Algebraicity of some Hilbert-Kunz multiplicities (modulo a conjecture)

Paul Monsky

Brandeis University, Waltham MA 02454-9110, USA. monsky@brandeis.edu

Abstract

Let $F$ be a finite field of characteristic 2 and $h$ be the element $x^3 + y^3 + xyz$ of $F[[x, y, z]]$. In an earlier paper we made a precise conjecture as to the values of the colengths of the ideals $(x^q, y^q, z^q, h^j)$ for $q$ a power of 2. We also showed that if the conjecture holds then the Hilbert-Kunz series of $H = uv + h$ is algebraic (of degree 2) over $\mathbb{Q}(w)$, and that $\mu(h)$ is algebraic (explicitly, $\frac{4}{3} + \frac{5}{14\sqrt{7}}$). In this note, assuming the same conjecture, we use a theory of infinite matrices to rederive this result, and we extend it to a wider class of $H$; for example $H = g(u, v) + h$. In a follow-up paper, under the same hypothesis, we will show that transcendental Hilbert-Kunz multiplicities exist.

1 A product on $X$

In this section we develop some general results about Hilbert-Kunz series and multiplicities for characteristic 2 power series. (There are similar results, implicit in [5], in all finite characteristics but they are harder to prove.)

Definition 1.1 $X$ is the vector space of functions $I \rightarrow Q$ where $I = [0, 1] \cap Z[\frac{1}{2}]$. If $f \neq 0$ is in the maximal ideal of $F[[u_1, \ldots, u_r]]$, char $F = 2$, then $\phi_f$ in $X$ is the function $\frac{i}{q} \rightarrow q^{-r} \deg(u_1^q, \ldots, u_r^q, f^i)$; here $q$ denotes a power of 2 and $\deg$ is colength in $F[[u_1, \ldots, u_r]]$. Note that $\phi_f$ is well-defined.

Definition 1.2 $\alpha$ in $X$ is convex if for all $i$ and $q$ with $0 < i < q$, $2\alpha\left(\frac{i}{q}\right) \geq \alpha\left(\frac{i-1}{q}\right) + \alpha\left(\frac{i+1}{q}\right)$.

Note that $\phi_f(0) = 0$, $\phi_f(1) = 1$, $\phi_f$ is convex and $\phi_f$ is Lipschitz. The first two assertions are clear. If we set $J = (u_1^q, \ldots, u_r^q)$ then multiplication by $f$ induces a map of $(J, f^{i-1})/(J, f^i)$ onto $(J, f^i)/(J, f^{i+1})$, yielding convexity. Finally, as Lipschitz constant we can take the Hilbert-Kunz multiplicity, $\mu$, of $f$. 

Definition 1.3 Suppose that $\alpha$ in $X$ is convex Lipschitz with $\alpha(0) = 0$ and $\alpha(1) = 1$. Then $\mu(\alpha) = \lim_{n \to \infty} \alpha(2^{-n}) \cdot 2^n$, while $S_\alpha$ is the element $\sum \alpha(2^{-n})(2w)^n$ of $Q[[w]]$. (The convexity of $\alpha$ shows that $n \to 2^n\alpha(2^{-n})$ is non-decreasing. Since $\alpha$ is Lipschitz, the function is bounded and the limit exists.)

Remarks When $\alpha = \phi_f$, $\mu(\alpha)$ and $S_\alpha(2^{-1}w)$ are just the Hilbert-Kunz multiplicity and Hilbert-Kunz series of $f$. Note that if $\alpha$ is as in Definition 1.3 then $\mu(\alpha) = \lim_{w \to 1^-} (1 - w)S_\alpha(w)$. For convexity shows that the co-efficients of the power series $(1 - w)S_\alpha(w)$ are $\geq 0$. So the limit is the value of this power series at 1. And we note that $\alpha(1) + (2\alpha(\frac{1}{2}) - \alpha(1)) + (4\alpha(\frac{1}{4}) - 2\alpha(\frac{1}{2})) + \cdots$ converges to $\mu(\alpha)$.

We next define a bilinear product $\# : X \times X \to X$ and show that if $f \neq 0$ and $g \neq 0$ are in the maximal ideals of $F[[u_1, \ldots, u_r]]$ and $F[[v_1, \ldots, v_s]]$, then $\phi_f \# \phi_g = \phi_h$, where $h$ is the element $f(u) + g(v)$ of $F[[u_1, \ldots, u_r, v_1, \ldots, v_s]]$. (There is a similar construction, implicit in [5], in any finite characteristic.)

Definition 1.4 Suppose $\alpha$ and $\beta$ are in $X$. We define $\alpha \# \beta(t)$ by induction on the denominator of $t$ in $I$, according to the following procedure:

Let $\alpha_0$ and $\alpha_1$ be the elements $t \to \alpha(\frac{1}{2})$ and $t \to \alpha(\frac{1}{2} + t)$ of $X$; define $\beta_0$ and $\beta_1$ similarly. Then:

(1) $\alpha \# \beta(0) = 0$; $\alpha \# \beta(1) = (\alpha(1) - \alpha(0))(\beta(1) - \beta(0))$

(2) If $0 \leq t \leq \frac{1}{2}$

$\alpha \# \beta(t) = \alpha_0 \# \beta_0(2t) + \alpha_1 \# \beta_1(2t)$

(3) If $\frac{1}{2} \leq t \leq 1$

$\alpha \# \beta(t) = \alpha_0 \# \beta_0(1) + \alpha_1 \# \beta_1(1) + \alpha_0 \# \beta_1(2t - 1) + \alpha_1 \# \beta_0(2t - 1)$

Note that when $t = 0$, $\frac{1}{2}$ or 1 the two definitions of $\alpha \# \beta(t)$ given by the above scheme coincide, so that $\alpha \# \beta$ is a well-defined element of $X$. $\#$ is evidently bilinear and symmetric; one can show that it is associative. It’s easy to see that if $\alpha$ is constant then $\alpha \# \beta = 0$, while if $\alpha$ is the identity function $t$, $\alpha \# \beta = (\beta(1) - \beta(0))t$. In particular, $t \# t = t$.

