting $\mathbf{F}_p[T]$ act on K using the Carlitz polynomials: for $M \in \mathbf{F}_p[T]$ and $\alpha \in K$, define

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

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

Example 3.2. In the Carlitz action, $T \cdot \alpha = [T](\alpha) = \alpha^p + T\alpha$, which is not $T\alpha$ (if $\alpha \neq 0$).

Example 3.3. In the Carlitz action, $c \cdot \alpha = [c](\alpha) = c\alpha$ for $c \in \mathbf{F}_p$. For nonconstant M in $\mathbf{F}_p[T]$, $M \cdot \alpha$ is essentially never the same as $M\alpha$ when $\alpha \neq 0$.

That the Carlitz action is an $\mathbf{F}_p[T]$-module structure on K amounts to several identities:

$$M \cdot (\alpha + \beta) = M \cdot \alpha + M \cdot \beta$$

for $M \in \mathbf{F}_p[T]$ and α and β in K, and

$$(M_1 + M_2) \cdot \alpha = M_1 \cdot \alpha + M_2 \cdot \alpha, \quad M_1 \cdot (M_2 \cdot \alpha) = (M_1 M_2) \cdot \alpha, \quad 1 \cdot \alpha = \alpha$$

for M_i in $\mathbf{F}_p[T]$ and $\alpha \in K$. These identities follow from Corollary 2.6 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 α and β in K we get $[M](\alpha + \beta) = [M](\alpha) + [M](\beta)$, which says $M \cdot (\alpha + \beta) = M \cdot \alpha + M \cdot \beta$.

There are two ways to make the additive group of K into an $\mathbf{F}_p[T]$-module, namely ordinary multiplication of $\mathbf{F}_p[T]$ on K and the action of $\mathbf{F}_p[T]$ on K through Carlitz polynomials, so to avoid ambiguity we want to denote K differently when it is an $\mathbf{F}_p[T]$-module in each way. A plain K will mean K as an $\mathbf{F}_p[T]$-module by multiplication of $\mathbf{F}_p[T]$, while $C(K)$ will mean K as an $\mathbf{F}_p[T]$-module by the Carlitz action. The second way, as $C(K)$, is more interesting.

The Carlitz action on the field $\overline{\mathbf{F}_p(T)}$, an algebraic closure of $\mathbf{F}_p(T)$, will be of particular importance, and the $\mathbf{F}_p[T]$-module $C(\overline{\mathbf{F}_p(T)})$ is called the *Carlitz module*; it's $\overline{\mathbf{F}_p(T)}$ with a subtle $\mathbf{F}_p[T]$-module structure, and 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 $m > 0$; 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 $M \neq 0$, and we will see in Section 5 that such α generate abelian extensions of $\mathbf{F}_p(T)$.

Definition 3.4. Let $\Lambda_M = \{\lambda \in \overline{\mathbf{F}_p(T)} : [M](\lambda) = 0\}$. This is called the *$M$-torsion* of the Carlitz module. The *Carlitz torsion* is the union of Λ_M over all nonzero $M \in \mathbf{F}_p[T]$.

Example 3.5. Since $[T](X) = X^p + TX = X(X^{p-1} + T)$,

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

which is analogous to $\mu_p = \{z \in \overline{\mathbf{Q}} : z^p = 1\} = \{1\} \cup \{z \in \overline{\mathbf{Q}} : \Phi_p(z) = 0\}$.

The polynomial $X^{p-1} + T$ is irreducible since it is Eisenstein with respect to T. If α is one root of $X^{p-1} + T$, then all the roots are $\{c\alpha : c \in \mathbf{F}_p^\times\}$, so $\mathbf{F}_p(T, \Lambda_T) = \mathbf{F}_p(T, \alpha)$, which is a Kummer extension of $\mathbf{F}_p(T)$ with degree $p - 1$. It is analogous to the pth cyclotomic extension $\mathbf{Q}(\zeta_p)/\mathbf{Q}$, also of degree $p - 1$ (and both have cyclic Galois group).

Example 3.6. Since $[T^2](X) = [T]([T](X)) = (X^p + TX)^p + T(X^p + TX)$,

$$\Lambda_{T^2} = \{\lambda \in \overline{\mathbf{F}_p(T)} : \lambda^p + T\lambda \in \Lambda_T\} = \Lambda_T \cup \{\lambda \in \overline{\mathbf{F}_p(T)} : (\lambda^p + T\lambda)^{p-1} = -T\}.$$

This is analogous to $\mu_{p^2} = \{z \in \overline{\mathbf{Q}} : z^p \in \mu_p\} = \mu_p \cup \{z \in \overline{\mathbf{Q}} : \Phi_p(z^p) = 0\}$.

If $\beta \in \Lambda_{T^2} - \Lambda_T$ (the elements of Λ_{T^2} not in Λ_T) then β is a root of $(X^p + TX)^{p-1} + T$, which is irreducible over $\mathbf{F}_p(T)$ since it is Eisenstein with respect to T, so $[\mathbf{F}_p(T, \beta) : \mathbf{F}_p(T)] = p(p-1)$. Moreover, $\alpha := \beta^p + T\beta$ is a nonzero element of Λ_T and $\Lambda_{T^2} = \{a\alpha + b\beta : a, b \in \mathbf{F}_p\}$, so $\mathbf{F}_p(T, \Lambda_T) = \mathbf{F}_p(T, \alpha)$, $\mathbf{F}_p(T, \Lambda_{T^2}) = \mathbf{F}_p(T, \beta)$, and $\mathbf{F}_p(T, \beta)/\mathbf{F}_p(T)$ is a Galois extension. What does the Galois group look like?

All elements of $\Lambda_{T^2} - \Lambda_T$ have the same minimal polynomial (namely $(X^p + TX)^{p-1} + T$), and $a\alpha + b\beta \notin \Lambda_T$ when $b \neq 0$, so the $\mathbf{F}_p(T)$-conjugates of β are all $a\alpha + b\beta$ with $a \in \mathbf{F}_p$ and $b \in \mathbf{F}_p^\times$. This is analogous to saying if $\zeta_{p^2} \in \mu_{p^2} - \mu_p$ then $\zeta_p := \zeta_{p^2}^p$ is a nontrivial element of μ_p and $\mu_{p^2} = \{\zeta_p^a \zeta_{p^2}^b : 0 \leq a, b \leq p - 1\} = \{\zeta_{p^2}^{ap+b}\}$, with $\zeta_{p^2}^{ap+b} \notin \mu_p$ when $b \neq 0$.

The $\mathbf{F}_p(T)$-conjugates of β can be written as

$$a\alpha + b\beta = a(\beta^p + T\beta) + b\beta = [aT + b](\beta),$$

where $a \in \mathbf{F}_p$ and $b \in \mathbf{F}_p^\times$. Then $\mathrm{Gal}(\mathbf{F}_p(T, \Lambda_{T^2})/\mathbf{F}_p(T)) \cong (\mathbf{F}_p[T]/(T^2))^\times$ by $\sigma \mapsto aT + b \bmod T^2$, where $\sigma(\beta) = [aT + b](\beta)$. This is analogous to $\mathrm{Gal}(\mathbf{Q}(\zeta_{p^2})/\mathbf{Q}) \cong (\mathbf{Z}/(p^2))^\times$.

By Theorem 2.9, $[M](X)$ has $p^{\deg M}$ different 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 μ_m for all m. Since $[M](X)$ is a p-polynomial in X, its roots Λ_M form a finite $\mathbf{F}_p$-vector space. But Λ_M is more than a vector space:

Theorem 3.7. *The set Λ_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 2.6,

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

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

Example 3.8. In Example 3.5 we saw $\Lambda_T = \mathbf{F}_p\alpha$ when α is one root of $X^{p-1} + T$, so Λ_T 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 $A = TQ + A(0)$,

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

Here is a Carlitz analogue of the group isomorphism $\mu_{mn} \cong \mu_m \times \mu_n$ when $(m, n) = 1$.

Theorem 3.9. *If M and N are relatively prime in $\mathbf{F}_p[T]$ then $\Lambda_{MN} \cong \Lambda_M \oplus \Lambda_N$ as $\mathbf{F}_p[T]$-modules.*

Proof. Let $\Lambda_M \oplus \Lambda_N \to \Lambda_{MN}$ by ordinary addition: $(\lambda, \lambda') \mapsto \lambda + \lambda'$. This map makes sense since Λ_M and Λ_N are submodules of Λ_{MN}. The map is $\mathbf{F}_p[T]$-linear and both $\Lambda_M \oplus \Lambda_N$ and Λ_{MN} have the same finite size. Therefore to be an isomorphism it suffices to be injective. If $\lambda + \lambda' = 0$ then $\lambda = -\lambda'$, so λ and λ' belong to $\Lambda_M \cap \Lambda_N$, which is $\{0\}$ since M and N are relatively prime. Therefore $(\lambda, \lambda') = (0, 0)$. $\square$

The group structure on μ_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 μ_m is cyclic, and if ζ generates μ_m then ζ^a is a generator of μ_m if and only if $(a, m) = 1$. Exactly the same properties apply to Λ_M:

Theorem 3.10. *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 Λ_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 $\lambda_0 \in \Lambda_M$ that is a Carlitz generator:*

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

and the generators of Λ_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 $M \mid A$.

To prove Λ_M has a generator as an $\mathbf{F}_p[T]$-module, we will adapt a proof that μ_m is a cyclic group. Here is a proof that the group μ_m is cyclic. In a finite abelian group, if there is an element of order n_1 and an element of order n_2 then there is an element whose order is the least common multiple $[n_1, n_2]$. Writing n for the largest order of the elements of μ_m, and n' for the order of some element of μ_m, there is an element in μ_m of order $[n, n']$. Since $[n, n'] \geq n$ and n is the maximal order, we must have $[n, n'] = n$, so $n' \mid n$: all orders divide the largest order. Therefore every element x of μ_m satisfies $x^n = 1$, so the polynomial

$X^n - 1$ has at least m roots, which implies $m \leq n$. Also $n \mid m$, since the order of an element divides the size of the group, so $n = m$: there is an element of μ_m with order m.

Consider now Λ_M instead of μ_m. While μ_m is a finite abelian group, Λ_M is a finitely generated (even finite) torsion $\mathbf{F}_p[T]$-module. In every finitely generated torsion $\mathbf{F}_p[T]$-module Λ, we can associate to each element $\lambda \in \Lambda$ its $\mathbf{F}_p[T]$-*order*, which is the unique monic generator of the annihilator ideal

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

As with finite abelian groups, if N_1 and N_2 are $\mathbf{F}_p[T]$-orders of elements of Λ then there is an element of Λ whose $\mathbf{F}_p[T]$-order is the least common multiple $[N_1, N_2]$. It follows that the $\mathbf{F}_p[T]$-order with largest degree is divisible by the $\mathbf{F}_p[T]$-order of every element of Λ. What this means in the case of Λ_M is that if N denotes the $\mathbf{F}_p[T]$-order of largest degree in Λ_M then every $\lambda \in \Lambda_M$ satisfies $[N](\lambda) = 0$, so $|\Lambda_M| \leq \deg([N](X)) = p^{\deg N}$, or equivalently $p^{\deg M} \leq p^{\deg N}$. Also $N \mid M$ (analogue of all orders in a group dividing the size of the group), so N is the monic scalar multiple of M. Letting $\lambda_0 \in \Lambda_M$ have maximal $\mathbf{F}_p[T]$-order N, $\mathrm{Ann}_{\Lambda_M}(\lambda_0) = (N) = (M)$, so the $\mathbf{F}_p[T]$-submodule that λ_0 generates in Λ_M has size

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

which shows λ_0 is a generator of Λ_M and there is an $\mathbf{F}_p[T]$-module isomorphism $\mathbf{F}_p[T]/M \cong \Lambda_M$ given by $A \bmod M \mapsto [A](\lambda_0)$. In particular, $[A](\lambda_0)$ generates Λ_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 ζ of μ_m gives a noncanonical group isomorphism $\mathbf{Z}/(m) \cong \mu_m$ by $a \bmod m \mapsto \zeta^a$, and in the same way choosing a generator λ_0 of Λ_M leads to a noncanonical $\mathbf{F}_p[T]$-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.

Corollary 3.11. *The $\mathbf{F}_p[T]$-submodules of Λ_M are all Λ_D where D divides M.*

Proof. Fix a generator λ_0 of Λ_M. Then $\mathbf{F}_p[T]/M \cong \Lambda_M$ as $\mathbf{F}_p[T]$-modules by $A \bmod M \mapsto [A](\lambda_0)$, so the result is a consequence of the submodules of $\mathbf{F}_p[T]/M$ being $D\mathbf{F}_p[T]/M$ for $D \mid M$, with the submodule $D\mathbf{F}_p[T]/M$ corresponding to $\Lambda_{M/D}$. □

4. Structure of $\mathbf{F}_p[T]/M$ with Carlitz Action

The Carlitz analogue of the cyclic group $\mathbf{Z}/(m)$ is the $\mathbf{F}_p[T]$-module Λ_M, which is (noncanonically) isomorphic to $\mathbf{F}_p[T]/M$. A Carlitz analogue of $(\mathbf{Z}/(m))^\times$ is the additive group $\mathbf{F}_p[T]/M$ with a new $\mathbf{F}_p[T]$-module structure: $N \cdot (A \bmod M) = [N](A) \bmod M$ for $N \in \mathbf{F}_p[T]$. We denote $\mathbf{F}_p[T]/M$ with this Carlitz action by $\mathbf{F}_p[T]$ as $C(\mathbf{F}_p[T]/M)$.

Example 4.1. The $\mathbf{F}_3[T]$-module $C(\mathbf{F}_3[T]/(T^2+1))$ is generated by 1. See Table 1.

