Nilpotence order growth of
recursion operators in characteristic *p*

Anna Medvedovsky

Nilpotence order growth of recursion operators in characteristic p

Anna Medvedovsky

We prove that the killing rate of certain degree-lowering "recursion operators" on a polynomial algebra over a finite field grows slower than linearly in the degree of the polynomial attacked. We also explain the motivating application: obtaining a lower bound for the Krull dimension of a local component of a big mod p Hecke algebra in the genus-zero case. We sketch the application for $p = 2$ and $p = 3$ in level one. The case $p = 2$ was first established in by Nicolas and Serre in 2012 using different methods.

1. Introduction

The main goal of this document is to prove the following *nilpotence growth theorem*, about the killing rate of a recursion operator on a polynomial algebra over a finite field under repeated application:

Theorem A (nilpotence growth theorem; see also Theorem 1). *Let $\mathbb{F}$ be a finite field of characteristic p, and suppose that $T : \mathbb{F}[y] \to \mathbb{F}[y]$ is a degree-lowering $\mathbb{F}$-linear operator satisfying the following condition*:

> *The sequence $\{T(y^n)\}_n$ of polynomials in $\mathbb{F}[y]$ satisfies a linear recursion over $\mathbb{F}[y]$ whose companion polynomial $X^d + a_1 X^{d-1} + \cdots + a_d \in \mathbb{F}[y][X]$ has both total degree d and y-degree d.*[i]

[i]The *companion polynomial* of a linear recurrence $s_n = a_1 s_{n-1} + \cdots + a_d s_{n-d}$ satisfied by a sequence $\{s_n\}_n$ for all $n \geq d$ is $X^d - a_1 X^{d-1} - \cdots - a_d$. See Section 2D.

Then there exists a constant $\alpha < 1$ *so that the minimum power of* T *that kills* y^n *is* $O(n^\alpha)$.

To prove this theorem, we reduce to the case where the companion polynomial of the recursion has an "empty middle" in its degree-d homogeneous part: that is, when for some $a \in \mathbb{F}$ it has the form $X^d + ay^d +$ (terms of total degree $< d$). Then we prove this empty-middle case (see Theorem 4 below) by constructing a function $c : \mathbb{F}[y] \to \mathbb{N} \cup \{-\infty\}$ that grows like $(\deg f)^\alpha$ and whose value is lowered by every application of T. In the special case where d is a power of p, the function c takes y^n to the integer obtained by writing n in base d and then reading the expansion in some smaller base, so that the sequence $\{c(y^n)\}_n$ is p-regular in the sense of Allouche and Shallit [1992]. The proof that $c(T(y^n)) < c(y^n)$, by strong induction, uses higher-order recurrences depending on n, so that n is compared to numbers whose base-d expansion is not too different.

It is the author's hope that ideas from p-automata theory can eventually be used to sharpen and generalize the nilpotence growth theorem.