Now let $T_0$ and $T_1 : X \to X$ be the maps taking $\alpha$ to $t \to \alpha(\frac{1}{2})$ and $t \to \alpha(\frac{1}{2} + t)$. Replacing $t$ by $\frac{1}{2}$ in (2) above and by $\frac{1}{2} + t$ in (3) above gives:

Theorem 1.5 If $\gamma = \alpha \# \beta$ then:

$$T_0(\gamma) = (T_0(\alpha) \# T_0(\beta)) + (T_1(\alpha) \# T_1(\beta))$$

$$T_1(\gamma) = \gamma \left(\frac{1}{2}\right) + (T_0(\alpha) \# T_1(\beta)) + (T_1(\alpha) \# T_0(\beta))$$

We now recall some notation used in both [1] and [5]. By an $F[T]$-module we’ll mean a finitely generated $F[T]$-module annihilated by a power of $T$. $\Gamma$ is the
Grothendieck group of the set of isomorphism classes of such modules. There is a multiplication on \( \Gamma \) making it into a commutative ring; if \( V \) and \( W \) are \( F[T] \)-modules, a representative of their product is \( V \otimes W \), with \( T \) acting by \((T_V \otimes \text{id}) + (\text{id} \otimes T_W)\). There is a \( Z \)-basis \( \lambda_0, \lambda_1, \ldots \) of \( \Gamma \) with the following property. If \( V \) is an \( F[T] \)-module then the class of \( V \) in \( \Gamma \) is \( \sum c_i \lambda_i \) where \( c_i = (-1)^i \dim(T^i V / T^{i+1} V) \). Because \( \text{char } F = 2 \), the multiplicative structure of \( \Gamma \) is very simple; \( \lambda_i \lambda_j = \lambda_k \) where \( k \) is the “Nim-sum” of \( i \) and \( j \).

**Lemma 1.7** Suppose now that \( q = 2^n \) and \( 0 \leq i < q \). Since the Nim-sum of \( i \) and \( q \) is \( q + i \), \( \lambda_i \lambda_q = \lambda_{q+i} \) giving:

\[
\lambda_{n+1}(\alpha) = \lambda_n(\alpha_0) + \lambda_q \lambda_n(\alpha_1)
\]

**Theorem 1.8** If \( \gamma = \alpha \neq \beta \), \( \lambda_n(\gamma) = \lambda_n(\alpha) \cdot \lambda_n(\beta) \).

**Proof** We argue by induction on \( n \). Since \( \gamma(1) - \gamma(0) = (\alpha(1) - \alpha(0)) (\beta(1) - \beta(0)) \) the result holds for \( n = 0 \). Suppose that it’s true for a given \( n \). Lemma 1.7, Theorem 1.5 and the induction hypothesis show that \( \lambda_{n+1}(\gamma) = \lambda_n(\alpha_0) \lambda_n(\beta_0) + \lambda_n(\alpha_1) \lambda_n(\beta_1) + \lambda_q \lambda_n(\alpha_0) \lambda_n(\beta_1) + \lambda_q \lambda_n(\alpha_1) \lambda_n(\beta_0) \). But this is \((\lambda_n(\alpha_0) + \lambda_q \lambda_n(\alpha_1)) \cdot (\lambda_n(\beta_0) + \lambda_q \lambda_n(\beta_1)) \) which is \( \lambda_{n+1}(\alpha) \cdot \lambda_{n+1}(\beta) \) by Lemma 1.7.

**Theorem 1.9** Suppose \( h = f(u_1, \ldots, u_r) + g(v_1, \ldots, v_s) \). Then:

1. For each \( n \), \( \lambda_n(\phi_h) = \lambda_n(\phi_f) \cdot \lambda_n(\phi_g) \)
2. \( \phi_h \neq \phi_f \neq \phi_g \)

**Proof** With \( q = 2^n \), let \( V \) be as in the paragraph following Definition 1.6. As we’ve seen \( V \) represents the element \( q^r \lambda_i \lambda_j \) of \( \Gamma \). Replacing \( f \) by \( g \) we get a \( W \) representing the element \( q^r \lambda_i \lambda_j \) of \( \Gamma \). Then \( q^{r+s} \lambda_i \lambda_j \) is represented by \( F[[u_1, \ldots, u_r, v_1, \ldots, v_s]]/(u_1^q, \ldots, v_s^q) \) with \( T \) acting by multiplication by \( f(u_1, \ldots, u_r) + g(v_1, \ldots, v_s) = h \). Since this \( F[T] \)-module represents \( q^{r+s} \lambda_i \lambda_j \) we get (1). Suppose now that \( \phi_h(t) \neq \phi_f \neq \phi_g(t) \) for some \( t = \frac{i}{2^n} \). Choose such a \( t \) with \( i \) as small as possible. Then \( i \neq 0 \), and the co-efficients of \( \lambda_{i-1} \) in \( \lambda_n(\phi_h) \) and \( \lambda_n(\phi_f \neq \phi_g) \) differ. Theorem 1.8 then shows that \( \lambda_n(\phi_h) \neq \lambda_n(\phi_f) \lambda_n(\phi_g) \), contradicting (1).
Theorem 1.10 If $\alpha$ and $\beta$ are Lipschitz with Lipschitz constant $m$, then $\gamma = \alpha \# \beta$ is Lipschitz with Lipschitz constant $m^2$.