A	0	1	2	T	$T+1$	$T+2$	$2T$	$2T+1$	$2T+2$
$[A](1) \bmod T^2+1$	0	1	2	$T+1$	$T+2$	T	$2T+2$	$2T$	$2T+1$

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

More generally, if $\pi(T)$ is a monic quadratic irreducible in $\mathbf{F}_p[T]$, then 1 is a generator of $C(\mathbf{F}_p[T]/(\pi(T)))$: $[a_1T + a_2](1) = a_1T + (a_1 + a_2) \equiv 0 \bmod \pi(T)$ only if $a_1 = 0$ and $a_2 = 0$.

Example 4.2. The $\mathbf{F}_2[T]$-module $C(\mathbf{F}_2[T]/(T^3+T+1))$ is not generated by 1: show $[T^2+T](1) \equiv 0 \bmod T^3+T+1$ and T^2 is a generator.

The passage from K to $C(K)$, where K is a field extension of $\mathbf{F}_p(T)$, and from $\mathbf{F}_p[T]/M$ to $C(\mathbf{F}_p[T]/M)$, are special cases of a more general construction: for each $\mathbf{F}_p[T]$-algebra $\mathcal{A}$ (such as a field extension of $\mathbf{F}_p(T)$ or the ring $\mathbf{F}_p[T]/M$), it makes sense to evaluate polynomials in $\mathbf{F}_p[T][X]$ at elements of $\mathcal{A}$, and setting $M \cdot a = [M](a)$ makes the additive group $\mathcal{A}$ 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 polynomial values acting on $\mathcal{A}$. (One might call $C(\mathcal{A})$ the "Carlitzification" of $\mathcal{A}$.) The proof of Theorem 2.11, for instance, was 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 $C(\mathcal{A})$ respects 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 (a $\mathbf{Z}$-algebra) to its unit group $A^\times$ (a $\mathbf{Z}$-module). In particular, if $\mathcal{A}$ and $\mathcal{B}$ are isomorphic $\mathbf{F}_p[T]$-algebras then $C(\mathcal{A})$ and $C(\mathcal{B})$ are isomorphic $\mathbf{F}_p[T]$-modules. For instance, the Chinese remainder theorem shows $\mathbf{F}_p[T]/M_1M_2 \cong \mathbf{F}_p[T]/M_1 \times \mathbf{F}_p[T]/M_2$ as $\mathbf{F}_p[T]$-algebras if M_1 and M_2 are relatively prime, so $C(\mathbf{F}_p[T]/M_1M_2) \cong C(\mathbf{F}_p[T]/M_1) \times C(\mathbf{F}_p[T]/M_2)$ as $\mathbf{F}_p[T]$-modules. Thus $C(\mathbf{F}_p[T]/M)$ decomposes into a direct product of $\mathbf{F}_p[T]$-modules $C(\mathbf{F}_p[T]/\pi^k)$ for monic irreducible π, so understanding the structure of $C(\mathbf{F}_p[T]/M)$ as an $\mathbf{F}_p[T]$-module boils down to the case when $M = \pi^k$. Think of $C(\mathbf{F}_p[T]/\pi^k)$ as analogous to $(\mathbf{Z}/(p^k))^\times$ when p is prime, since it makes the results below on $C(\mathbf{F}_p[T]/\pi^k)$ reasonable.

Let's first treat $k = 1$.

Theorem 4.3. *For monic irreducible π in $\mathbf{F}_p[T]$, the $\mathbf{F}_p[T]$-module $C(\mathbf{F}_p[T]/\pi)$ is cyclic. It is isomorphic to $\mathbf{F}_p[T]/(\pi-1)$ as $\mathbf{F}_p[T]$-modules.*

This is like $(\mathbf{Z}/(p))^\times$ being a cyclic group of order $p-1$ if p is prime: $(\mathbf{Z}/(p))^\times \cong \mathbf{Z}/(p-1)$ as abelian groups. (Note $(\mathbf{F}_p[T]/\pi)^\times \not\cong \mathbf{F}_p[T]/(\pi-1)$ as abelian groups since the sizes don't match.)

Proof. Let's recall a proof that $(\mathbf{Z}/(p))^\times$ is cyclic and then adapt it to the Carlitz setting. Writing $p-1 = q_1^{e_1}\cdots q_k^{e_k}$ with distinct primes q_i and $e_i \geq 1$, we will find an $a_i \in (\mathbf{Z}/(p))^\times$ with order $q_i^{e_i}$. Then the product $a_1\cdots a_k$ will be an element of order $p-1$.

The polynomial $X^{p-1}-1$ splits completely in $\mathbf{F}_p[X]$ with distinct roots, so its factor $X^{q_i^{e_i}}-1$ also splits completely over $\mathbf{F}_p$ with distinct roots. Therefore $X^{q_i^{e_i}}-1$ has a root that is not a root of $X^{q_i^{e_i-1}}-1$, and such a root a_i will have order $q_i^{e_i}$.

In the Carlitz setting, $C(\mathbf{F}_p[T]/\pi)$ is an $\mathbf{F}_p[T]/(\pi-1)$-module since $[\pi-1](A) \equiv 0 \bmod \pi$ for all $A \in \mathbf{F}_p[T]$ (Corollary 2.16).[2] We seek $A_0 \bmod \pi \in C(\mathbf{F}_p[T]/\pi)$ with annihilator ideal $(\pi-1)$, so $\mathbf{F}_p[T]/(\pi-1) \cong C(\mathbf{F}_p[T]/\pi)$ as $\mathbf{F}_p[T]$-modules by $M \bmod \pi-1 \mapsto [M](A_0) \bmod \pi$.

Factor $\pi-1$ in $\mathbf{F}_p[T]$ as $\pi_1^{e_1}\cdots\pi_k^{e_k}$ with π_i being distinct monic irreducibles and $e_i \geq 1$. For $i = 1,\dots,k$ we will find an $A_i \bmod \pi \in C(\mathbf{F}_p[T]/\pi)$ with annihilator ideal $(\pi_i^{e_i})$. Then the *sum* $A_1 + A_2 + \cdots + A_k \bmod \pi$ will have annihilator ideal $(\pi-1)$.

[2] While π acts by ordinary multiplication on $\mathbf{F}_p[T]/\pi$ as 0, since $\pi A \equiv 0 \bmod \pi$, π acts by the Carlitz action on $C(\mathbf{F}_p[T]/\pi)$ as the *identity*, since $[\pi](A) \equiv A \bmod \pi$ (Theorem 2.15).

Since $\pi_i^{e_i} \mid (\pi-1)$, $[\pi_i^{e_i}](X) \mid [\pi-1](X)$ (Corollary 2.6). The polynomial $[\pi-1](X)$ has X-degree $p^{\deg(\pi-1)} = p^{\deg\pi} = |\mathbf{F}_p[T]/\pi|$ and vanishes at each element of $\mathbf{F}_p[T]/\pi$, so $[\pi-1](X)$ splits completely with distinct roots over $\mathbf{F}_p[T]/\pi$. Therefore $[\pi_i^{e_i}](X)$ also splits completely over $\mathbf{F}_p[T]/\pi$ with distinct roots. By comparing degrees, $[\pi_i^{e_i}](X)$ has a root in $\mathbf{F}_p[T]/\pi$ that is not a root of its factor $[\pi_i^{e_i-1}](X)$, and every such root in $\mathbf{F}_p[T]/\pi$ has annihilator ideal $(\pi_i^{e_i})$. The sum of such roots for each i gives us a generator of $C(\mathbf{F}_p[T]/\pi)$. □

The structure of $C(\mathbf{F}_p[T]/\pi^k)$ as an $\mathbf{F}_p[T]$-module for $k \geq 2$ resembles the structure of $(\mathbf{Z}/(p^k))^\times$ as an abelian group, so let's recall what that is as motivation.

Lemma 4.4. *Let $k \geq 2$.*

1) $(\mathbf{Z}/(2^k))^\times = \langle -1 \bmod 2^k\rangle \times \langle 5 \bmod 2^k\rangle \cong \mathbf{Z}/(2) \times \mathbf{Z}/(2^{k-2})$.
2) *For odd prime p, $(\mathbf{Z}/(p^k))^\times = C_{p-1} \times \langle 1+p \bmod p^k\rangle \cong \mathbf{Z}/(p-1)\times\mathbf{Z}/(p^{k-1})$, where C_{p-1} is cyclic of order $p-1$.*[3]

Proof. 1) By induction, $5^{2^r} \equiv 1+2^{r+2} \bmod 2^{r+3}$ for all $r \geq 0$. Therefore $5 \bmod 2^k$ has order 2^{k-2} for $k \geq 2$. The powers of 5 mod 2^k are all $\equiv 1 \bmod 4$, so they don't include $-1 \bmod 2^k$. Therefore by counting we get $(\mathbf{Z}/(2^k))^\times = \langle -1, 5 \bmod 2^k\rangle \cong \langle -1\rangle \times \langle 5 \bmod 2^k\rangle$.

2) To write down an element with order $p-1$ in $(\mathbf{Z}/(p^k))^\times$, we will use the fact that raising to the pth power is a well-defined function $\mathbf{Z}/(p^r) \to \mathbf{Z}/(p^{r+1})$ for $r \geq 1$:

$$a \equiv b \bmod p^r \Rightarrow a^p \equiv b^p \bmod p^{r+1}. \tag{4.1}$$

This follows from the intermediate binomial coefficients in $(X+Y)^p$ all being divisible by p. We are interested in this function on units: define $f_{p,r}\colon (\mathbf{Z}/(p^r))^\times \to (\mathbf{Z}/(p^{r+1}))^\times$ by $f_{p,r}(a \bmod p^r) = a^p \bmod p^{r+1}$, which is a homomorphism. For $p \neq 2$ $f_{p,r}$ is *injective*. See Table 2 for data when $p=3$. To show $f_{p,r}$ is injective, let $u \bmod p^r$ be in the kernel, so $u^p \equiv 1 \bmod p^{r+1}$. Then $u^p \equiv 1 \bmod p$, so $u \equiv 1 \bmod p$. Suppose for some positive integer $i \leq r-1$ that $u \equiv 1 \bmod p^i$. Write $u \equiv 1 + cp^i \bmod p^{i+1}$. For $p \neq 2$,

$$u \equiv 1+cp^i \bmod p^{i+1} \text{ for } i \geq 1 \Longrightarrow u^p \equiv 1 + cp^{i+1} \bmod p^{i+2}. \tag{4.2}$$

Since $i+2 \leq r+1$, we get $1+cp^{i+1} \equiv 1 \bmod p^{i+2}$, so $p \mid c$, and thus $u \equiv 1 \bmod p^{i+1}$. Repeating this for $i = 1, 2, \ldots, r-1$ gives us $u \equiv 1 \bmod p^r$.

a mod 9	1	2	4	5	7	8
a^3 mod 27	1	8	10	17	19	26

TABLE 2. Values of $f_{3,2}$ on $(\mathbf{Z}/(9))^\times$.

The implication (4.2) breaks down for $i=1$ if $p=2$, and of course squaring $(\mathbf{Z}/(2^r))^\times \to (\mathbf{Z}/(2^{r+1}))^\times$ is not injective when $r \geq 2$.

Each $f_{p,r}$ preserves orders of elements by injectivity, so for $a \bmod p$ in $(\mathbf{Z}/(p))^\times$ with order $p-1$, $a^{p^{k-1}} \bmod p^k$ has order $p-1$ in $(\mathbf{Z}/(p^k))^\times$ since $a^{p^{k-1}} \bmod p^k$ equals $f_{p,k-1}\circ\cdots\circ f_{p,2}\circ f_{p,1}(a \bmod p)$. From $1+p \equiv 1 \bmod p$ we get $(1+p)^{p^{k-1}} \equiv 1 \bmod p^k$, and from $1+p \not\equiv 1 \bmod p^2$ and injectivity of $f_{p,r}$ for $p \neq 2$ we get $(1+p)^{p^{k-2}} \not\equiv 1 \bmod p^k$, so $1+p \bmod p^k$ has order p^{k-1}. Counting orders, $(\mathbf{Z}/(p^k))^\times = \langle a^{p^{k-1}}, 1+p\rangle \cong \langle a^{p^{k-1}}\rangle \times \langle 1+p\rangle$. □

[3]For $p \neq 2$, there is *no explicit formula* for a generator of C_{p-1}, although such a number for $k > 1$ can be expressed in terms of a choice of a generator of $(\mathbf{Z}/(p))^\times$.

Theorem 4.5. *Let π be monic irreducible in $\mathbf{F}_p[T]$ and $k \geq 2$.*

1) *For $\pi = T$ or $T+1$ in $\mathbf{F}_2[T]$, $C(\mathbf{F}_2[T]/\pi^k) = C_1 \times C_2$ where $C_1 \cong \mathbf{F}_2[T]/(T^2+T)$ and $C_2 \cong \mathbf{F}_2[T]/(\pi^{k-2})$ as $\mathbf{F}_2[T]$-modules, with respective generators 1 and π^2.*
2) *If $(p, \deg \pi) \neq (2,1)$ then $C(\mathbf{F}_p[T]/\pi^k) = C_1 \times C_2$, where $C_1 \cong \mathbf{F}_p[T]/(\pi - 1)$ and $C_2 \cong \mathbf{F}_p[T]/\pi^{k-1}$ as $\mathbf{F}_p[T]$-modules, with C_2 generated by $\pi \bmod \pi^k$.*

Since $\pi - 1$ and π^{k-1} are relatively prime, part 2 says $C(\mathbf{F}_p[T]/\pi^k) \cong \mathbf{F}_p[T]/((\pi-1)\pi^{k-1})$ as $\mathbf{F}_p[T]$-modules, which resembles $(\mathbf{Z}/p^k)^\times \cong \mathbf{Z}/((p-1)p^{k-1})$ as abelian groups for $p \neq 2$. (Note $(\mathbf{F}_p[T]/\pi^k)^\times \not\cong \mathbf{F}_p[T]/((\pi-1)\pi^{k-1})$ as abelian groups since the sizes don't match.)