Motivating application of the nilpotence growth theorem. The motivating application for the nilpotence growth theorem (Theorem A above) is the *nilpotence method* for establishing lower bounds on dimensions of local components of Hecke algebras acting on mod p modular forms of tame level N. These Hecke algebra components were first studied by Jochnowitz [1982] in the 1970s, but the first full structure theorem did not appear until over thirty years later. In 2012 Nicolas and Serre used recurrences satisfied by Hecke operators (see (5-1)) to describe the Hecke action on modular forms modulo 2 completely explicitly [Nicolas and Serre 2012a], leading to a Hecke algebra structure result for $p = 2$ and $N = 1$ [Nicolas and Serre 2012b]. Unfortunately their explicit formulas appear not to generalize beyond $p = 2$. The structure of mod p Hecke algebras for $p \geq 5$ was subsequently established by very different techniques by Bellaïche and Khare [2015] for $N = 1$ and later generalized by Deo [2017] to all N. The Bellaïche–Khare method deduces information about mod p Hecke algebra components from corresponding characteristic-zero Hecke algebra components, which are known to be big by the Gouvêa–Mazur "infinite fern" construction ([Gouvêa and Mazur 1998]; see also [Emerton 2011, Corollary 2.28]). The nilpotence method is yet a third technique, coming out of an idea of Bellaïche for tackling the case $p = 3$ and $N = 1$ as outlined in [Bellaïche and Khare 2015, Appendix], and implemented and developed in level one for $p = 2, 3, 5, 7, 13$ in the present author's Ph.D. dissertation [Medvedovsky 2015]. Like the Nicolas–Serre approach, the nilpotence method stays entirely in characteristic p and makes use of Hecke recurrences; but instead of explicit Hecke action formulas, the nilpotence growth theorem now plays the crucial dimension-bounding role. See Section 5 below for a taste of this method for $p = 2, 3$, which completes the determination of the structure of the Hecke algebra for $p = 3$ begun in [Bellaïche and Khare 2015, Appendix] and recovers the Nicolas–Serre result for $p = 2$. In fact, the nilpotence method via the nilpotence growth theorem in its current form gives lower bounds on dimensions of mod p Hecke algebras of level N so long as the genus of the modular curve $X_0(Np)$ is zero; see [Medvedovsky 2015] and the forthcoming [Medvedovsky $\geq$ 2018] for details.

Structure of this document. After a few preliminary definitions in Section 2, we state and discuss a more general version of the nilpotence growth theorem (NGT); see Theorem 1 in Section 3. In Section 4, we prove a toy version of the NGT (Theorem 2). In Section 5, we use the toy version of NGT to prove that the mod p level-one Hecke algebra for $p = 2, 3$ has the form $\mathbb{F}_p[\![x, y]\!]$. This section illustrates the motivating application of the nilpotence growth theorem and is not required for the rest of the document. This is a reasonable stopping point for a first reading.

In Section 6 the proof of the NGT begins in earnest. There is a short overview of the structure of the proof in Section 6A. In Section 6B, we reduce to working over a finite field. In Section 6C, we reduce to the empty-middle NGT (Theorem 4). In Section 6D, we give the inductive argument that reduces the proof of the empty-middle NGT to finding a *nilgrowth witness* function satisfying certain properties. The next three sections are combinatorial in nature, as we construct a nilgrowth witness. In Section 7, we discuss base-b representation of numbers and introduce the *content* function. In Section 8, we prove a number of technical inequalities about the content function. In Section 9 we finally construct a nilgrowth witness out of the content function, finishing the proof of the empty-middle NGT, and hence of the NGT in full. Finally in Section 10, we state a more precise version of the toy NGT and speculate on the optimality of some bounds.

2. Preliminaries

This section contains a brief review of a few unconnected algebraic notions. All rings and algebras are assumed to be commutative, with unity. We use the convention that the set of natural numbers starts with zero: $\mathbb{N} = \{0, 1, 2, \ldots\}$. Below, R is always a ring.

2A. ***Structure of finite rings.*** If R is finite, then R is artinian, hence a finite product of finite local rings. If R is a finite local ring with maximal ideal $\mathfrak{m}$, then the residue field $R/\mathfrak{m}$ is a finite field of characteristic p. Moreover, the graded pieces $\mathfrak{m}^n/\mathfrak{m}^{n+1}$ are finite $R/\mathfrak{m}$-vector spaces, so that R has cardinality a power of p. Basic examples of finite local rings include $\mathbb{F}_p[t]/(t^k)$ and $\mathbb{Z}/p^k\mathbb{Z}$.

2B. ***Degree filtration on a polynomial algebra.*** If $0 \neq f = \sum_{n\geq 0} c_n y^n$ is a polynomial in $R[y]$, then its *y-degree*, or just *degree*, is as usual defined to be $\deg f := \max\{n : c_n \neq 0\}$. For $f = 0$, set $\deg f := -\infty$.

The degree function gives $R[y]$ the structure of a *filtered algebra*. Let $R[y]_n := \{f \in R[y] : \deg f \leq n\}$, and then $R[y] = \bigcup_{n\geq 0} R[y]_n$ and multiplication preserves the filtration as required.