Proof We show that if $0 \leq j < 2q$ then $|\gamma \left(\frac{j+1}{2q}\right) - \gamma \left(\frac{j}{2q}\right)| \leq \frac{m^2}{2q}$, arguing by induction on $q$. Note first that $\alpha_0, \alpha_1, \beta_0$ and $\beta_1$ are all Lipschitz with Lipschitz constant $\frac{m}{2q}$. We claim that when $j < q$ the values of $\alpha_0 \# \beta_0$ (and of $\alpha_1 \# \beta_1$) at $\frac{j+1}{2q}$ and $\frac{j}{2q}$ differ by at most $\frac{m^2}{4q}$. (When $q = 1$, $j = 0$, and this is clear. When $q > 1$ we use the fact that $\alpha_0$ and $\beta_0$ (and $\alpha_1$ and $\beta_1$) have Lipschitz constant $\frac{m}{2q}$, together with the induction hypothesis.) Theorem 1.5 then shows that $\gamma \left(\frac{j+1}{2q}\right)$ and $\gamma \left(\frac{j}{2q}\right)$ differ by at most $\frac{m^2}{4q} + \frac{m^2}{4q} = \frac{m^2}{2q}$. The argument is similar when $j \geq q$, but now we make use of the values of $\alpha_0 \# \beta_1$ (and of $\alpha_1 \# \beta_0$) at $\frac{j+1}{2q}$ and $\frac{j}{2q}$. □

Lemma 1.11 Let $\delta_r, r \geq 1$, be the class of $F[T]/T^r$ in $\Gamma$; note that $\delta_r = \lambda_0 - \lambda_1 + \lambda_2 \cdots + (-)^{r-1}\lambda_{r-1}$. Then for $\alpha$ in $X$ the following are equivalent:

(1) $\alpha$ is convex.

(2) For each $n$, $L_n(\alpha) = \sum_{i=0}^{q} c_i (-)^i \lambda_i$ with $c_0 \geq c_1 \geq \cdots \geq c_{q-1}$.

(3) For each $n$, $L_n(\alpha)$ is a linear combination of $\delta_1, \ldots, \delta_q$ with the co-efficients of $\delta_1, \ldots, \delta_{q-1} \geq 0$.

Proof Since the $c_i$ in (2) is $\alpha \left(\frac{i+1}{q}\right) - \alpha \left(\frac{i}{q}\right)$, (1) and (2) are equivalent. Suppose (2) holds. If we set $c_q = 0$, then the formula for $\delta_r$ given above shows that $L_n(\alpha) = \sum_{i=0}^{q-1} (c_i - c_{i+1}) \delta_i$. Since $c_0 - c_1, \ldots, c_{q-2} - c_{q-1}$ are all $\geq 0$ we get (3). That (2) follows from (3) is easy. □

Lemma 1.12 Suppose $1 \leq r, s \leq q$. Then, in $\Gamma$, $\delta_r \delta_s$ is a linear combination of $\delta_1, \ldots, \delta_q$ with non-negative integer co-efficients. Furthermore $\delta_r \delta_q = r \delta_q$.

Proof Let $V$ and $W$ be the $F[T]$-modules $F[T]/T^r$ and $F[T]/T^s$ representing $\delta_r$ and $\delta_s$. Writing $V \otimes W$ (with $T$ acting by $T_V \otimes \text{id} + \text{id} \otimes T_W$) as a direct sum of cyclic $F[T]$-modules we get the first assertion. The second is an easy calculation. □

Theorem 1.13 If $\alpha$ and $\beta$ in $X$ are convex, then so is $\alpha \# \beta$.

Proof By Lemma 1.11, $L_n(\alpha)$ and $L_n(\beta)$ are each linear combinations of $\delta_1, \ldots, \delta_q$ with the co-efficients of $\delta_1, \ldots, \delta_{q-1} \geq 0$. By Lemma 1.12 the same is true of $L_n(\alpha) \cdot L_n(\beta)$. Theorem 1.8 and Lemma 1.11 then show that $\alpha \# \beta$ is convex. □

Theorem 1.14 Suppose that $\alpha$ in $X$ is convex Lipschitz with $\alpha(0) = 0$ and $\alpha(1) = 1$. Suppose further that $S_\alpha = \sum \alpha(2^{-n})(2w)^n$ lies in a finite extension, $L$, of $Q(w)$. (We extend the imbedding of $Q[w]$ in $Q[[w]]$ to their fields of fractions.) Then $\mu(\alpha)$ is algebraic over $Q$ of degree $\leq [L : Q(w)]$. In fact there
is a valuation ring containing $Q[w]$ in $L$ whose maximal ideal contains $w - 1$
and whose residue class field contains a copy of $Q(\mu(\alpha))$.

**Proof** Take $H$ irreducible in $Q[W, T]$ so that $H(w, (1 - w)S_a) = 0$. Then for
any $z$ in the open unit disc, $H(z, (1 - z)S_a(z)) = 0$. The remarks following Def-
nition 1.3 show that $H(1, \mu(\alpha)) = 0$. Since $H(1, T) \neq 0$, $\mu(\alpha)$ is algebraic over
$Q$. Let $g$ be $\text{Irr}(\mu(\alpha), Q)$. Then $(W - 1, g(T))$ is a maximal ideal in $Q[W, T]/H$
and we take a valuation ring in $L$ that contains $Q[w, (1 - w)S_a] = Q[W, T]/H$,
and whose maximal ideal contracts to the above maximal ideal. \qed

2 A calculation from [2], revisited

Let $f$ be the element $x^3 + y^3 + xyz$ of $Z/2 [x, y, z]$, defining a nodal cubic.
The values of $\phi_f$ at $\frac{1}{q}$ are known, and in particular, $\mu(f) = \frac{7}{3}$. In [2] we
conjectured a precise value for all $\phi_f \left( \frac{1}{q} \right)$, and showed that the conjecture
implied that $\mu(wv + f)$ is $\frac{4}{3} + \frac{5}{4\sqrt{7}}$. In this section we’ll rework this result using
infinite matrix techniques from [3]; this approach will give rise to more general
theorems.