Proof. 1) It suffices to give the proof for $\pi = T$; the proof for $\pi = T+1$ is the same by using $T+1$ in place of T everywhere[4] since $\mathbf{F}_2[T] = \mathbf{F}_2[T+1]$, $T^2 + T = (T+1)^2 + (T+1)$, and $[T+1](X) = X^2 + (T+1)X$.

By induction, in $\mathbf{F}_2[T]$ we have $[T^r](T^2) \equiv T^{r+2} \bmod T^{r+3}$ for all $r \geq 0$. Therefore in $C(\mathbf{F}_2[T]/(T^k))$ the annihilator ideal of T^2 is (T^{k-2}) for $k \geq 2$, so the submodule C_2 generated by T^2 has size $|\mathbf{F}_2[T]/(T^{k-2})| = 2^{k-2}$. This submodule is contained in $T^2\mathbf{F}_2[T]/(T^k)$, whose cardinality is 2^{k-2}, so $C_2 = T^2\mathbf{F}_2[T]/(T^k)$ as a subset of $C(\mathbf{F}_2[T]/(T^k))$.

The submodule C_1 of $C(\mathbf{F}_2[T]/(T^k))$ generated by 1 is $\{0, 1, T, T+1 \bmod T^k\}$ since $[1](1) = 1$ and $[T^r](1) = T+1$ for $r \geq 1$. Since $|C_1| = 4$ and $C_1 \cap C_2 = \{0\}$, $C(\mathbf{F}_2[T]/(T^k)) = C_1 + C_2 \cong C_1 \times C_2$ as $\mathbf{F}_p[T]$-modules.[5]

2) Because $[\pi](X) = X^{p^{\deg \pi}} + \cdots + \pi X$ has all non-leading coefficients divisible by π, we get an analogue of (4.1): for all A and B in $\mathbf{F}_p[T]$, and $r \geq 1$,

$$A \equiv B \bmod \pi^r \Rightarrow [\pi](A) \equiv [\pi](B) \bmod \pi^{r+1}. \tag{4.3}$$

Let $L_{\pi,r} \colon C(\mathbf{F}_p[T]/\pi^r) \to C(\mathbf{F}_p[T]/\pi^{r+1})$ for $r \geq 1$ by $L_{\pi,r}(A \bmod \pi^r) = [\pi](A) \bmod \pi^{r+1}$. This makes sense by (4.3) and it is $\mathbf{F}_p[T]$-linear. Except when $p = 2$ and $\deg \pi = 1$ ($\Rightarrow \pi = T$ or $T+1$) we will show $L_{\pi,r}$ is *injective* by checking its kernel is 0. It is not injective if $p = 2$, $\deg \pi = 1$, and $r \geq 2$ since $[\pi](\pi) = 0$ for $\pi = T$ or $T+1$ in $\mathbf{F}_2[T]$, and thus $[\pi](A + \pi) = [\pi](A)$. See examples in Table 3.

$A \bmod T^2$	0	1	T	$T+1$	$A \bmod (T+1)^2$	0	1	T	$T+1$
$[T](A) \bmod T^3$	0	$T+1$	0	$T+1$	$[T+1](A) \bmod (T+1)^3$	0	T	T	0

TABLE 3. Noninjectivity of $L_{T,2}$ and $L_{T+1,2}$ when $p = 2$.

Suppose $A \bmod \pi^r$ is in the kernel of $L_{\pi,r}$, so $[\pi](A) \equiv 0 \bmod \pi^{r+1}$. Reducing both sides mod π, we get $A^{p^{\deg \pi}} \equiv 0 \bmod \pi$, so $\pi \mid A$. If $r = 1$ then $A \equiv 0 \bmod \pi^r$, so we're done. Take $r \geq 2$. To show $A \equiv 0 \bmod \pi^r$, assume otherwise, and write the highest power of π that divides A as π^d, so $1 \leq d \leq r-1$. We will compute the highest power of π dividing the polynomial $[\pi](A)$ and compare the result with π^{r+1}, which we know divides $[\pi](A)$.

In $[\pi](A) = A^{p^{\deg \pi}} + \cdots + \pi A$, the first term is divisible by $\pi^{dp^{\deg \pi}}$ and all intermediate terms are divisible at least by π^{1+pd} since $[\pi](X)$ is a p-polynomial whose intermediate

[4] The Carlitz modules $C(\mathbf{F}_2[T]/T^k)$ and $C(\mathbf{F}_2[T]/(T+1)^k)$ are not isomorphic for $k \geq 4$, since one module has annihilator ideal $T^{k-2}(T+1)$ and the other has annihilator ideal $(T+1)^{k-2}T$.

[5] The "universal" subgroup $\pm 1 \bmod 2^k$ in each $(\mathbf{Z}/(2^k))^\times$ is analogous to the "universal" submodule $\{0, 1, T, T+1 \bmod T^k\}$ in each $C(\mathbf{F}_2[T]/T^k)$, and this submodule is generated by 1.

coefficients are divisible by π. The last term πA is divisible by π^{1+d} but not by a higher power of π. Since $1+pd > 1+d$ and $dp^{\deg \pi} > 1+d$ *unless* $p=2$, $\deg \pi = 1$, and $d=1$, the highest power of π dividing $[\pi](A)$ is π^{1+d} unless $p=2$, $\deg\pi=1$, and $d=1$. We have $\pi^{r+1} \mid [\pi](A)$ by hypothesis, so $\pi^{r+1} \mid \pi^{1+d}$ unless $p=2$, $\deg\pi = 1$, and $d=1$. Since $d \leq r-1$ we get a contradiction unless $p=2$, $\deg \pi = 1$, and $d=1$. In Table 3, $L_{\pi,2}(\pi \bmod \pi^2) \equiv [\pi](\pi) \equiv 0 \bmod \pi^3$.

An injective $\mathbf{F}_p[T]$-linear map preserves annihilator ideals, so if we let $A_0 \bmod \pi$ generate $C(\mathbf{F}_p[T]/\pi)$ then applying $L_{\pi,r}$ to it for $r=1,2,\ldots,k-1$ tells us $[\pi^{k-1}](A_0) \bmod \pi^k$ as an element of $C(\mathbf{F}_p[T]/\pi^k)$ has annihilator ideal $(\pi - 1)$. From $\pi \equiv 0 \bmod \pi$ we get $[\pi^{k-1}](\pi) \equiv 0 \bmod \pi^k$, and by injectivity of $L_{\pi,r}$'s we get $\pi \not\equiv 0 \bmod \pi^2 \Rightarrow [\pi^{k-2}](\pi) \not\equiv 0 \bmod \pi^k$, so $\pi \bmod \pi^k$ as an element of $C(\mathbf{F}_p[T]/\pi^k)$ has annihilator ideal (π^{k-1}). Let C_1 and C_2 be the submodules of $C(\mathbf{F}_p[T]/\pi^k)$ generated by $[\pi^{k-1}](A_0)$ and π, respectively. Since $C_1 \cap C_2 = \{0\}$, by counting we have $C(\mathbf{F}_p[T]/\pi^k) = C_1 + C_2 \cong C_1 \times C_2$. □

The last analogy we will develop between the structure of $(\mathbf{Z}/m)^\times$ and $C(\mathbf{F}_p[T]/M)$ is a Carlitz analogue of $\varphi(m) = |(\mathbf{Z}/m)^\times|$ that was shown to me by Darij Grinberg.

The φ-function admits two formulas (product over primes, sum over positive divisors):

$$\varphi(m) = m\prod_{p|m}\left(1-\frac{1}{p}\right) = \sum_{d|m} m\frac{\mu(d)}{d},$$

where μ is the Möbius function. For monic M in $\mathbf{F}_p[T]$, define $\varphi_C(M) \in \mathbf{F}_p[T]$ to be the polynomial (product over irreducible monic factors, sum over all monic factors)

$$\varphi_C(M) = M\prod_{\pi|M}\left(1-\frac{1}{\pi}\right) = \sum_{D|M} M\frac{\mu(D)}{D},$$

where $\mu(D) \in \{0,1,-1\}$ is defined in the same way as in the integers: $\mu(D)$ is $(-1)^r$ if D is squarefree with r monic irreducible factors, and $\mu(D)$ is 0 otherwise. For example, if π is monic irreducible then

$$\varphi_C(\pi^k) = \pi^k\left(1-\frac{1}{\pi}\right) = \pi^k - \pi^{k-1}.$$

Theorem 4.6. *The function φ_C has the following properties:*

(1) *For relatively prime monic A and B, $\varphi_C(AB) = \varphi_C(A)\varphi_C(B)$.*
(2) *For monic M, $\sum_{D|M}\varphi_C(D) = M$, where D runs over monic factors of M.*
(3) *For monic M and all A in $\mathbf{F}_p[T]$, $[\varphi_C(M)](A) \equiv 0 \bmod M$.*
(4) *For monic M and monic A in $\mathbf{F}_p[T]$, $[M](A)$ is monic and $M \mid \varphi_C([M](A))$ when $p \neq 2$, and also when $p=2$ and $\deg A \geq 2$.*

These four properties of $\varphi_C(M)$ are analogues of properties of $\varphi(m)$, namely

- for relatively prime positive integers a and b, $\varphi(ab) = \varphi(a)\varphi(b)$,
- for $m \geq 1$, $\sum_{d|m}\varphi(d) = m$, where d runs over positive factors of m,
- for $a \bmod m \in (\mathbf{Z}/m)^\times$, $a^{\varphi(m)} \equiv 1 \bmod m$,
- for $k \geq 1$ and $a > 1$, $k \mid \varphi(a^k - 1)$. (The order of $a \bmod a^k - 1$ is k.)

In the fourth part of the theorem, the constraints put on A are meant to avoid A being Carlitz torsion in $\mathbf{F}_2[T]$ (*e.g.*, $\Lambda_{T(T+1)} = \{0,1,T,T+1\}$ when $p=2$).[6] This is analogous to supposing $a > 1$ in the fourth property of $\varphi(m)$. As an example of what can go wrong

[6]The Carlitz torsion in $\mathbf{F}_p[T]$ when $p \neq 2$ is 0 by Theorem 2.18.

in the fourth part of the theorem when $p = 2$ and $\deg A = 1$, let $M = T^2 + 1$ and $A = T$. Then $[M](A) = T$, so $\varphi_C([M](A)) = \varphi_C(T) = T + 1$, which is not divisible by M.

Proof. (1) This is clear from the first formula defining φ_C.

(2) This follows from Möbius inversion for monic polynomials in $\mathbf{F}_p[T]$, using the second formula for $\varphi_C(M)$.

(3) If M_1 and M_2 are relatively prime and monic, then in $\mathbf{F}_p[T][X]$

$$[\varphi_C(M_1M_2)](X) = [\varphi_C(M_1)\varphi_C(M_2)](X) = [\varphi_C(M_1)]([\varphi_C(M_2)](X)),$$

so to prove the third property it suffices to check the case $M = \pi^k$ for monic irreducible π. In that case the congruence is $[\pi^k - \pi^{k-1}](A) \stackrel{?}{\equiv} 0 \bmod \pi^k$. We know by Theorem 2.15 that $[\pi](A) \equiv A \bmod \pi$, and then (4.3) turns this congruence into $[\pi^k](A) \equiv [\pi^{k-1}](A) \bmod \pi^k$ by induction on k, so π^k divides $[\pi^k - \pi^{k-1}](A) = [\varphi_C(\pi^k)](A)$.

(4) In $\mathbf{F}_p[T][X]$, the leading X-term in $[T^j](X)$ is X^{p^j}, so we may expect for monic $A \in \mathbf{F}_p[T]$ that the leading T- term of $[T^j](A) \in \mathbf{F}_p[T]$ is the same as the leading T-term of A^{p^j}. However, there is a subtlety: $[T^j](X)$ is a polynomial in T and X, which don't interact, so when we substitute $A \in \mathbf{F}_p[T]$ for X in $[T^j](X)$ to get a polynomial entirely in T we need to check that the terms in $[T^j](A)$ other than A^{p^j} all have smaller degree than A^{p^j}. Using the basic equations

$$[T^0](A) = A, \quad [T](A) = A^p + TA, \quad [T^{j+1}](A) = [T^j](A)^p + T([T^j](A))$$

check by induction that $[T^j](A)$ has the same leading T-term as A^{p^j} if $p \neq 2$ and $\deg A \geq 1$, or if $p = 2$ and $\deg A \geq 2$. Therefore if $M = T^m + \sum_{i=0}^{m-1} c_iT^i$ is monic, so

$$[M](A) = [T^m](A) + \sum_{i=0}^{m-1} c_i[T^i](A),$$

the leading term of $[M](A)$ is the same as the leading term of A^{p^m} if $p \neq 2$ and $\deg A \geq 1$, or if $p = 2$ and $\deg A \geq 2$. That implies $[M](A)$ is monic when M and A are monic under the conditions claimed in part (4) except for the case $p \neq 2$ and $\deg A = 0$, *i.e.*, $A = 1$, which we now check separately. By induction on i, $[T^i](1)$ has leading term $T^{p^{i-1}}$ if $p \neq 2$ and $i \geq 1$, so $[M](1)$ is monic if $p \neq 2$ and M is monic with $\deg M \geq 1$. If $\deg M = 0$ then $M = 1$ and $[M](1) = 1$ is also monic.