2C. ***Local nilpotence and the nilpotence index.*** Let M be any R-module and $T \in \mathrm{End}_R(M)$ an R-linear endomorphism. (In applications to Hecke algebras, R will be a finite field, M an infinite-dimensional R-algebra of modular forms, and T a Hecke operator.)

The operator $T : M \to M$ is *locally nilpotent* on M if every element of M is annihilated by some power of T. If T is locally nilpotent and f in M is nonzero, we define the *nilpotence index of f with respect to T* as

$$N_T(f) := \max\{k \geq 0 : T^k f \neq 0\}.$$

Also set $N_T(0) := -\infty$.

Suppose $R = K$ is a field, $M = K[y]$, and $T : M \to M$ preserves the degree filtration; that is, $T(K[y]_n) \subset K[y]_n$. Then T is locally nilpotent if and only if T strictly lowers degrees, in which case we also have $N_T(f) \leq \deg f$.

For example, $T = \frac{d}{dy}$ is locally nilpotent on $K[y]$. If K has characteristic zero, then $N_T(f) = \deg f$; otherwise $N_T(f) \leq \operatorname{char} K - 1$.

2D. *Linear recurrences and companion polynomials.*

Now suppose that M is an R-algebra, and M' is an M-module (we will usually take $M' = M$). A sequence $s = \{s_n\} \in M'^{\mathbb{N}}$ satisfies an M-*linear recurrence of order d* if there exist elements $a_0, a_1, \ldots, a_d \in M$ so that

$$a_0 s_n = a_1 s_{n-1} + \cdots + a_d s_{n-d} \quad \text{for all } n \geq d. \tag{2-1}$$

Unlike some authors, we do not assume that a_d is nonzero or not a zero divisor, but we do insist that the recursion already hold for $n = d$. The *companion polynomial* of this linear recurrence is $P(X) = a_0 X^d - a_1 X^{d-1} - \cdots - a_d \in M[X]$. If $a_0 = 1$, then the recurrence is said to be *monic*; we will always assume below that our linear recurrences are monic unless stated otherwise.

Example. The sequence $s = \{0, 1, y, y^2, y^3, y^4, \ldots\} \in R[y]^{\mathbb{N}}$ satisfies an $R[y]$-linear recursion of minimal order 2; we have $s_n = y s_{n-1}$ for all $n \geq 2$, but not for $n = 1$. The companion polynomial of the recurrence is therefore $X^2 - yX$.

Given any sequence s in $M'^{\mathbb{N}}$, the set of companion polynomials of (not necessarily monic) M-linear recurrences satisfied by s forms an ideal of $M[X]$. We record this observation in the following form:

Fact. If a sequence $s \in M'^{\mathbb{N}}$ satisfies the recurrence defined by some monic $P \in M[X]$, then it also satisfies the recurrence defined by PQ for any other monic $Q \in M[X]$.

In characteristic p we get the following corollary, of which we will make crucial use:

Corollary 2.1. *If R has characteristic p and $s \in M'^{\mathbb{N}}$ satisfies the order-d recurrence*

$$s_n = a_1 s_{n-1} + a_2 s_{n-2} + \cdots + a_d s_{n-d} \quad \text{for all } n \geq d,$$

then for every $k \geq 0$ the sequence s also satisfies the order-dp^k deeper recurrence

$$s_n = a_1^{p^k} s_{n-p^k} + a_2^{p^k} s_{n-2p^k} + \cdots + a_d^{p^k} s_{n-dp^k} \quad \text{for all } n \geq dp^k. \tag{2-2}$$

Proof. Let $P = X^d - a_1 X^{d-1} - \cdots - a_d$ be the companion polynomial of a recursion satisfied by s. By the fact above, the sequence s also satisfies the recurrence whose companion polynomial is

$$P^{p^k} = X^{dp^k} - a_1^{p^k} X^{dp^k - p^k} - a_2^{p^k} X^{dp^k - 2p^k} - \cdots - a_d^{p^k},$$

which is exactly what is expressed in (2-2). $\qquad \square$

If M is a domain embedded into a field K, we have the following well-known characterization of power sequences in $\overline{K}^{\mathbb{N}}$ satisfying a fixed M-linear recurrence:

Fact. An element α in $\overline{K}$ is a root of monic $P \in M[X]$ if and only if the sequence $\{\alpha^n\}_n = \{1, \alpha, \alpha^2, \ldots\}$ satisfies the linear recurrence with companion polynomial P.