**Definition 2.1** $t$ and $\epsilon$ will denote the elements $t \to 1$, $t \to t$ and $t \to t - t^2$
of $X$.

**Definition 2.2** For $m = 0, 1, 2, \ldots$ and $t$ in $I$, $\phi_m(t)$ is defined by induction on
the denominator of $t$ as follows:

1. $\phi_m(0) = \phi_m(1) = 0$
2. If $0 \leq t \leq \frac{1}{2}$, $8\phi_m(t) = \phi_{m+1}(2t) + (8m + 6)t$ for $m$ even, and $\phi_{m-1}(2t) +
\epsilon(2t) + (8m + 6)t$ for $m$ odd.
3. If $\frac{1}{2} \leq t \leq 1$, $8\phi_0(t) = \phi_0(2t - 1) + 6(1 - t)$
4. If $\frac{1}{2} \leq t \leq 1$, $8\phi_m(t) = \phi_{m-1}(2t - 1) + \epsilon(2t - 1) + (8m + 6)(1 - t)$ for
$m \neq 0$ even, and $\phi_{m+1}(2t - 1) + (8m + 6)(1 - t)$ for $m$ odd.

When $t = 0, \frac{1}{2}$ or 1, the two definitions of $\phi_m(t)$ given by the above scheme
evidently coincide. So the $\phi_m$ are well-defined elements of $X$. Replacing $t$ by
$\frac{t}{2}$ in (2) and by $\frac{1 + t}{2}$ in (3) and (4) we get the “magnification rules”:

1. $8T_0(\phi_0) = \phi_1 + 3t$
2. When $m \neq 0$ is even,
$8T_0(\phi_m) = \phi_{m+1} + (4m + 3)t$
3. When $m$ is odd,
$8T_0(\phi_m) = \phi_{m-1} + \epsilon + (4m + 3)t$

Note also that $4T_0(\epsilon) = \epsilon + t$ and that $4T_1(\epsilon) = \epsilon + (1 - t)$. 

5
Conjecture 2.3 If \( f = x^3 + y^3 + xyz \), then \( \phi_f = t + \phi_0 \) with \( \phi_0 \) as above.

In [2] we presented evidence for a conjecture easily seen to be equivalent to this. We noted in particular that both sides agree at all \( \frac{1}{q} \) and at each \( \frac{1}{512} \).

**Theorem 2.4** If \( E_1 = \epsilon \# \phi_0 \) then \( \lim_{n \to \infty} E_1(2^{-n})2^n = \frac{1}{3} + \frac{5}{14\sqrt{7}} \).

Suppose now that Conjecture 2.3 holds. Then \( t + E_1 = (t + \epsilon) \# (t + \phi_0) = \phi_{uv} \# \phi_f = \phi_{uv + f} \). So Theorem 2.4 tells us that the Hilbert-Kunz multiplicity of \( uv + x^3 + y^3 + xyz \) is \( \lim_{n \to \infty} (2^{-n} + E_1(2^{-n}))2^n = \frac{4}{3} + \frac{5}{14\sqrt{7}} \), an observation made in [2]. We now give a proof of Theorem 2.4 using the techniques of [3].

**Lemma 2.5** Let \( T : X \to X \) be \( 32T_0 \). Set \( E_k = \epsilon \# \phi_{k-1} \). Then:

1. \( T(E_1) = E_1 + E_2 + 6t \)
2. \( T(E_k) = E_{k-1} + E_{k+1} + (8k - 2)t + (\epsilon \# \epsilon) \) for \( k > 1 \)
3. \( T(\epsilon \# \epsilon) = 4(\epsilon \# \epsilon) + 4t \), and \( T(t) = 16t \)

**Proof** Suppose \( k \) is even. Then \( T(E_k) = 32T_0(\epsilon \# \phi_{k-1}) = (4T_0(\epsilon \# 8T_0(\phi_{k-1}))) + (4T_1(\epsilon \# 8T_1(\phi_{k-1}))) \). The magnification rules following Definition 2.2 show that this is \( (\epsilon + t) \# (\phi_{k-2} + \epsilon + (4k-1)t) + (\epsilon + 1 - t) \# (\phi_k + (4k-1)(1-t)) \). Expanding out we get \( (\epsilon \# \phi_{k-2}) + (4k-1)t + (\epsilon \# \phi_k) + (4k-1)t + (\epsilon \# \epsilon) = E_{k-1} + E_{k+1} + (8k - 2)t + (\epsilon \# \epsilon) \). The other parts of the lemma are derived similarly.

**Lemma 2.6** Let \( S \) be the power series \( \sum E_1(2^{-n})(32w)^n \). Then \( (1 - 16w)(1 - 4w)(1 - 2w)^2S = 4w(1 - 2w)^2 + (2w - 12w^2)\sqrt{1 - 4w^2} \).

**Proof** Let \( l : X \to Q \) be evaluation at 1, so that \( l(E_k) = 0 \) for each \( k \), and \( l(\epsilon \# \epsilon) = 0 \), while \( l(t) = 1 \). Then \( E_1(2^{-n})32^n = l(T^n(E_1)) \) and \( S \) is just \( \sum l(T^n(E_1))w^n \). If we take \( Y \) to be the subspace of \( X \) spanned by \( \epsilon \# \epsilon \) and \( t \), Lemma 2.5 shows that we are in the situation of Example 5.12 of [3]. The final line of that paper is the desired result.