To show $M \mid \varphi_C([M](A))$, set $N = [M](A)$, which is monic. Using N in place of M in (3), $[\varphi_C(N)](A) \equiv 0 \bmod N$. Therefore in $C(\mathbf{F}_p[T]/N)$, the annihilator ideal of $A \bmod N$ contains $\varphi_C(N)$. Also $[M](A) \equiv 0 \bmod N$ by the definition of N, so M is in the annihilator ideal too. To show $M \mid \varphi_C(N)$, we will show M generates that annihilator ideal by showing no monic D dividing M with $\deg D < \deg M$ can satisfy $[D](A) \equiv 0 \bmod N$. We know $[D](A) \neq 0$ by Theorem 2.18 (since we are avoiding the case $p = 2$ and $\deg A \leq 1$). From $\deg D < \deg M$ we have $\deg([D](A)) < \deg([M](A))$ by the degree formulas in Theorem 2.18 (this is simple to check unless $p \neq 2$ and $\deg A = 0$, when the inequality is true but a little tedious to confirm), so $[D](A) \not\equiv 0 \bmod N$. □

One further analogy between $\varphi(m)$ and $\varphi_C(M)$ has to do with counting. For a finite abelian group, decomposed into a product of cyclic subgroups as $\mathbf{Z}/(n_1) \times \cdots \times \mathbf{Z}/(n_r)$, its cardinality is $|n_1 \cdots n_r|$. For a finitely generated torsion $\mathbf{F}_p[T]$-module $\mathcal{M}$, decomposed into a product of cyclic modules $\mathbf{F}_p[T]/(f_1) \times \cdots \times \mathbf{F}_p[T]/(f_r)$, the $\mathbf{F}_p[T]$-*cardinality* of $\mathcal{M}$ is defined to be the ideal $(f_1 \cdots f_r)$ and this ideal is independent of the cyclic decomposition.

For monic irreducible π and $k \geq 1$, the $\mathbf{F}_p[T]$-cardinality of $C(\mathbf{F}_p[T]/\pi^k)$ can be read off from the cyclic decomposition in Theorem 4.5, and it is $(\pi^{k-1}(\pi-1))$, whose monic generator is $\pi^{k-1}(\pi-1) = \pi^k - \pi^{k-1} = \varphi_C(\pi^k)$. From this and the Chinese remainder theorem it follows that the $\mathbf{F}_p[T]$-cardinality of $C(\mathbf{F}_p[T]/M)$ is $(\varphi_C(M))$ when M is monic, and this is a Carlitz analogue of the definition of $\varphi(m)$ as the cardinality of $(\mathbf{Z}/m)^\times$.

The groups $(\mathbf{F}_p[T]/\pi)^\times$ and the $\mathbf{F}_p[T]$-modules $C(\mathbf{F}_p[T]/\pi)$ present us with two ways to extend results about $(\mathbf{Z}/(p))^\times$ to the polynomial setting. For example, Artin's primitive root conjecture about integers generating infinitely many $(\mathbf{Z}/(p))^\times$ can be formulated for polynomials generating infinitely many groups $(\mathbf{F}_p[T]/\pi)^\times$ or generating infinitely many $\mathbf{F}_p[T]$-modules $C(\mathbf{F}_p[T]/\pi)$, with good answers in both cases; see [8] and [10, Chap. 10].

5. Carlitz extensions of $\mathbf{F}_p(T)$

We now adjoin Λ_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 [10, 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 Λ_M to F gives a Galois extension of F. We only need to adjoin a generator of Λ_M to F, since the other elements of Λ_M are polynomials in the generator (with $\mathbf{F}_p[T]$-coefficients). Each element of $\mathrm{Gal}(F(\Lambda_M)/F)$ permutes the roots Λ_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 mth roots of unity, we anticipate that each element of $\mathrm{Gal}(F(\Lambda_M)/F)$ acts on Λ_M by a Carlitz polynomial. To make this explicit, we use a generator of Λ_M.

Choose $\sigma \in \mathrm{Gal}(F(\Lambda_M)/F)$. Letting λ_0 be a generator of Λ_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 Λ_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 3.10). That σ acts like A on λ_0 propagates to all of Λ_M: every $\lambda \in \Lambda_M$ 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 σ has the same effect by the Carlitz action on all the elements of Λ_M. Write A as A_σ to indicate its dependence on σ: to each $\sigma \in \mathrm{Gal}(F(\Lambda_M)/F)$ we get a unit $A_\sigma \in (\mathbf{F}_p[T]/M)^\times$ that describes through its Carlitz polynomial how σ permutes the elements of Λ_M.

Theorem 5.1. *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 σ and τ in $\mathrm{Gal}(F(\Lambda_M)/F)$ and $\lambda \in \Lambda_M$, $(\sigma\tau)(\lambda)$ 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 Λ_M. Therefore $A_{\sigma\tau} \equiv A_\sigma A_\tau \bmod M$ (Theorem 3.10), which shows we have a homomorphism from $\mathrm{Gal}(F(\Lambda_M)/F)$ to $(\mathbf{F}_p[T]/M)^\times$.

When σ 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 σ is the identity on Λ_M, so σ 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 5.2. *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 [9, p. 278].

Both Λ_M and $(\mathbf{F}_p[T]/M)^\times$ are unchanged if we scale M by an element of $\mathbf{F}_p^\times$, so without loss of generality M is monic.

Pick a generator λ_0 of Λ_M. The image of $\mathrm{Gal}(F(\Lambda_M)/F) \to (\mathbf{F}_p[T]/M)^\times$ is all $A \bmod M$ such that $[A](\lambda_0)$ is F-conjugate to λ_0, so the map $\mathrm{Gal}(F(\Lambda_M)/F) \to (\mathbf{F}_p[T]/M)^\times$ is surjective when $[A](\lambda_0)$ is F-conjugate to λ_0 for all A that are relatively prime to M. Let λ_0 have minimal polynomial $f(X) \in F[X]$. The F-conjugates of λ_0 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 every monic irreducible $\pi \in \mathbf{F}_p[T]$ not dividing M.

Pick a monic irreducible π in $\mathbf{F}_p[T]$ that does not divide M, and let $g(X)$ be the minimal polynomial of $[\pi](\lambda_0)$ in $F[X]$. We want to show $g(X) = f(X)$. Since λ_0 and $[\pi](\lambda_0)$ are in Λ_M, both $f(X)$ and $g(X)$ divide $[M](X)$ in $F[X] = \mathbf{F}_p(T)[X]$. Since M is monic in $\mathbf{F}_p[T]$, $[M](X)$ is monic in X, and every monic factor of $[M](X)$ in $F[X]$ is in $\mathbf{F}_p[T][X]$. (This is analogous to every monic factor in $\mathbf{Q}[X]$ of a monic in $\mathbf{Z}[X]$ having to be in $\mathbf{Z}[X]$). Therefore $f(X)$ and $g(X)$ are in $\mathbf{F}_p[T][X]$.

Since $g([\pi](\lambda_0)) = 0$, $g([\pi](X))$ has λ_0 as a root, so $f(X) \mid g([\pi](X))$ in $F[X]$. Both $f(X)$ and $g([\pi](X))$ are monic X-polynomials in $\mathbf{F}_p[T][X]$ (because π 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 $u(X)$ and $v(X)$ are monic in $\mathbf{Z}[X]$ and $u(X) \mid v(X)$ in $\mathbf{Q}[X]$ then $u(X) \mid v(X)$ 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 π and use Theorem 2.17 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 π,

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

in $(\mathbf{F}_p[T]/\pi)[X]$. This is impossible: the right side has a multiple irreducible factor (each 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 2.11). So $g(X) = f(X)$, which shows $f([\pi](\lambda_0)) = 0$. ∎

Example 5.3. Let $M = T$. The isomorphism $\mathrm{Gal}(F(\Lambda_T)/F) \to (\mathbf{F}_p[T]/T)^\times$ sends each σ to 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 3.8) 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 5.2 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 all $\lambda \in \Lambda_T - \{0\}$, and $F(\Lambda_T)/F$ is a Kummer extension (Example 3.5), 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 of not only $\mathbf{F}_p(T)$ but every characteristic p field K not algebraic over $\mathbf{F}_p$: denote an element of K transcendental over $\mathbf{F}_p$ as T, so $K \supset \mathbf{F}_p(T)$. With this T we get 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 Galois with the effect of the Galois group on Λ_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).

6. More cyclotomic and Carlitz analogies

The roots of the polynomials $X^m - 1$ and $[M](X)$ have similar features (*e.g.*, the first is a cyclic group of size m and the second is a 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 $\mathrm{Gal}(F(\Lambda_M)/F) \cong (\mathbf{F}_p[T]/M)^\times$, that are are more profound. We explore analogies between these Galois extensions in this section.

By Theorem 5.2, $[F(\Lambda_M) : F] = |(\mathbf{F}_p[T]/M)^\times|$ for every $M \neq 0$, 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 $\varphi(M)$.[7] 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

$$\varphi(ab) = \frac{\varphi(a)\varphi(b)(a,b)}{\varphi((a,b))}, \quad \varphi(AB) = \frac{\varphi(A)\varphi(B)\,\mathrm{N}((A,B))}{\varphi((A,B))}. \tag{6.1}$$

Let's put the two formulas in (6.1) to work toward analogous goals. Classically, two cyclotomic fields $\mathbf{Q}(\mu_m)$ and $\mathbf{Q}(\mu_n)$ with $m \le n$ 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 result for cyclotomic extensions of $\mathbf{Q}$ and then just translate the argument over to Carlitz extensions of F.

Theorem 6.1. *Let m and n be positive integers.*

(1) *The number of roots of unity in $\mathbf{Q}(\mu_m)$ is $[2, m]$.*
(2) *We have $\mathbf{Q}(\mu_m) = \mathbf{Q}(\mu_n)$ if and only if $[2, m] = [2, n]$, which for $m \neq n$ is the same as saying $\min(m, n) = k$ and $\max(m, n) = 2k$ for some odd k.*

Proof. (1) Our argument is based on [1, p. 158]. The root of unity $-\zeta_m$ is in $\mathbf{Q}(\mu_m)$ and has order $2m$ if m is odd and m if m is even, hence order $[2, m]$ in general. Thus $\mu_{[2,m]} \subset \mathbf{Q}(\mu_m)$.

If $\mathbf{Q}(\mu_m)$ contains an rth root of unity then $\mathbf{Q}(\mu_r) \subset \mathbf{Q}(\mu_m)$, and taking degrees over $\mathbf{Q}$ shows $\varphi(r) \le \varphi(m)$. As $r \to \infty$, $\varphi(r) \to \infty$ (albeit erratically)[8] so there is a largest r

[7] Don't confuse $\varphi(M)$, a positive integer, with $\varphi_C(M)$ from Section 4, which is a polynomial.

[8] A bound $\varphi(r) \le B$ implies an upper bound on r. For each prime power p^e dividing r, $\varphi(p^e) \le B$, so $p^{e-1}(p - 1) \le B$. Then $2^{e-1} \le B$ and $p - 1 \le B$, so we get upper bounds on p and on e, which gives an upper bound on r by unique factorization.

satisfying $\mu_r \subset \mathbf{Q}(\mu_m)$. Since $\mu_m\mu_r = \mu_{[m,r]}$ is in $\mathbf{Q}(\mu_m)$ we have $[m,r] \leq r$, so $[m,r] = r$. Write $r = ms$. Then by (6.1),

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

Since $\mathbf{Q}(\mu_m) = \mathbf{Q}(\mu_r)$ for the maximal 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 number of roots of unity in $\mathbf{Q}(\mu_m)$ is either m or $2m$. If m is even then $\varphi(2m) = 2\varphi(m) > \varphi(m)$, so $r \neq 2m$. Thus when m is even the number of roots of unity in $\mathbf{Q}(\mu_m)$ is m. If m is odd then $-\zeta_m$ has order $2m$, so the number of roots of unity in $\mathbf{Q}(\mu_m)$ is $2m$. In general the number of roots of unity in $\mathbf{Q}(\mu_m)$ is $[2,m]$.

(2) If $\mathbf{Q}(\mu_m) = \mathbf{Q}(\mu_n)$ then counting roots of unity in this field implies $[2,m] = [2,n]$. Conversely, since $\mu_{[2,m]}$ is μ_m for even m and it is $\pm\mu_m$ for odd m, $\mathbf{Q}(\mu_m) = \mathbf{Q}(\mu_{[2,m]})$ for all m. Therefore if $[2,m] = [2,n]$ then $\mathbf{Q}(\mu_m) = \mathbf{Q}(\mu_n)$. For $m \neq n$, the condition $[2,m] = [2,n]$ becomes $m = [2,n]$ for even m (so n is odd and $m = 2n$), and $2m = [2,n]$ for odd m (so n is even and $n = 2m$). □

Theorem 6.2. *Let M and N be nonzero in $\mathbf{F}_p[T]$.*

1. *The full Carlitz torsion in $F(\Lambda_M)$ is Λ_M if $p \neq 2$ and $\Lambda_{[T(T+1),M]}$ if $p = 2$.*
2. (a) *When $p \neq 2$, $F(\Lambda_M) = F(\Lambda_N)$ if and only if $N = cM$ where $c \in \mathbf{F}_p^\times$.*
   (b) *When $p = 2$, $F(\Lambda_M) = F(\Lambda_N)$ if and only if*[9] $[M, T(T+1)] = [N, T(T+1)]$.

Proof. (1) For every monic R in $\mathbf{F}_p[T]$, 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 a monic R with $\Lambda_R \subset F(\Lambda_M)$ and $\mathrm{N}(R)$ as large as possible. Also $\Lambda_M + \Lambda_R = \Lambda_{[M,R]}$ is in $F(\Lambda_M)$, which implies $\Lambda_{[M,R]} = \Lambda_R$, so $[M,R] = R$. Write $R = MS$. Then by (6.1),

$$\varphi(R) = \varphi(MS) = \varphi(M)\varphi(S)\frac{\mathrm{N}((M,S))}{\varphi((M,S))} \geq \varphi(M)\varphi(S). \tag{6.2}$$

Since $F(\Lambda_M) = F(\Lambda_R)$ for the maximal R, computing degrees over F shows $\varphi(M) = \varphi(R) \geq \varphi(M)\varphi(S)$, so $1 \geq \varphi(S)$. Thus $R = SM$ with $\varphi(S) = 1$. If $(S,M) \neq 1$ then the calculation of $\varphi(MS)$ in (6.2) shows $\varphi(R) > \varphi(M)$ since $\mathrm{N}((M,S)) > \varphi((M,S))$. This contradicts the necessity of $\varphi(M) = \varphi(R)$, so $(S,M) = 1$.