If the companion polynomial of such an M-linear recurrence has no repeated roots in $\overline{K}$, it follows from the proposition that *every* solution to the recurrence is a linear combination of such power sequences on the roots of the companion polynomial. One can further describe all $\overline{K}$-sequences satisfying a general M-linear recursion — see, for example, [Conrad 2016], particularly the historical references on page 2 — but we will not need this below.

3. The nilpotence growth theorem (NGT)

3A. *Statement of the NGT.* We are now ready to state the most general version of the nilpotence growth theorem (NGT). From now on, we will assume R to be a *finite* ring, and $M = R[y]$.

Theorem 1 (nilpotence growth theorem). *Let R be a finite ring, and suppose that $T : R[y] \to R[y]$ is an R-linear operator satisfying the following two conditions*:

(1) *T lowers degrees*: $\deg T(f) < \deg f$ *for every nonzero f in $R[y]$.*

(2) *The sequence $\{T(y^n)\}_n$ satisfies a* filtered *linear recursion over $R[y]$: there exist $a_1, \ldots, a_d \in R[y]$, with $\deg a_i \leq i$ for each i, so that for all $n \geq d$,*

$$T(y^n) = a_1 T(y^{n-1}) + \cdots + a_d T(y^{n-d}).$$

Suppose further that

(3) *the coefficient of y^d in a_d is invertible in R.*

Then there exists a constant $\alpha < 1$ so that $N_T(y^n) \ll n^\alpha$.

In other words, Theorem 1 implies that, under a mild technical assumption (condition (3)), the nilpotence index of a degree-lowering operator defined by a filtered linear recursion grows *slower than linearly* in the degree. The mild technical assumption is necessary in the theorem as stated; see the discussion in (4) in Section 3B below.

3B. *Discussion of the NGT.* (1) *Connection with Theorem A*: If $T : R[y] \to R[y]$ satisfies the conditions of Theorem 1, then the companion polynomial of the recursion satisfied by the sequence $\{T(y^n)\}_n$ is

$$P_T = X^d - a_1 X^{d-1} - \cdots - a_d \in R[y, X].$$

The condition $\deg a_i \leq i$ from (2) guarantees that the total degree of P_T is exactly d. In particular, in the case where $R = \mathbb{F}$ is a finite field, condition (3) implies that $\deg_y P_T = \deg a_d = d$. In other words, Theorem 1 over a finite field reduces to Theorem A.

(2) *Condition (1) guarantees that T is locally nilpotent*: Moreover, $N_T(y^n) \leq n$, so that the function $n \mapsto N_T(y^n)$ a priori grows no faster than linearly.

(3) *Condition (2) and connection to recursion operators*: The condition that the sequence $\{T(y^n)\}_n$ satisfies a linear recurrence is the definition of a *recursion operator*, a notion that will be explored in a future paper. A natural source of *filtered* recursion operators (that is, satisfying additional degree bounds as in condition (2) above) comes from the action of Hecke operators on algebras of modular forms of a fixed level. Namely, *if f is a modular form of weight k and level N and T is a Hecke operator acting on the algebra M of forms of level N, then the sequence $\{T(f^n)\}_n$ satisfies an M-linear recursion with companion polynomial $X^d + a_1 X^{d-1} + \cdots + a_d$, where a_i comes from weight ki.*

See (5-1) and (5-2) below for examples over $\mathbb{F}_p$, [Medvedovsky 2015, chapter 6] for a proof in level one when T is a prime Hecke operator, or [Medvedovsky $\geq$ 2018] for more details.