Theorem 2.4 is now easily proved. Lemma 2.6 shows that the value, \( \lambda \), of \( (1 - 16w)S \) at \( w = \frac{1}{16} \) is \( \left( \frac{4}{3} \cdot \frac{64}{49} \right) \left( \frac{4}{16} \cdot \frac{49}{64} \cdot \frac{64}{61} \cdot \frac{61}{64} \right) = \frac{1}{3} + \frac{5}{14\sqrt{7}} \). Furthermore, \( S - \frac{\lambda}{1 - 16w} \) is holomorphic in the disc \( |w| < \frac{1}{4} \). It follows that \( S \left( \frac{w}{16} \right) - \frac{\lambda}{1 - w} \) is holomorphic in \( |w| < 4 \), and so the co-efficients in its power series expansion \( \to 0 \). So \( E_1(2^{-n}) \cdot 2^n - \lambda \to 0 \), the desired result.

We conclude this section by showing that the \( \phi_m \) of Definition 2.2 are convex and Lipschitz.

**Lemma 2.7** \( \phi_m \left( \frac{1}{q} \right) \leq \frac{4m+4}{3q} \) for even \( m \) and \( \frac{4m+3}{3q} \) for odd \( m \).
Proof When \( q = 2 \), \( \phi_{m} \left( \frac{1}{q} \right) = \frac{4m+3}{4q} \). We argue by induction. Suppose \( q \geq 2 \).

If \( m \) is even, \( \phi_{m} \left( \frac{1}{2q} \right) = \frac{1}{8} \phi_{m+1} \left( \frac{1}{q} \right) + \frac{4m+3}{8q} \). By the induction hypothesis this is \( \leq \frac{4m+7}{24q} + \frac{4m+3}{8q} = \frac{4m+4}{3(2q)} \).

If \( m \) is odd, \( \phi_{m} \left( \frac{1}{2q} \right) = \frac{1}{8} \phi_{m+1} \left( \frac{1}{q} \right) + \frac{1}{8q} - \frac{1}{8q^{2}} + \frac{4m+3}{8q} \).

By the induction hypothesis this is \( \leq \frac{4m+4}{24q} + \frac{4m+4}{8q} = \frac{4m+3}{3(2q)} \). \( \square \)

Lemma 2.8 \( \phi_{m} \left( 1 - \frac{1}{q} \right) \leq \frac{4m+4}{3q} \) for odd \( m \) and \( \frac{4m+3}{3q} \) for even \( m \).

Proof \( q = 2 \) is clear. Suppose \( q \geq 2 \); we argue by induction. If \( m \) is odd, \n
\[ \phi_{m} \left( 1 - \frac{1}{2q} \right) = \frac{1}{8} \phi_{m+1} \left( 1 - \frac{1}{q} \right) + \frac{4m+3}{8q} \]

while if \( m \neq 0 \) is even, \( \phi_{m} \left( 1 - \frac{1}{2q} \right) = \frac{1}{8} \phi_{m+1} \left( 1 - \frac{1}{q} \right) + \frac{1}{8q} - \frac{1}{8q^{2}} + \frac{4m+3}{8q} \), and we continue as in the proof of Lemma 2.7. Finally, \( \phi_{0} \left( 1 - \frac{1}{2q} \right) = \frac{1}{8} \phi_{0} \left( 1 - \frac{1}{q} \right) + \frac{3}{8q} \).

By the induction hypothesis this is \( \leq \frac{1}{8q} + \frac{3}{8q} = \frac{1}{2q} \). \( \square \)

Lemma 2.9 \( \phi_{m} \left( \frac{s+1}{2q} \right) \) and \( \phi_{m} \left( \frac{s-1}{2q} \right) \) are \( \leq \phi_{m} \left( \frac{1}{2} \right) \).

Proof If \( m \) is odd, \( 8 \left( \phi_{m} \left( \frac{1}{2} \right) - \phi_{m} \left( \frac{s+1}{2q} \right) \right) = \frac{4m+3}{q} - \phi_{m+1} \left( \frac{1}{q} \right) \). By Lemma 2.7 this is \( \geq \frac{4m+3}{q} - \frac{4m+8}{3q} \geq 0 \).

Also \( 8 \left( \phi_{m} \left( \frac{1}{2} \right) - \phi_{m} \left( \frac{s-1}{2q} \right) \right) \geq \frac{4m+3}{q} - \phi_{m-1} \left( 1 - \frac{1}{q} \right) \).

By Lemma 2.8 this is \( \geq \frac{4m+3}{q} - \frac{4m-1}{3q} \geq 0 \). The argument for even \( m \) is similar. \( \square \)

Theorem 2.10 The \( \phi_{m} \) are convex and Lipschitz.

Proof To prove convexity, we show that if \( 0 < j < 2q \), then \( 2\phi_{m} \left( \frac{j}{2q} \right) - \phi_{m} \left( \frac{j-1}{2q} \right) - \phi_{m} \left( \frac{j+1}{2q} \right) \geq 0 \), arguing by induction on \( q \). The case \( q = 1 \) is immediate. When \( j < q \) the induction assumption tells us that \( 2\phi_{s} \left( \frac{j}{q} \right) - \phi_{s} \left( \frac{j-1}{q} \right) - \phi_{s} \left( \frac{j+1}{q} \right) \geq 0 \) for each \( s \); this and the fact that \( \epsilon \) and \( t \) are convex gives the result. When \( j > q \), the induction assumption tells us that \( 2\phi_{s} \left( \frac{j}{q} \right) - \phi_{s} \left( \frac{j-1}{q} \right) - \phi_{s} \left( \frac{j+1}{q} \right) \geq 0 \); this and the convexity of \( \epsilon \) and \( 1-t \) give the result. Finally the case \( j = q \) is handled by Lemma 2.9. Note also that Lemmas 2.7 and 2.8 show that \( |\phi_{m} \left( \frac{1}{q} \right) - \phi_{m}(0)| \) and \( |\phi_{m} \left( 1 - \frac{1}{q} \right) - \phi_{m}(1)| \) are each \( \leq \frac{4m+4}{3q} \). Since \( \phi_{m} \) is convex, it follows that it is Lipschitz with Lipschitz constant \( \frac{4m+4}{3q} \). \( \square \)