In $\mathbf{F}_p[T]$, $\varphi(\pi) > 1$ for all irreducible π except when $p = 2$ and $\pi = T$ or $T+1$. Therefore when $p \neq 2$ the condition $\varphi(S) = 1$ for $S \in \mathbf{F}_p[T]$ implies $S \in \mathbf{F}_p^\times$, so $\Lambda_R = \Lambda_{SM} = \Lambda_M$. When $p = 2$, $\varphi(\pi) > 1$ if $\deg \pi \geq 2$ but $\varphi(T) = \varphi(T+1) = 1$. It is left to the reader to check $\varphi(S) = 1$ for $S \in \mathbf{F}_2[T]$ only when S is $1, T, T+1$, or $T(T+1)$ and conclude $\Lambda_R = \Lambda_{[T(T+1),M]}$.

(2) Verifying the description when $F(\Lambda_M) = F(\Lambda_N)$ is left to the reader. □

Remark 6.3. This theorem lets us find the M such that $\Lambda_M \subset \mathbf{F}_p(T)$, by solving $F(\Lambda_M) = F(\Lambda_1)$. Except when $p = 2$ and $M \mid T(T+1)$, necessarily $M \in \mathbf{F}_p^\times$, so $\Lambda_M = \{0\}$. If $p = 2$ and $M \mid T(T+1)$ then $\Lambda_M \subset \Lambda_{T(T+1)} = \{0, 1, T, T+1\}$. This explains when $\Lambda_M \subset \mathbf{F}_p(T)$ in a more conceptual way than Theorem 2.18.

[9] Least common multiples in $\mathbf{F}_q[T]$ are defined to be monic.

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 mth cyclotomic polynomial:

$$\Phi_m(X) = \prod_{\substack{1 \le a \le m \\ (a,m)=1}} (X - \zeta^a) = \prod_{\substack{\zeta^m=1 \\ \text{order } m}} (X - \zeta),$$

where ζ is a root of unity of order m in the first product and in the second product ζ runs over all roots of unity of order m. For example, if p is prime then $\Phi_p(X) = (X^p - 1)/(X - 1)$: every pth 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 Λ_M have the same minimal polynomial over $F = \mathbf{F}_p(T)$, which is an analogue of the cyclotomic polynomials:

$$\Phi_M(X) = \prod_{\substack{\deg A < \deg M \\ (A,M)=1}} (X - [A](\lambda_0)) = \prod_{\substack{[M](\lambda)=0 \\ \mathbf{F}_p[T]-\text{order } M}} (X - \lambda),$$

where λ_0 is a chosen generator of Λ_M and the second product is taken over roots λ of $[M](X)$ which have $\mathbf{F}_p[T]$-order M: $[D](\lambda) \neq 0$ for every monic proper divisor D of M. (Such λ are the *generators* of Λ_M, just as roots of unity of order m are the generators of μ_m.)

Example 6.4. If π is irreducible in $\mathbf{F}_p[T]$ then $\Phi_\pi(X) = [\pi](X)/X$ since $[\pi](X)/X$ is Eisenstein with respect to π (Corollary 2.12) and thus is irreducible over $\mathbf{F}_p(T) = F$. Comparing degrees, roots, and leading coefficients, for all $k \ge 1$ 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 π, and $[\pi^{k-1}](X)$ also has all non-leading X-coefficients divisible by π (Corollary 2.14), $\Phi_{\pi^k}(X)$ has all non-leading X-coefficients divisible by π. Therefore $\Phi_{\pi^k}(X)$ is Eisenstein with respect to π for all k.

Remark 6.5. It was noted in Remark 2.13 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 π 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 is $\Phi_M(0) = 1$.

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 [7]. 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 $d \ge 1$, $n \ge 1$, and $M \in \mathbf{F}_p[T]$, where Λ_{1/T^n} is the set of roots of the Carlitz polynomial $[1/T^n](X)$ built with $1/T$ in place of T: $[1/T](X) = X^p + (1/T)X$ and $[1/T^k](X) = [1/T]([1/T^{k-1}](X))$.[10]

Example 6.6. 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 that generate the same extension of

[10]The family of polynomials $[1/T^n](X)$ does not interact well with $[M](X)$ for $M \in \mathbf{F}_p[T]$, *e.g.*, $[1/T]([T](X)) \neq X$ and $[T]([1/T](X)) \neq X$.

$\mathbf{F}_p(T)$ 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 ∞) that is not satisfied by subfields of $\mathbf{F}_{p^d}(T, \Lambda_M)$, so it is not inside such a field.

Table 4 summarizes some of the analogous features we have seen with μ_m and Λ_M.

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

TABLE 4. Analogies between μ_m and Λ_M

7. Quadratic Reciprocity in $\mathbf{F}_p[T]$, $p \neq 2$

In this section, let p be odd. For every monic irreducible $\pi(T)$ in $\mathbf{F}_p[T]$ and every A in $\mathbf{F}_p[T]$, define the Legendre symbol $(\frac{A}{\pi})$ to be 0 or ± 1 in $\mathbf{F}_p$ according to

$$\left(\frac{A}{\pi}\right) = \begin{cases} 1, & \text{if } A \equiv \square \bmod \pi, A \not\equiv 0 \bmod \pi, \\ -1, & \text{if } A \not\equiv \square \bmod \pi, \\ 0, & \text{if } A \equiv 0 \bmod \pi. \end{cases}$$

Our goal is to prove a reciprocity law for this symbol by adapting a proof of quadratic reciprocity in $\mathbf{Z}$, using $\mathbf{F}_p(T, \Lambda_\pi)/\mathbf{F}_p(T)$ in place of $\mathbf{Q}(\zeta_p)/\mathbf{Q}$. We assume the reader knows quadratic reciprocity in $\mathbf{Z}$, as otherwise the entire point of studying $(\frac{A}{\pi})$ will be lost.

Here are three basic properties of the Legendre symbol on $\mathbf{F}_p[T]$:

- If $A \equiv B \bmod \pi$ then $(\frac{A}{\pi}) = (\frac{B}{\pi})$ since $(\frac{A}{\pi})$ only depends on $A \bmod \pi$.
- (Euler's congruence): for all $A \in \mathbf{F}_p[T]$,

$$\left(\frac{A}{\pi}\right) \equiv A^{(\mathrm{N}(\pi)-1)/2} \bmod \pi. \tag{7.1}$$

- (Multiplicativity): for all A and B in $\mathbf{F}_p[T]$,

$$\left(\frac{AB}{\pi}\right) = \left(\frac{A}{\pi}\right)\left(\frac{B}{\pi}\right).$$

To prove (7.1), look at a proof of Euler's congruence in $\mathbf{Z}$, $(\frac{a}{p}) \equiv a^{(p-1)/2} \bmod p$ and it should carry over to $\mathbf{F}_p[T]$. That $(\frac{AB}{\pi}) = (\frac{A}{\pi})(\frac{B}{\pi})$ for all A and B follows from (7.1) in the same way that Euler's congruence in $\mathbf{Z}$ implies $(\frac{ab}{p}) = (\frac{a}{p})(\frac{b}{p})$ for all integers a and b.

By multiplicativity, a calculation of $(\frac{A}{\pi})$ for general A is reduced to two cases: $(\frac{c}{\pi})$ for $c \in \mathbf{F}_p^\times$ and $(\frac{\widetilde{\pi}}{\pi})$ for monic irreducible $\widetilde{\pi}$ not equal to π. The first case is analogous to the supplementary law for $(\frac{-1}{p})$, and we deal with it first.

Theorem 7.1. *For $c \in \mathbf{F}_p^\times$ and π a monic irreducible in $\mathbf{F}_p[T]$, $(\frac{c}{\pi}) = c^{(\mathrm{N}(\pi)-1)/2} = (\frac{c}{p})^{\deg \pi}$.*

Proof. By (7.1), $(\frac{c}{\pi}) \equiv c^{(\mathrm{N}(\pi)-1)/2} \bmod \pi$. Both sides are in $\mathbf{F}_p^\times$, and different elements of $\mathbf{F}_p$ can't be congruent mod π, so $(\frac{c}{\pi}) = c^{(\mathrm{N}(\pi)-1)/2}$.

Let $d = \deg \pi$, so $\frac{\mathrm{N}(\pi)-1}{2} = \frac{p-1}{2}(1 + p + \cdots + p^{d-1})$ and

$$c^{(\mathrm{N}(\pi)-1)/2} = c^{\frac{p-1}{2}(1+p+\cdots+p^{d-1})}.$$

Since this equation is in $\mathbf{F}_p$, we can replace $c^{(p-1)/2}$ by $(\frac{c}{p})$, so the exponent $1+p+\cdots+p^{d-1}$ only matters mod 2. This sum is d mod 2, so $c^{(\mathrm{N}(\pi)-1)/2} = (\frac{c}{p})^d$. $\square$

We turn now to the $\mathbf{F}_p[T]$-analogue of the main law of quadratic reciprocity.

Theorem 7.2. *For distinct monic irreducible π and $\widetilde{\pi}$ in $\mathbf{F}_p[T]$,*

$$\left(\frac{\widetilde{\pi}}{\pi}\right) = (-1)^{(\mathrm{N}\pi-1)/2\cdot(\mathrm{N}\widetilde{\pi}-1)/2}\left(\frac{\pi}{\widetilde{\pi}}\right). \tag{7.2}$$

The exponent $\frac{\mathrm{N}(\pi)-1}{2} \cdot \frac{\mathrm{N}(\widetilde{\pi})-1}{2}$ only matters mod 2, and by calculations from the proof of Theorem 7.1 this product is congruent to $\frac{p-1}{2}\deg\pi \cdot \frac{p-1}{2}\deg\widetilde{\pi} \equiv \frac{p-1}{2}\deg\pi\deg\widetilde{\pi}$ mod 2, so another way of writing (7.2) is

$$\left(\frac{\widetilde{\pi}}{\pi}\right) = (-1)^{\frac{p-1}{2}\deg\pi\deg\widetilde{\pi}}\left(\frac{\pi}{\widetilde{\pi}}\right).$$

Therefore if π or $\widetilde{\pi}$ has even degree then $\left(\frac{\widetilde{\pi}}{\pi}\right) = \left(\frac{\pi}{\widetilde{\pi}}\right)$.

Example 7.3. In $\mathbf{F}_3[T]$, is $T^3 - T - 1 \equiv \square \bmod T^4 + T + 2$? The polynomials $T^3 - T - 1$ and $T^4 + T + 2$ are both irreducible, so (7.2) says

$$\left(\frac{T^3 - T - 1}{T^4 + T + 2}\right) = \left(\frac{T^4 + T + 2}{T^3 - T - 1}\right).$$

Since $T^4 + T + 2 \equiv T^2 + 2T + 2 \bmod T^3 - T - 1$, and $T^2 + 2T + 2$ is irreducible in $\mathbf{F}_3[T]$,

$$\left(\frac{T^4 + T + 2}{T^3 - T - 1}\right) = \left(\frac{T^2 + 2T + 2}{T^3 - T - 1}\right) \overset{(7.2)}{=} \left(\frac{T^3 - T - 1}{T^2 + 2T + 2}\right).$$

We have $T^3 - T - 1 \equiv T \bmod T^2 + 2T + 2$, so

$$\left(\frac{T^3 - T - 1}{T^2 + 2T + 2}\right) = \left(\frac{T}{T^2 + 2T + 2}\right) \overset{(7.2)}{=} \left(\frac{T^2 + 2T + 2}{T}\right) = \left(\frac{2}{T}\right).$$

By Theorem 7.1, $(\frac{2}{T}) = (\frac{2}{3})^{\deg T} = -1,$[11] so $(\frac{T^3-T-1}{T^4+T+2}) = -1$ in $\mathbf{F}_3[T]$.

[11] Concretely, the value of $(\frac{2}{T})$ is related to whether or not 2 is a square in $\mathbf{F}_3[T]/(T) \cong \mathbf{F}_3$. Since 2 is not a square in $\mathbf{F}_3$, $(\frac{2}{T}) = -1$.

There are proofs of Theorem 7.2 that take advantage of features of $\mathbf{F}_p[T]$ that are unavailable in $\mathbf{Z}$. See, for instance, [5]. We will instead prove Theorem 7.2 by using the analogy between cyclotomic extensions of $\mathbf{Q}$ and Carlitz extensions of $\mathbf{F}_p(T)$.

The extension $\mathbf{Q}(\zeta_p)/\mathbf{Q}$ has cyclic Galois group $(\mathbf{Z}/(p))^\times$, of even order, so there is a unique quadratic extension of $\mathbf{Q}$ in $\mathbf{Q}(\zeta_p)$. It is $\mathbf{Q}(\sqrt{p^*})$, where $p^* = (-1)^{(p-1)/2}p$ (that is, $p^* = \pm p$, with sign chosen so that $p^* \equiv 1 \bmod 4$). Here are a few ways to show $\sqrt{p^*} \in \mathbf{Q}(\zeta_p)$:

(1) Gauss sums. Define $G_p = \sum_{a \bmod p} (\frac{a}{p})\zeta_p^a$. By construction this lies in $\mathbf{Q}(\zeta_p)$, and many textbook treatments of Gauss sums will provide a proof that $G_p^2 = p^*$.

(2) Ramification. The extension $\mathbf{Q}(\zeta_p)/\mathbf{Q}$ is ramified only at the prime p (among finite primes), so a quadratic field in $\mathbf{Q}(\zeta_p)$ can ramify only at p. The unique quadratic field ramified only at the *odd* prime p (among finite primes) is $\mathbf{Q}(\sqrt{p^*})$.

(3) Rewriting terms in a product. In $X^{p-1} + \cdots + X + 1 = \prod_{k=1}^{p-1}(X - \zeta_p^k)$ set $X = 1$ to get[12] $p = \prod_{k=1}^{p-1}(1-\zeta_p^k)$. The product of the terms at k and $p-k$ is $(1-\zeta_p^k)(1-\zeta_p^{-k}) = -(\zeta_p^k - 2 + \zeta_p^{-k}) = -(\zeta_p^{k/2} - \zeta_p^{-k/2})^2$, where $\pm k/2$ are interpreted in $\mathbf{Z}/(p)$ (or write $\zeta_p^{k/2}$ as $\zeta_p^{k(p+1)/2}$). This is -1 times a square, so $(-1)^{(p-1)/2}p$ is a square in $\mathbf{Q}(\zeta_p)$.

Let $F = \mathbf{F}_p(T)$. The extension $F(\Lambda_\pi)/F$ has a Galois group $(\mathbf{F}_p[T]/\pi)^\times$ that is cyclic of even order $\mathrm{N}(\pi) - 1$, so there is a unique quadratic extension of F inside $F(\Lambda_\pi)$.

Lemma 7.4. *The quadratic extension of F in $F(\Lambda_\pi)$ is $F(\sqrt{\pi^*})$, where $\pi^* = (-1)^{(\mathrm{N}\pi-1)/2}\pi$.*

Proof. The first two methods of showing $\sqrt{p^*} \in \mathbf{Q}(\zeta_p)$ don't carry over to the Carlitz setting:

- The sum $\sum_{A \bmod \pi}(\frac{A}{\pi})[A](\lambda)$ for a fixed nonzero λ in Λ_π resembles a Gauss sum, but it is useless because it is 0 (surprise!) if $\mathrm{N}(\pi) > 3$ (not if $\mathrm{N}(\pi) = 3$).
- A quadratic extension of F is not uniquely determined if it has one ramified prime π, as illustrated by $F(\sqrt{\pi})$ and $F(\sqrt{c\pi})$ for nonsquare c in $\mathbf{F}_p^\times$.

The third method does adapt. For λ in $\Lambda_\pi - \{0\}$, in $[\pi](X)/X = \prod_{A \not\equiv 0 \bmod \pi}(X - [A](\lambda))$ set $X = 0$ and get $\pi = \prod_{A \not\equiv 0 \bmod \pi}[A](\lambda)$.[13] The product of the terms at A and $-A$ is $[A](\lambda)[-A](\lambda) = -[A](\lambda)^2$ because $[-A](X) = [-1]([A](X)) = -[A](X)$. Therefore up to a square factor in $F(\Lambda_\pi)^\times$, π equals $(-1)^{(\mathrm{N}\pi-1)/2}$, so $(-1)^{(\mathrm{N}\pi-1)/2}\pi$ is a square in $F(\Lambda_\pi)$. $\square$

The proof of quadratic reciprocity in $\mathbf{Z}$ that we will model our proof on in $\mathbf{F}_p[T]$ is the one using Gauss sums. Let's review it. For an odd prime p, the Gauss sum $G_p = \sum_{a \bmod p}(\frac{a}{p})\zeta_p^a$ satisfies $G_p^2 = p^*$ and $G_p \in \mathbf{Z}[\zeta_p]$. For an odd prime $q \neq p$ the quotient ring[14] $\mathbf{Z}[\zeta_p]/(q)$ has prime characteristic q. We will compute $G_p^q \bmod q\mathbf{Z}[\zeta_p]$ in two ways.

(1) Since the q-th power map is additive in characteristic q and $\varepsilon^q = \varepsilon$ for $\varepsilon = \pm 1$, we have $G_p^q \equiv \sum_{a \bmod p}(\frac{a}{p})\zeta_p^{aq} \bmod q\mathbf{Z}[\zeta_p]$. Changing variables by $a \mapsto a/q$ on the summation indices in $\mathbf{Z}/(p)$, $G_p^q \equiv \sum_{a \bmod p}(\frac{aq^{-1}}{p})\zeta_p^a \equiv (\frac{q}{p})G_p \bmod q\mathbf{Z}[\zeta_p]$.

(2) From $G_p^2 = p^*$, $G_p^q = (G_p^2)^{(q-1)/2}G_p = (p^*)^{(q-1)/2}G_p \equiv (\frac{p^*}{q})G_p \bmod q\mathbf{Z}[\zeta_p]$.

From the two calculations of G_p^q in $\mathbf{Z}[\zeta_p]/(q)$, $\left(\frac{q}{p}\right)G_p \equiv \left(\frac{p^*}{q}\right)G_p \bmod q\mathbf{Z}[\zeta_p]$. Multiplying both sides by G_p, $\left(\frac{q}{p}\right)p^* \equiv \left(\frac{p^*}{q}\right)p^* \bmod q\mathbf{Z}[\zeta_p]$. Since $p^* = \pm p$ is invertible in $\mathbf{Z}/(q)$,

---

[12] In terms of the field norm, this says $p = \mathrm{N}_{\mathbf{Q}(\zeta_p)/\mathbf{Q}}(1 - \zeta_p)$.

[13] This says $\pi = \mathrm{N}_{F(\Lambda_\pi)/F}(\lambda)$.

[14] This quotient ring is usually not a field.

and thus in $\mathbf{Z}[\zeta_p]/(q)$, $\left(\frac{q}{p}\right) \equiv \left(\frac{p^*}{q}\right) \bmod q\mathbf{Z}[\zeta_p]$. Both sides of this congruence are ± 1, and $1 \not\equiv -1 \bmod q\mathbf{Z}[\zeta_p]$ (otherwise $2/q$ would be an algebraic integer, which is false), so in $\mathbf{Z}$

$$\left(\frac{q}{p}\right) = \left(\frac{p^*}{q}\right) = \left(\frac{(-1)^{(p-1)/2}p}{q}\right) = \left(\frac{-1}{q}\right)^{(p-1)/2}\left(\frac{p}{q}\right) = (-1)^{(p-1)/2\cdot(q-1)/2}\left(\frac{p}{q}\right).$$

In the Carlitz setting, it's attractive to adapt this proof by working with the sum $G_\pi := \sum_{A \bmod \pi}(\frac{A}{\pi})[A](\lambda)$, where λ is a fixed nonzero element of Λ_π, but this has a fatal flaw that we already indicated in the proof of Lemma 7.4: $G_\pi = 0$ (unless $\mathrm{N}(\pi) = 3$). Let's stop for a moment and see where this vanishing comes from. The argument we will use was shown to me by Darij Grinberg.

By the property $[M+N](X) = [M](X)+[N](X)$, we can write G_π as $\left[\sum_{A \bmod \pi}(\frac{A}{\pi})A\right](\lambda)$. Since $[\pi](\lambda) = 0$, each $[M](\lambda)$ only depends on $M \bmod \pi$. We will show $\sum_{A \bmod \pi}(\frac{A}{\pi})A \equiv 0 \bmod \pi$, so $G_\pi = [0](\lambda) = 0$. For $B \not\equiv 0 \bmod \pi$,

$$\sum_{A \bmod \pi}\left(\frac{A}{\pi}\right)A \equiv \sum_{A \bmod \pi}\left(\frac{AB}{\pi}\right)AB \equiv \left(\frac{B}{\pi}\right)B\sum_{A \bmod \pi}\left(\frac{A}{\pi}\right)A \bmod \pi.$$

When $\mathrm{N}(\pi) > 3$, there is a $B \not\equiv 0 \bmod \pi$ such that $B \not\equiv \pm 1 \bmod \pi$, so $(\frac{B}{\pi})B \not\equiv 1 \bmod \pi$. Using that B above, we conclude that $\sum_{A \bmod \pi}(\frac{A}{\pi})A \equiv 0 \bmod \pi$.

We've met a peculiar non-analogy between $\mathbf{Z}$ and $\mathbf{F}_p[T]$: the classical Gauss sum G_p squares to p^* but the sum G_π vanishes (except in one case). To prove quadratic reciprocity in $\mathbf{F}_p[T]$ by the Gauss sum method without using a Gauss sum, we will examine a *product* (not a sum) that squares to π^*.

Let H be a set of representatives of $(\mathbf{F}_p[T]/(\pi))^\times/\{\pm 1\}$ (a "half-system" mod π). By the proof of Lemma 7.4, $\pi = \prod_{A\in H}[A](\lambda)[-A](\lambda) = (-1)^{(\mathrm{N}(\pi)-1)/2}\prod_{A\in H}[A](\lambda)^2$, where λ is a fixed nonzero element of Λ_π. The product $\Gamma_\pi = \prod_{A\in H}[A](\lambda)$ over the half-system H is going to be our replacement for the Gauss sum:

$$\Gamma_\pi \in \mathbf{F}_p[T,\lambda] \quad \text{and} \quad \Gamma_\pi^2 = (-1)^{(\mathrm{N}(\pi)-1)/2}\pi = \pi^*.$$

For monic irreducible $\widetilde{\pi}$ distinct from π, we will compute $\Gamma_\pi^{\mathrm{N}(\widetilde{\pi})}$ in $\mathbf{F}_p[T,\lambda]/(\widetilde{\pi})$ in two ways. (This is the analogue of computing G_p^q in $\mathbf{Z}[\zeta_p]/(q)$ in two ways.)

(1) From the definition of Γ_π, $\Gamma_\pi^{\mathrm{N}(\widetilde{\pi})} = \prod_{A\in H}[A](\lambda)^{\mathrm{N}(\widetilde{\pi})}$. By Theorem 2.17, $f(X)^{\mathrm{N}(\widetilde{\pi})} \equiv f([\widetilde{\pi}](X)) \bmod \widetilde{\pi}\mathbf{F}_p[T,X]$ for all $f(X) \in \mathbf{F}_p[T][X]$. Using $[A](X)$ as $f(X)$ and setting $X = \lambda$,

$$\begin{aligned}\prod_{A\in H}[A](\lambda)^{\mathrm{N}(\widetilde{\pi})} &\equiv \prod_{A\in H}[A]([\widetilde{\pi}](\lambda)) \bmod \widetilde{\pi}\mathbf{F}_p[T,\lambda]\\ &\equiv \prod_{A\in H}[A\widetilde{\pi}](\lambda) \bmod \widetilde{\pi}\mathbf{F}_p[T,\lambda].\end{aligned}$$

The value of $[A\widetilde{\pi}](\lambda)$ only depends on $A\widetilde{\pi}$ modulo π. (This is an analogue of ζ_p^{aq} only depending on $aq \bmod p$.) As A runs over the half-system H, the products $A\widetilde{\pi}$ are also a half-system. It may not be H itself, but we can match terms with H up to sign: $A\widetilde{\pi} \equiv \varepsilon_A B_A \bmod \pi$ for $\varepsilon_A = \pm 1$ and $B_A \in H$. Since $[\varepsilon_A B_A](X) = \varepsilon_A[B_A](X)$,

$$\prod_{A\in H}[A\widetilde{\pi}](\lambda) = \prod_{A\in H}[\varepsilon_A B_A](\lambda) = \prod_{A\in H}\varepsilon_A\cdot\prod_{A\in H}[B_A](\lambda) = \prod_{A\in H}\varepsilon_A\cdot\prod_{B\in H}[B](\lambda) = \left(\prod_{A\in H}\varepsilon_A\right)\Gamma_\pi.$$

The product $\prod_{A\in H}\varepsilon_A$ equals $(\frac{\widetilde{\pi}}{\pi})$; this is an analogue of the classical Gauss lemma that expresses a Legendre symbol as a product of signs coming from scaling all the terms in a half-system for $(\mathbf{Z}/(p))^\times/\{\pm 1\}$ (the proof is the same, and left to the reader). Therefore

$$\Gamma_\pi^{\mathrm{N}(\widetilde{\pi})} = \prod_{A\in H}[A](\lambda)^{\mathrm{N}(\widetilde{\pi})} \equiv \left(\frac{\widetilde{\pi}}{\pi}\right)\Gamma_\pi \bmod \widetilde{\pi}\mathbf{F}_p[T,\lambda]. \tag{7.3}$$

(2) Since $\Gamma_\pi^2 = \pi^*$,

$$\Gamma_\pi^{\mathrm{N}(\widetilde{\pi})} = (\Gamma_\pi^2)^{(\mathrm{N}(\widetilde{\pi})-1)/2}\Gamma_\pi = (\pi^*)^{(\mathrm{N}(\widetilde{\pi})-1)/2}\Gamma_\pi \equiv \left(\frac{\pi^*}{\widetilde{\pi}}\right)\Gamma_\pi \bmod \widetilde{\pi}\mathbf{F}_p[T,\lambda]. \tag{7.4}$$

Comparing (7.3) and (7.4), $\left(\frac{\widetilde{\pi}}{\pi}\right)\Gamma_\pi \equiv \left(\frac{\pi^*}{\widetilde{\pi}}\right)\Gamma_\pi \bmod \widetilde{\pi}\mathbf{F}_p[T,\lambda]$. Multiply both sides of this congruence by Γ_π and cancel the resulting factor π^* on both sides (since $\pi^* = \pm\pi$ is invertible in $\mathbf{F}_p[T]/(\widetilde{\pi})$) to obtain $\left(\frac{\widetilde{\pi}}{\pi}\right) \equiv \left(\frac{\pi^*}{\widetilde{\pi}}\right) \bmod \widetilde{\pi}\mathbf{F}_p[T,\lambda]$. (This is the analogue of $\left(\frac{q}{p}\right) \equiv \left(\frac{p^*}{q}\right) \bmod q\mathbf{Z}[\zeta_p]$.) We have $1 \not\equiv -1 \bmod \widetilde{\pi}\mathbf{F}_p[T,\lambda]$ (otherwise, since λ is integral over $\mathbf{F}_p[T]$, $2/\widetilde{\pi}$ would be integral over $\mathbf{F}_p[T]$, which is false), so in $\mathbf{F}_p^\times$

$$\left(\frac{\widetilde{\pi}}{\pi}\right) = \left(\frac{\pi^*}{\widetilde{\pi}}\right).$$

Theorem 7.2 follows from this by computing the right side using the definition of π^* and the supplementary law Theorem 7.1 when $c = -1$.