3 Algebraicity results

We generalize the calculations of Section 2 to show:

Theorem 3.1 Suppose \( \beta_{1} \) lies in a finite dimensional subspace of \( X \) stable under \( T_{0} \) and \( T_{1} \), and is convex Lipschitz. Set \( E_{1} = \beta_{1} \neq \phi_{0} \) with \( \phi_{0} \) as in
**Definition 2.2.** Then the power series \( S_{t+E_1}(w) \) is algebraic over \( Q(w) \), and \( \mu(t+E_1) \) is algebraic over \( Q \).

**Proof** Since \( \beta_1 \) and \( \phi_0 \) are convex Lipschitz, the same is true of \( t + E_1 \). In view of Theorem 1.14 we only need to prove the result for \( S \). We shall mimic the proof of Theorem 2.4. Take \( \beta_1, \ldots, \beta_l, 1, t \) spanning a space stable under \( T_0 \) and \( T_1 \). We are free to modify each \( \beta_j \) by a linear combination of \( 1 \) and \( t \) and so may assume \( \beta_j(0) = \beta_j(1) = 0 \). Then

\[
T_0(\beta_j) = (a \text{ linear combination of } \beta_j) + a \text{ multiple of } t,
\]

while

\[
T_1(\beta_j) = (a \text{ linear combination of } \beta_j) + a \text{ multiple of } (1-t).
\]

Since \( T_0(\beta_j)(1) = T_1(\beta_j)(0) = \beta_j \left( \frac{1}{2} \right) \) we get:

\[
T_0(\beta_j) = \sum r_{i,j} \beta_i + c_j t
\]

\[
T_1(\beta_j) = \sum s_{i,j} \beta_i + c_j (1-t)
\]

with the \( r_{i,j} \), the \( s_{i,j} \) and the \( c_j \) all in \( Q \).

We proceed in several steps:

I) Let \( R \) and \( S \) be the elements \( |r_{i,j}| \) and \( |s_{i,j}| \) of \( M_l(Q) \). We define an infinite matrix \( V \) with rows and columns indexed by the positive integers as follows. \( V \) is built up out of \( l \) by \( l \) blocks. The initial diagonal block is \( S \) while all succeeding diagonal blocks are matrices of zeroes. The blocks just below the diagonal blocks are alternately \( R \) and \( S \), as are the blocks just to the right of the diagonal blocks. All other entries are zero.

II) Let \( \phi_m \) be as in Definition 2.2. If \( m \geq 0 \) and \( 1 \leq j \leq l \) let \( E_{j+lm} = \beta_j \# \phi_m \); note that \( E_1 = \beta_1 \# \phi_0 \) in accord with the statement of the theorem. \( Y \subset X \) is the subspace spanned by \( t \) and the \( \beta_j \# \epsilon \), and we define \( y_1, y_2, \ldots \) in \( Y \) as follows. If \( 1 \leq j \leq l \), \( y_j = 6c_j t \). If \( m > 0 \),

\[
y_{j+lm} = (8m+6)c_j t = \sum r_{i,j} (\beta_i \# \epsilon)
\]

for odd \( m \) and \( \sum s_{i,j} (\beta_i \# \epsilon) \) for even \( m \). Note that \( 4T_0(\beta_j \# \epsilon) = T_0(\beta_j) \# (\epsilon + t) + T_1(\beta_j) \# (\epsilon + 1 - t) \), so that \( Y \) is stable under \( T_0 \).

III) With notation as above we claim that \( 8T_0(E_j) = \sum s_{i,j} E_i + y_j \). This amounts to:

1. If \( 1 \leq j \leq l \), \( 8T_0(E_j) = \sum s_{i,j} E_i + \sum r_{i,j} E_{i+t} + y_j \)
2. If \( m \) is odd, \( 8T_0(E_{j+lm}) = \sum r_{i,j} E_{i+lm-t} + \sum s_{i,j} E_{i+lm+t} + y_{j+lm} \)
3. If \( m > 0 \) is even, \( 8T_0(E_{j+lm}) = \sum s_{i,j} E_{i+lm-t} + \sum r_{i,j} E_{i+lm+t} + y_{j+lm} \)

Note that the left hand side of (3) is \( 8T_0(\beta_j \# \phi_m) = (\sum s_{i,j} \beta_i + c_j t) \# (\phi_{m+1} + (4m+3) t) + (\sum s_{i,j} \beta_i + c_j (1-t)) \# (\phi_{m-1} + \epsilon + (4m+3)(1-t)) \). Expanding out and using the definition of \( y_{j+lm} \) we get (3). Similar calculations give (1) and (2).

IV) Now set \( s = 2l \). It’s convenient to view the matrix \( V \) of \( I \) as built up out of \( s \) by \( s \) blocks. Set \( D = \begin{pmatrix} S & R \\ R & 0 \end{pmatrix} \) and \( B = \begin{pmatrix} 0 & R \\ R & 0 \end{pmatrix} \) in \( M_s(Q) \). Then the diagonal blocks of \( V \) are a single \( D \) followed by \( B \)’s. If we take \( A = \begin{pmatrix} 0 & S \\ S & 0 \end{pmatrix} \) and \( C = \begin{pmatrix} S & 0 \\ 0 & 0 \end{pmatrix} \), then the blocks just below the diagonal blocks are all \( A \)’s,
while those just to the right of the diagonal blocks are all $C$’s. And all other entries are zero.