Remark 7.5. Our first method of calculating $\Gamma_\pi^{\mathrm{N}(\widetilde{\pi})}$ in $\mathbf{F}_p[T,\lambda]/(\widetilde{\pi})$, leading to (7.3), can bypass the need for Gauss's lemma by using Galois theory. For $A \in (\mathbf{F}_p[T]/(\pi))^\times$, let $\sigma_A \in \mathrm{Gal}(F(\Lambda_\pi)/F)$ be the automorphism such that $\sigma_A(\lambda) = [A](\lambda)$ for all $\lambda \in \Lambda_\pi$. Since $\Gamma_\pi^2 \in \mathbf{F}_p[T]$, $\sigma_A(\Gamma_\pi) = \pm\Gamma_\pi$. Write $\sigma_A(\Gamma_\pi) = \delta_A\Gamma_\pi$, where $\delta_A = \pm 1$. Then $\delta_{AB} = \delta_A\delta_B$, so

$$A \bmod \pi \mapsto \delta_A$$

is a homomorphism $(\mathbf{F}_p[T]/(\pi))^\times \to \{\pm 1\}$, and it is surjective since $-\Gamma_\pi$ is an $\mathbf{F}_p(T)$-conjugate of Γ_π (that is, $-\Gamma_\pi = \sigma(\Gamma_\pi)$ for some $\sigma \in \mathrm{Gal}(F(\Lambda_\pi)/F) \cong (\mathbf{F}_p[T]/(\pi))^\times$). Since $(\mathbf{F}_p[T]/(\pi))^\times$ is cyclic, it has only one nontrivial homomorphism onto $\{\pm 1\}$, and the homomorphism $A \bmod \pi \mapsto (\frac{A}{\pi})$ is an example, so $\delta_A = (\frac{A}{\pi})$ for all A. That is, $\sigma_A(\Gamma_\pi) = (\frac{A}{\pi})\Gamma_\pi$ for all $A \not\equiv 0 \bmod \pi$.

What does this tell us when $A = \widetilde{\pi}$? Since

$$\sigma_{\widetilde{\pi}}(\lambda) = [\widetilde{\pi}](\lambda) \equiv \lambda^{\mathrm{N}(\widetilde{\pi})} \bmod \widetilde{\pi}\mathbf{F}_p[T,\lambda]$$

by Theorem 2.17, and the maps $\sigma_{\widetilde{\pi}}$ and raising to the $\mathrm{N}(\widetilde{\pi})$-th power are $\mathbf{F}_p[T]$-algebra endomorphisms on $\mathbf{F}_p[T,\lambda]/(\widetilde{\pi})$, $\sigma_{\widetilde{\pi}}(f(\lambda)) \equiv f(\lambda)^{\mathrm{N}(\widetilde{\pi})} \bmod \widetilde{\pi}\mathbf{F}_p[T,\lambda]$ for all $f(X) \in \mathbf{F}_p[T,X]$. Thus $\sigma_{\widetilde{\pi}}(\Gamma_\pi) \equiv \Gamma_\pi^{\mathrm{N}(\widetilde{\pi})} \bmod \widetilde{\pi}\mathbf{F}_p[T,\lambda]$, so (7.3) follows from the general formula $\sigma_A(\Gamma_\pi) = (\frac{A}{\pi})\Gamma_\pi$ at $A = \widetilde{\pi}$.

This point of view can be extended to prove quadratic reciprocity in $\mathbf{F}_p[T]$ by using the method of Frobenius elements, similar to one of the other proofs of quadratic reciprocity in $\mathbf{Z}$ [11, Sect. 6.5].

8. 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 π (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 ξ 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

$$f(X) := X \prod_{\substack{A\in\mathbf{F}_p[T]\\ A\neq 0}}\left(1 - \frac{X}{\xi A}\right), \tag{8.1}$$

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.

By viewing $f(X)$ in (8.1) as a limit of the finite products $f_d(X) = X\prod_{\deg A\leq d}(1 - X/\xi A)$ as $d \to \infty$, we have $f(X+Y) = f(X) + f(Y)$ and $f(cX) = cf(X)$ for all $c \in \mathbf{F}_p$.[15] The roots of $f(TX)$ are

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

and all roots are simple. Because $f(X)$ is additive and vanishes on $\xi\mathbf{F}_p[T]$, on each 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 ξ) for the two functions to match:

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

[15] The coefficient of X^j in $f(X)$ is the limit of the jth 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 form an $\mathbf{F}_p$-vector space. Every such polynomial is a p-polynomial by Appendix A, so $f_d(X+Y) = f_d(X) + f_d(Y)$ and $f_d(cX) = cf_d(X)$ for $c \in \mathbf{F}_p$. Now let $d \to \infty$.

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

$$f(TX) = f(X)^p + Tf(X). \tag{8.2}$$

The condition $T = -f(\xi/T)^{p-1}$ nearly determines ξ. By the product defining $f(X)$ in (8.1), 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 $\mathbf{F}_p((1/T))$ using the $(1/T)$-adic absolute value, where for a nonzero Laurent series $c_d/T^d + c_{d+1}/T^{d+1} + \cdots$ with $c_d \in \mathbf{F}_p^\times$ we set

$$\left|\frac{c_d}{T^d} + \frac{c_{d+1}}{T^{d+1}} + \cdots\right| = \left(\frac{1}{p}\right)^d.$$

For example, $|1/T^n| = 1/p^n$ and $|A| = p^{\deg A}$ for all nonzero $A \in \mathbf{F}_p[T]$. We have $|1/(TA)| = (1/p)^{1+\deg A}$ for every 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

$$\xi^{p-1} = \frac{-T^p}{\prod_{A\neq 0}(1 - 1/TA)^{p-1}}. \tag{8.3}$$

This product converges in $\mathbf{F}_p((1/T))$ and ξ is algebraic over $\mathbf{F}_p((1/T))$. We can use this equation to define ξ (at last). Since ξ appears in (8.3) through ξ^{p-1}, the equation only determines ξ 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 ξ (so that (8.2) is satisfied). The product defining $f(X)$ has first term X, so we need $a_0 = 1$. Substituting the series for $f(X)$ into (8.2) 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 $j \geq 1$. By induction on j, $\deg(D_j) = jp^j$ for all $j \geq 0$.

Definition 8.1. 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 8.2. It can be shown for all $j \geq 0$ that

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

Theorem 8.3 (Carlitz). *There is an infinite product decomposition*

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

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

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

Proof. See [6, Cor. 3.2.9]. □

Remark 8.4. Carlitz gave another expression for ξ_p:

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

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 $e_C(X)$ that is as important for it as the differential equation is for e^X is not $e'_C(X) = 1$ but rather (8.2): $e_C(TX) = e_C(X)^p + Te_C(X)$.

The parameter ξ_p doesn't appear in the coefficients of the Carlitz exponential series in Definition 8.1, just like π doesn't appear in the definition of the usual exponential series. The value ξ_p is a characteristic p analogue of $2\pi i$. Wade [12] proved ξ_p 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 all $x \in \mathbf{F}_p((1/T))$, the $1/T$-adic absolute value of the general term in the series $e_C(x)$ is $|x|^{p^j}/|D_j| = |x|^{p^j}/(1/p)^{-jp^j} = (|x|/p^j)^{p^j}$ because $\deg D_j = jp^j$. This tends to 0 as $j \to \infty$ for each choice of $|x|$, so the series $e_C(x)$ converges for all x. Taking $x = 1$, for instance, we get the $1/T$-adic power series

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

We need to enlarge the domain of $e_C(X)$ beyond $\mathbf{F}_p((1/T))$ to find its full kernel $\xi_p\mathbf{F}_p[T]$ unless $p = 2$ (because $\xi_p \notin \mathbf{F}_p((1/T))$ unless $p = 2$). When $p \neq 2$, 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 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 a p-power series in X, for all $M \in \mathbf{F}_p[T]$ it follows that

$$e_C(MX) = [M](e_C(X)). \tag{8.4}$$

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 (8.4) 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 (8.4) analytic content is the next result, which is the analogue for Carlitz torsion of the complex-analytic parametrization of mth roots of unity: $\mu_m = \{e^{2\pi i a/m} : a \in \mathbf{Z}\}$.

Theorem 8.5. *For nonzero M in $\mathbf{F}_p[T]$, $\Lambda_M = \{e_C((A/M)\xi_p) : A \in \mathbf{F}_p[T]\}$.*

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

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

so $e_C((A/M)\xi_p) \in \Lambda_M$. To show these Carlitz exponential values fill up Λ_M, we count the values. If $e_C((A/M)\xi_p) = e_C((B/M)\xi_p)$ then subtracting shows $e_C(((A-B)/M)\xi_p) = 0$, so $(A-B)/M \in \mathbf{F}_p[T]$ by Theorem 8.3. Thus $A \equiv B \bmod M$, so the number of values for $e_C((A/M)\xi_p)$ as A varies is $|\mathbf{F}_p[T]/M| = |\Lambda_M|$. □

Remark 8.6. The proof that $e_C(x) = \sum_{j\geq 0} x^{p^j}/D_j$ converges for all $x \in \mathbf{F}_p((1/T))$ carries over to convergence of $e_C(x)$ for all x in any complete valued extension field of $\mathbf{F}_p((1/T))$, so *a priori* the formal series $e_C((A/M)\xi_p)$ make sense in $\mathbf{F}_p((1/T))((-T)^{1/(p-1)})$, a finite extension of the complete field $\mathbf{F}_p((1/T))$; every finite extension of a complete valued field is complete with respect to a unique extension of the absolute value on the smaller field.

Corollary 8.7. *As $\mathbf{F}_p[T]$-modules, $\bigcup_M \Lambda_M \cong \mathbf{F}_p(T)/\mathbf{F}_p[T]$.*

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

Corollary 8.7 is analogous to the isomorphism of $\mathbf{Q}/\mathbf{Z}$ with all roots of unity in $\overline{\mathbf{Q}}^\times$ by $r \mapsto e^{2\pi i r}$.

The Carlitz exponential helps us describe the coefficients of $[M](X)$ when $M \neq 0$. 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

$$[m](X) = (1+X)^m - 1 = \sum_{j=1}^m c_{j,m} X^j, \tag{8.5}$$

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

$$[m](X) = e^{m\log(1+X)} - 1 = \sum_{j\geq 1} \frac{m^j(\log(1+X))^j}{j!}. \tag{8.6}$$

The right side is in $\mathbf{Q}[[X]]$ and makes sense since $(\log(1+X))^j = X^j + \text{higher order terms}$. 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 $P_j(Y) \in \mathbf{Q}[Y]$. Because $(\log(1+X))^j = X^j + \text{higher order terms}$, $\deg P_j(Y) = j$. 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 (8.5), we observe that

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

The first part tells us $c_{j,m}$ is some universal polynomial of degree j that is evaluated at 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

$$P_j(Y) = \frac{Y(Y-1)\cdots(Y-(j-1))}{j(j-1)\cdots(j-(j-1))} = \frac{Y(Y-1)\cdots(Y-(j-1))}{j!}. \tag{8.7}$$

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. For nonzero M, write

$$[M](X) = \sum_{j=0}^{\deg M} a_{j,M}(T)X^{p^j}, \quad a_{j,M}(T) \in \mathbf{F}_p[T]. \tag{8.8}$$

(In Theorem 2.5 we wrote the coefficients as $a_j(T)$ rather than as $a_{j,M}(T)$, but the coefficients depend on M and now we need to keep track of that information.) We know by Theorem 2.5 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 (8.6). 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 (8.4) is $\log_C([M](X)) = M\log_C(X)$. Explicitly, Carlitz found the formula

$$\log_C(X) = X - \frac{X^p}{T^p - T} + \frac{X^{p^2}}{(T^{p^2}-T)(T^p-T)} - \cdots = \sum_{j\ge 0}(-1)^j\frac{X^{p^j}}{L_j},$$

where $L_0 = 1$ and $L_j = (T^{p^j} - T)(T^{p^{j-1}} - T)\cdots(T^p - T)$ for $j \ge 1$. We will not 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

$$[M](X) = e_C(M\log_C(X)) = \sum_{j\ge 0}\frac{M^{p^j}(\log_C(X))^{p^j}}{D_j}. \tag{8.9}$$

This series is in $\mathbf{F}_p(T)[[X]]$. If we replace X with X/M in the first equation of (8.9), then

$$[M](X/M) = e_C(M\log_C(X/M)). \tag{8.10}$$

As $\deg M \to \infty$, $1/M \to 0$ in $\mathbf{F}_p((1/T))$, so $M\log_C(X/M) = X + \sum_{j\ge 1}(-1)^j X^{p^j}/(L_j M^{p^j-1})$ tends to X in $\mathbf{F}_p((1/T))[[X]]$ as $\deg M \to \infty$. Therefore (8.10) implies

$$\lim_{\deg M\to\infty}[M](X/M) = e_C(X),$$

where the convergence on the left is coefficientwise convergence as a series in X. This is a Carlitz analogue of the formula $\lim_{n\to\infty}(1 + x/n)^n = e^x$. I thank Darij Grinberg for this observation.

Returning to (8.9), replace M with a variable Y:

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

which defined the polynomials $E_j(Y) \in \mathbf{F}_p(T)[Y]$. (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 $Y = M \in \mathbf{F}_p[T] - \{0\}$ in (8.9), we get

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

so a comparison with (8.8) gives for $M \neq 0$ that

$$a_{j,M}(T) = E_j(M) \text{ for } 0 \leq j \leq \deg M, \text{ and } E_j(M) = 0 \text{ for } j > \deg M. \tag{8.11}$$

Since $[0](X) = 0$, also $E_j(0) = 0$. That and (8.11) implies $E_j(X)$ is divisible by

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

where the product runs over all h, not just monic h, and includes $h = 0$. This product has degree $p^j = \deg E_j(X)$, so it differs from $E_j(X)$ by a factor in $\mathbf{F}_p(T)^\times$. Since $E_j(T^j) = a_{j,T^j}(T) = 1$ (because T^j is monic), we obtain

$$E_j(X) = \frac{\prod_{\deg h<j}(X-h)}{\prod_{\deg h<j}(T^j-h)} = \frac{\prod_{\deg h<j}(X-h)}{\prod_{\substack{h \text{ monic} \\ \deg h=j}} h} = \frac{\prod_{\deg h<j}(X-h)}{D_j}, \tag{8.12}$$

where the last formula comes from Remark 8.2 and the product in the numerator includes $h = 0$ for every j. Therefore when $M \neq 0$ and $j \leq \deg M$,

$$a_{j,M} = E_j(M) = \frac{\prod_{\deg h<j}(M-h)}{D_j}, \tag{8.13}$$

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. (In particular, $a_{j,0} = 0$.) This is the analogue of (8.7) and suggests that, on account of the degrees involved, $E_j(X)$ is an analogue of $\binom{X}{p^j}$ and D_j, the denominator in the coefficient of X^{p^j} in $e_C(X)$, is an analogue of $(p^j)!$.

Example 8.8. $E_0(X) = X/D_0 = X$ and $E_1(X) = \prod_{c\in\mathbf{F}_p}(X-c)/D_1 = (X^p-X)/(T^p-T)$, so $a_{0,M} = E_0(M) = M$ and $a_{1,M} = E_1(M) = (M^p - M)/(T^p - T)$.

We can extend the polynomials $E_j(X)$ to an analogue of all $\binom{X}{n}$ as follows. For $n \geq 0$, write n in base p, say $n = c_0 + c_1p + \cdots + c_kp^k$ where $0 \leq c_i \leq p-1$. Define

$$\mathcal{E}_n(X) = E_0(X)^{c_0}E_1(X)^{c_1}\cdots E_k(X)^{c_k} \in \mathbf{F}_p(T)[X].$$

Then $\deg \mathcal{E}_n(X) = c_0 + c_1p + \cdots + c_kp^k = n$ and $\mathcal{E}_{p^j}(X) = E_j(X)$. One justification of the analogy $\mathcal{E}_n(X) \leftrightarrow \binom{X}{n}$ occurs in the description of integral-valued polynomials.

- For all $m \in \mathbf{Z}$, $\binom{m}{n} \in \mathbf{Z}$, and the polynomials $\binom{X}{n}$ are a basis of the integral-valued polynomials: $f(X) \in \mathbf{Q}[X]$ satisfies $f(\mathbf{Z}) \subset \mathbf{Z}$ if and only if $f(X) = \sum_{n=0}^{d} c_n\binom{X}{n}$ with $c_n \in \mathbf{Z}$, and the coefficients are unique.

- $\mathcal{E}_n(\mathbf{F}_p[T]) \subset \mathbf{F}_p[T]$ since $E_j(\mathbf{F}_p[T]) \subset \mathbf{F}_p[T]$, by (8.11). The $\mathcal{E}_n(X)$'s are an $\mathbf{F}_p[T]$-basis of all polynomials $f(X) \in \mathbf{F}_p(T)[X]$ satisfying $f(\mathbf{F}_p[T]) \subset \mathbf{F}_p[T]$, and the $E_j(X)$'s are an $\mathbf{F}_p[T]$-basis of the p-polynomials $f(X)$ satisfying $f(\mathbf{F}_p[T]) \subset \mathbf{F}_p[T]$.

Remark 8.9. A wider context for this construction of a basis of integral-valued polynomials via digit expansions (enlarging the set of E_j's to the $\mathcal{E}_n$'s by writing n in base p) is in [4].

We said already that D_j, which is the common denominator of the coefficients of $E_j(X)$, is a Carlitz analogue of $(p^j)!$. Carlitz considered the common denominator of the coefficients of $\mathcal{E}_n(X)$ to be an analogue of $n!$. This denominator is

$$n!_C := D_0^{c_0} D_1^{c_1} \cdots D_k^{c_k}$$

where $n = \sum_{i=0}^k c_i p^i$ with $0 \le c_i \le p-1$. (Since $D_0 = 1$, c_0 plays no role in $n!_C$.) For example, $n!_C = 1$ for $0 \le n \le p-1$ and $p!_C = D_1 = T^p - T$. To see a genuine analogy between $n!$ and $n!_C$, let's compare their factorizations. Legendre proved

$$n! = \prod_p p^{\sum_{s\ge 1}[n/p^s]}. \tag{8.14}$$

The irreducible factorization of $n!_C$ was determined by Sinnott and looks just like (8.14).

Theorem 8.10. *For all $n \ge 0$, $n!_C = \displaystyle\prod_{\text{monic } \pi} \pi^{\sum_{s\ge 1}[n/\mathrm{N}(\pi)^s]}$.*

In Theorem 8.10 and below, a product over π is understood to have π irreducible.

Proof. Since $n!_C$ is a product of powers of the polynomials D_j, first we write down an explicit formula for the D_j's. We know D_0 doesn't matter. For $j \ge 1$, the recursive formula $D_j = (T^{p^j} - T)D_{j-1}^p$ from Definition 8.1 can be turned into an explicit formula:

$$D_j = \prod_{m=1}^{j} (T^{p^m} - T)^{p^{j-m}}.$$

Therefore when $n = \sum_{i=0}^k c_i p^i$,

$$\begin{aligned}
n!_C &= \prod_{i=1}^{k} D_i^{c_i} \\
&= \prod_{i=1}^{k} \prod_{m=1}^{i} (T^{p^m} - T)^{c_i p^{i-m}} \\
&= \prod_{m=1}^{k} \prod_{i=m}^{k} (T^{p^m} - T)^{c_i p^{i-m}} \\
&= \prod_{m=1}^{k} (T^{p^m} - T)^{c_m + c_{m+1}p + \cdots + c_k p^{k-m}} \\
&= \prod_{m=1}^{k} (T^{p^m} - T)^{[n/p^m]}.
\end{aligned}$$

In $\mathbf{F}_p[T]$, $T^{p^m}-T$ factors into the monic irreducibles π of degree dividing m, each irreducible appearing once. Therefore

$$n!_C = \prod_{m=1}^{k} \prod_{\deg \pi | m} \pi^{[n/p^m]} = \prod_{\deg \pi \le k} \pi^{\sum_{m \equiv 0 \bmod \deg \pi} [n/p^m]},$$

where the sum in the exponent on π runs over all $m \ge 1$ that are divisible by $\deg \pi$; we don't have to restrict to $m \le k$ since if $m > k$ then $p^m \ge p^{k+1} > n$, so $[n/p^m] = 0$. Writing $m = s \deg \pi$ with $s \ge 1$,

$$n!_C = \prod_{\deg \pi \le k} \pi^{\sum_{s \ge 1} [n/p^{s \deg \pi}]} = \prod_{\deg \pi \le k} \pi^{\sum_{s \ge 1} [n/\mathrm{N}(\pi)^s]}.$$

We can let the product run over all monic π, not just those with degree at most k, since if $\deg \pi \ge k+1$ then $\mathrm{N}(\pi) > n$ and therefore $[n/\mathrm{N}(\pi)^s] = 0$ for all $s \ge 1$. $\square$

Example 8.11. For all π we have $\mathrm{N}(\pi) \ge p$, so the product in Theorem 8.10 is 1 for $0 \le n < p$, which is also $n!_C$. For $n = p$ we have

$$\prod_{\text{monic } \pi} \pi^{\sum_{s \ge 1} [p/\mathrm{N}(\pi)^s]} = \prod_{\deg \pi = 1} \pi^{[p/p]} = \prod_{c \in \mathbf{F}_p} (T - c) = T^p - T = p!_C.$$

9. 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 every finite field $\mathbf{F}_q$. Set $[T](X) = X^q + TX$ rather than $X^p + TX$, and define $[T^n](X)$ by iteration and $[M](X)$ for $M \in \mathbf{F}_q[T]$ 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 in X (a linear combination of X, X^q, X^{q^2}, and so on) and its roots Λ_M in an algebraic closure $\overline{\mathbf{F}_q(T)}$ 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 polynomials and Carlitz torsion still work, by replacing p by q everywhere.[16] 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 $(\mathbf{F}_q[T]/M)^\times$.

Our treatment of quadratic reciprocity in $\mathbf{F}_p[T]$ for $p \ne 2$ extends to $\mathbf{F}_q[T]$ for odd q.

The Carlitz exponential $e_C(X)$ for $\mathbf{F}_q[T]$ has denominators $D_j = (T^{q^j} - T)D_{j-1}^q$ for $j \ge 1$ and $D_0 = 1$, and the zeros of this $e_C(X)$ are $\mathbf{F}_q[T]$-multiples of a transcendental ξ_q that is given by the same formula as ξ_p, with p replaced by q.

Appendix A. Additive Polynomials

Let A be an integral domain. A polynomial $f(X) \in A[X]$ is called *additive* if $f(X+Y) = f(X) + f(Y)$ in $A[X,Y]$. If A contains a field F then we say $f(X) \in A[X]$ is *F-linear* if it

[16] In Theorems 4.5 and 6.2, for instance, the distinction between $p = 2$ and $p \ne 2$ becomes $q = 2$ and $q \ne 2$, rather than q being even and q being odd.

is additive and $f(cX) = cf(X)$ for all $c \in F$.[17] 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 A be an integral domain of characteristic p.*

1) *If V is a finite additive subgroup of A then the product*

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

is an additive polynomial.

2) *If F is a finite field in A and V is an F-vector subspace of A, then the polynomial in part 1 is F-linear.*

Proof. 1) Call the product $f(X)$. For indeterminates X and Y, let $g(Y) = f(X+Y) - f(X) - f(Y)$ in $A[X][Y] = A[X,Y]$. We want to show $g(Y) = 0$ in $A[X,Y]$. 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 in A. For each $w \in V$,

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

We will show this is 0 in $A[X]$. 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 an arbitrary $u \in V$ (since $u + w$ and u are both roots of f), $g(w) = 0$ in $A[X]$. Since $g(w) = 0$ for each $w \in V$ and the Y-degree of g is less than $|V|$, $g = 0$ in $A[X,Y]$.

2) To show $f(cX) = cf(X)$ for all $c \in F$, we can assume $c \neq 0$ since the result is obvious for $c = 0$. Since multiplication by c permutes V,

$$f(cX) = \prod_{v \in V}(cX - v) = \prod_{v \in V}(cX - cv) = c^{|V|}f(X).$$

Let $q = |F|$ and $d = \dim_F(V)$, so $|V| = q^d$. From $c^q = c$ we get $c^{|V|} = c$, so $f(cX) = cf(X)$. □

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

To prove the converse, let $f(X)$ be additive in $A[X]$. 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 $f'(X) = f'(0) \in A$. 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 A 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 A 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 $f(X) = a_0X + g(X^p)$.

[17] Don't confuse this notion of an F-linear polynomial with the high school notion of a linear polynomial as one with degree at most 1.

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 $f(X) \equiv a_0 X \bmod X^p$. We have

$$\begin{aligned} g(X^p + Y^p) &= g((X+Y)^p) \\ &= f(X+Y) - a_0(X+Y) \\ &= f(X) + f(Y) - a_0 X - a_0 Y \\ &= g(X^p) + g(Y^p). \end{aligned}$$

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

$$f(X) = a_0 X + a_1 X^p + h(X^{p^2}).$$

If $h(X) = 0$ then we are done. If $h(X) \neq 0$ then $h(X)$ is divisible by X, so $f(X) \equiv a_0 X + a_1 X^p \bmod X^{p^2}$. By a similar argument from before, $h(X)$ is additive and this lets us pull out an $a_2 X^{p^2}$ term. Repeating this argument enough times, we eventually see $f(X)$ has the desired form since it is a polynomial. $\square$

Corollary A.3. *Let A be an integral domain containing a field F. If F is infinite, a polynomial $f(X) \in A[X]$ is F-linear if and only if $f(X) = aX$. If F is finite with size q then $f(X) \in A[X]$ is F-linear if and only if it is of the form $f(X) = b_0 X + b_1 X^q + b_2 X^{q^2} + \cdots + b_n X^{q^n}$ 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 corollary (aX for infinite F and $b_0 X + b_1 X^q + b_2 X^{q^2} + \cdots + b_n X^{q^n}$ when $|F| = q$) are F-linear.

To prove the converse, first suppose A has characteristic 0, so F is necessarily infinite. Then additivity alone already forces $f(X) = aX$. When A has characteristic p, additivity implies $f(X) = a_0 X + a_1 X^p + a_2 X^{p^2} + \cdots + a_m X^{p^m}$ for some m. The F-linearity says $f(cX) = cf(X)$ for all $c \in F$, so $a_i c^{p^i} = c a_i$, which means $c^{p^i} = c$ for all $c \in F$ and all 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 $(X^q - X) \mid (X^{p^i} - X)$ 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. $\square$

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] K. Conrad, *The digit principle*, J. Number Theory **84** (2000), 230–257.
[5] K. Conrad, Quadratic Reciprocity in Odd Characteristic, https://kconrad.math.uconn.edu/blurbs/ugradnumthy/QRcharp.pdf.
[6] D. Goss, "Basic Structures of Function Field Arithmetic", Springer-Verlag, Berlin, 1996.
[7] D. R. Hayes, *Explicit class field theory for rational function fields*, Bull. Amer. Math. Soc. **189** (1974), 77–91.
[8] C-N. Hsu, *On Artin's conjecture for the Carlitz module*, Compositio Math. **106** (1997), 247–266.
[9] S. Lang, "Algebra," 3rd revised ed., Springer-Verlag, New York, 2002.
[10] M. Rosen, "Number Theory in Function Fields," Springer-Verlag, New York, 2002.
[11] P. Samuel, "Algebraic Theory of Numbers," Dover, Mineloa, 2008.
[12] L. I. Wade, *Certain quantities transcendental over $GF(p^n, x)$*, Duke Math J. **8** (1941), 701–720.