The proof of Theorem 3.1 is now easy. III and IV tell us that we are in the situation of Theorem 5.11 of [3] with $T = 8T_0$ and $s, A, B, C, D$ as above. (Note that the $y_j$ are all in $Y$, that $Y$ is finite-dimensional and stable under $T$, and that the condition of Lemma 5.10 of [3] on the sequence $y_1, y_2, \ldots$ is trivially satisfied.) Let $l : X \to Q$ be evaluation at 1 so that each $l(E_j) = 0$. Then Theorem 5.11 of [3] shows that $\sum l(T^n(E_1))w^n = \sum E_1(2^{-n})(8w)^n$ is algebraic over $Q(w)$. So the same is true of

$$\frac{1}{1-w} + \sum E_1(2^{-n})(2w)^n = \sum (2^{-n} + E_1(2^{-n}))(2w)^n = S + E_1(w).$$

**Definition 3.2** $g \neq 0$ in the maximal ideal of $F[[u_1, \ldots, u_r]]$ is “strongly rational” if $\phi_g$ lies in a finite dimensional subspace of $X$ stable under $T_0$ and $T_1$.

The following is shown in [4] and [5]:

**Theorem 3.3**

1. If $F$ is finite and $r = 2$, $g$ is strongly rational.
2. If $g$ is strongly rational, the Hilbert-Kunz series of $g$ lies in $Q(w)$, and $\mu(g)$ is rational.
3. If $g(u_1, \ldots, u_r)$ and $h(v_1, \ldots, v_s)$ are strongly rational, then so are $g(u) + h(v)$, $g(u)h(v)$, and all powers of $g(u)$.

**Remark** Much of the above is easy to prove. (1) however makes use of a result on the finiteness of the number of ideal classes in certain 1-dimensional rings. And the proof of (3) for $g(u) + h(v)$ (or rather the generalization of this result to arbitrary finite characteristic $p$) isn’t easy. But when $p = 2$ there’s an immediate proof. Namely suppose that $V_1$ and $V_2$ are finite dimensional subspaces of $X$ containing $\phi_g$ and $\phi_h$ and stable under $T_0$ and $T_1$. Then the space spanned by 1 and $V_1 \# V_2$ is finite dimensional and stable under $T_0$ and $T_1$. Furthermore it contains $\phi_g \# \phi_h = \phi_{g(u) + h(v)}$.

If $g$ is strongly rational, Theorem 3.1 tells us that $S + (\phi_g \# \phi_0)$ is algebraic over $Q(w)$ and that $\mu(t + (\phi_g \# \phi_0))$ is algebraic. Now $t + (\phi_g \# \phi_0) = \phi_g \# (t + \phi_0)$. This gives:

**Theorem 3.4** Suppose that Conjecture 2.3 holds; that is to say that $t + \phi_0 = \phi_{x^3 + y^3 + x^2 y}$. Then if $g$ in $F[[u_1, \ldots, u_r]]$ is strongly rational, the Hilbert-Kunz series of $g(u_1, \ldots, u_r) + x^3 + y^3 + xyz$ is algebraic over $Q(w)$, and the Hilbert-Kunz multiplicity is algebraic. In particular using Theorem 3.3 we find that if we assume Conjecture 2.3 then these algebraicity results hold for $\sum g_i(u_i, v_i) + x^3 + y^3 + xyz$ whenever $F$ is finite over $Z/2$.  

9
In Theorem 3.1 it is possible in theory, once the \( r_{i,j} \), the \( s_{i,j} \) and the \( c_j \) are known, to get a polynomial relation between \( w \) and \( S_{t+E_1} \) and compute \( \mu(t+E_1) \) by using the methods of [3]. This is daunting in practice but we'll give one interesting partial result. Let \( M \) be the smallest subspace of \( X/(Q+Q\cdot t) \) that contains the image of \( \beta_1 \) and is stable under \( T_0 \) and \( T_1 \); our hypotheses show it to be finite dimensional. If \( J_0 \) and \( J_1 \) are maps \( M \rightarrow M \) let \( \Psi_{J_0,J_1}(x,w) \) be the 2-variable polynomial \( \det[xI-w^2(J_0+xJ_1)(J_1+xJ_0)] \).

**Theorem 3.5** In the situation of Theorem 3.1, \( \sum E_1(2^{-n})(8w)^n \) lies in the splitting field over \( Q(w) \) of \( \Psi_{T_0,T_1}(x,w) \).

**Proof** We adopt the notation of Theorem 3.1 and its proof. \( \sum E_1(2^{-n})(8w)^n = \sum (T^n(E_1))w^n \), and Theorem 5.11 of [3] shows that this power series lies in a certain extension \( \mathcal{L} \) of \( Q(w) \) constructed from the matrices \( A, B \) and \( C \). We saw in [3] that \( \mathcal{L} \subset \) a splitting field over \( Q(w) \) of \( \det[xI_w-w(Ax^2+Bx+C)] \).

This last matrix is

\[
\begin{pmatrix}
xI_l & -wx(R+xS) \\
-w(S+xR) & xI_l
\end{pmatrix}.
\]

So our determinant is just

\[
x^l \det \begin{pmatrix}
I_l & -w(R+xS) \\
-w(S+xR) & xI_l
\end{pmatrix}.
\]

Since \( R \) and \( S \) give the action of \( T_0 \) and \( T_1 \) on \( M \), this last determinant is \( \Psi_{T_0,T_1}(x,w) \). \( \Box \)

**4 A (very) partially worked example**

Suppose \( \beta_1 = \phi_9 \) with \( g = u^6 + u^3v^3 + v^6 \). The methods of [4] show that \( M \) is five dimensional, that the action of \( 4T_0 \) on \( M \) is given by \( \beta_1 \rightarrow \beta_2 \rightarrow \beta_3 \rightarrow \beta_1 \), \( \beta_4 \rightarrow \beta_5 \rightarrow 0 \), and that the action of \( 4T_1 \) is given by \( \beta_5 \rightarrow \beta_4 \rightarrow \beta_3 \rightarrow \beta_5 \), \( \beta_2 \rightarrow \beta_1 \rightarrow 0 \). A Maple calculation then shows that \( \Psi_{4T_0,4T_1}(x,w) = -x^2\Psi^* \) where \( \Psi^* \) is the reciprocal polynomial \( w^{10}(x^6+1) - (2w^8+w^4)(x^5+x) - (2w^8-3w^6-2w^2)(x^4+x^2) + (2w^{10}-w^8+2w^6-4w^4-1)x^3 \). In an algebraic closure of \( Q(w) \) let \( \rho, \sigma \) and \( \tau \) be the roots of \( \Psi^* \) having positive ord; the other 3 roots are \( \rho^{-1}, \sigma^{-1} \) and \( \tau^{-1} \). The Galois group of \( \Psi^* \) over \( Q(w) \) has order 48 and consists of those permutations of the roots that permute the sets \( \{\rho, \rho^{-1}\}, \{\sigma, \sigma^{-1}\}, \{\tau, \tau^{-1}\} \) among themselves.

Now Theorem 3.5 shows that \( \sum E_1(2^{-n})(32w)^n \) is in a splitting field of \( \Psi^* \) over \( Q(w) \). But as we saw in [3], the field \( \mathcal{L} \) attached to the matrices \( A, B \) and \( C \)
sits inside a certain subfield of the splitting field of det $|xI_s-w(Ax^2+Bx+C)|$. In our case $\mathcal{L}$ is the degree 8 extension of $Q(w)$ corresponding to the subgroup of the Galois group that stabilizes the set $\{\rho, \sigma, \tau\}$.

Let $u_1 = w^{10}(\rho - \rho^{-1})(\sigma - \sigma^{-1})(\tau - \tau^{-1})$ and $u_2 = w^{10}(\rho \sigma \tau + \rho^{-1} \sigma^{-1} \tau^{-1})$. Using Galois theory we find that $u_1^2$ is in $Q(w)$, that $u_2$ has degree 4 over $Q(w)$, and that $u_1$ and $u_2$ generate the degree 8 extension of $Q(w)$ mentioned above.

So $\sum E_1(2^{-n})(32w)^n$ lies in $Q(w,u_1,u_2)$. A short calculation shows that $u_1^2 = (w^2 - 1)^2(w^2 + 1)^2((1-w^2)^2 - 4w^6)$. One can also write down an irreducible equation for $u_2$ over $Q(w)$ but it’s messy. (Some of the primes of $Q[w]$ that ramify in $Q(w,u_2)$ are $(1 - w^2 + 2w^3)$, $(1 - w^2 - 2w^3)$ and $(4 + 8w^2 - 4w^4 - 12w^6 - 23w^8 - 18w^{10} + 81w^{12} + 108w^{14})$). Now the only fields between $Q(w)$ and $Q(w,u_1,u_2)$ are $Q(w)$, $Q(w,u_1,u_2) = Q(w,\sqrt{(1-w^2)^2 - 4w^6})$, $Q(w,u_2)$ and $Q(w,u_1,u_2)$. So $\sum E_1(2^{-n})(32w)^n$, and consequently the conjectured Hilbert-Kunz series of $u^6 + u^3 v^3 + v^6 + x^3 + y^3 + xyz$, generates one of these 4 extensions of $Q(w)$. I think it generates the full degree 8 extension, but verifying this would be a very nasty computation.

Now consider the integral closure of $Q[w]$ in $Q(w,u_1,u_2)$. There is just one prime ideal in this ring lying over $(1 - 16w)$, and the argument of Theorem 1.14 shows that $\mu(t + E_1)$, the putative Hilbert-Kunz multiplicity of $u^6 + u^3 v^3 + v^6 + x^3 + y^3 + xyz$ lies in the residue class field of this ideal.

The residue-class field is a degree 8 extension of $Q$ generated by the images, $\bar{u}_1$ and $\bar{u}_2$ of $u_1$ and $u_2$. $Q(\bar{u}_1)$ is just $Q(\sqrt{(13)(157)(2039)})$, while $Q(\bar{u}_2)$ is a degree 4 extension of $Q$ with discriminant $2^2 \cdot 3^3 \cdot 5^2 \cdot 13^2 \cdot 17^2 \cdot 31 \cdot 157^2 \cdot 2039^2 \cdot 780854102129687$. The only subfields of $Q(\bar{u}_1,\bar{u}_2)$ are $Q$, $Q(\bar{u}_1)$, $Q(\bar{u}_2)$ and $Q(\bar{u}_1,\bar{u}_2)$. So $\mu(t + E_1)$ generates one of these 4 extensions of $Q$. My belief is that it generates the full degree 8 extension.

References

[1] C. Han, P. Monsky, Some surprising Hilbert-Kunz functions, Math. Z. 214 (1993), 119–135.
[2] P. Monsky, Rationality of Hilbert-Kunz multiplicities: a likely counterexample, Michigan Math. J. 57 (2008), 605–613.
[3] P. Monsky, Generating functions attached to some infinite matrices, Preprint (2009), arXiv:math.CO/0906.1836.
[4] P. Monsky, P. Teixeira, $p$-Fractals and power series I, J. Algebra 280 (2004), 505–536.
[5] P. Monsky, P. Teixeira, $p$-Fractals and power series II, J. Algebra 304 (2006), 237–255.