The history of Hecke recurrences appears to be relatively brief. Hecke recurrences over $\mathbb{F}_2$ were crucially used by Nicolas and Serre to obtain the structure of the mod 2 Hecke algebra in level one [Nicolas and Serre 2012a; 2012b]. Earlier, Al Hajj Shehadeh, Jaafar, and Khuri-Makdisi [2009] had investigated two-dimensional Hecke recurrences over $\mathbb{Q}$ satisfied by the array $\{T_\ell(E_4^n E_6^m)\}_{n,m}$; Buzzard and Calegari [2005, p. 594] had used Hecke recurrences for U_2 acting on a power basis of overconvergent 2-adic modular functions in their study of slopes.

(4) *Condition (3) is necessary as stated*: Consider the operator $T : R[y] \to R[y]$ defined by $T(y^n) = s_n$, where $\{s_n\}$ is the sequence $\{0, 1, y, y^2, \ldots\}$ with companion polynomial $X^2 - yX$ from the example on page 696. All conditions except (3) are satisfied, and it is easy to see that $N_T(y^n) = n$ in this case.

For an example with $a_d \neq 0$, consider the operator T with the defining companion polynomial $P_T = X^2 + yX + y$ and initial values $[T(1), T(y)] = [0, 1]$. By induction, $\deg T(y^n) = n - 1$. Therefore $N_T(y^n) = n$.

Computationally, it appears that if $R = \mathbb{F}_p$ and $\deg a_d < d$ but there exists an i with $0 < i < d$ so that $\deg a_i = i$, then either N_T grows logarithmically or else it grows linearly. In that sense, it appears that "fullness" of degree at the end of P_T (that is, the presence of a y^d term) appears to be, at least generically, necessary to compensate for "fullness" of degree in the middle (that is, the presence of a $y^i X^{d-i}$ term for some $0 < i < d$), if one wants the growth of N_T to be sublinear but not degenerate. But the phenomenon is not well understood.

(5) *The constant α*: The power α depends on R and d only, and tends to 1 as $d \to \infty$. More precisely, the dependence on R is only through its maximal residue characteristic; the length of R as a module over itself affects only the implicit constant of the growth condition $N_T(y^n) \ll n^\alpha$. If the inequality $\deg a_i < i$ is strict for every $i < d$ ("empty middle" case) we can take α to be $\log_{p^k}(p^k - 1)$ for k satisfying $d \leq p^k$. See Theorems 2 or 4 below.

(6) *Finite characteristic is necessary — a counterexample in characteristic zero*: Consider the operator T on $\mathbb{Q}[y]$ with $P_T = X^2 - yX - y^2$ and degree-lowering initial terms $[T(1), T(y)] = [0, 1]$. This satisfies all three conditions of the NGT. It is easy to see that $T(y^n) = F_n y^{n-1}$, where F_n is the n-th Fibonacci number; the recursion is $s_n = y s_{n-1} + y^2 s_{n-2}$. Therefore

$$T^k(y^n) = F_n F_{n-1} \cdots F_{n-k+1} y^{n-k},$$

so that $N_T(y^n) = n$. (Compare to characteristic p, where the operator defined by $T(y^n) = F_n y^{n-1}$ on $\mathbb{F}_p[y]$ satisfies $T^{p+1} \equiv 0$.) See also Proposition 10.1 for a family of examples in any degree.

Computationally, it appears that generic characteristic-zero examples that do not degenerate (to $\log n$ growth) all exhibit linear growth. In contrast, over a finite field, one observes $O(n^\alpha)$ growth for various $\alpha < 1$.

(7) *Finiteness of R is necessary as stated — a counterexample over $\mathbb{F}_p(t)$, due to Paul Monsky*: Let $P_T = X^2 - tyX - y^2$ and start with $[0, 1]$ again. Then $T(y^n) = F_n(t)y^{n-1}$ with $F_n(t) \in \mathbb{F}_p[t]$ monic of degree $n - 1$, so that $N_T(y^n) = n$ again. However, see the empty-middle case (Theorem 4) for a special case that does hold for infinite rings of characteristic p.

4. A toy case of the NGT

Fix a prime p and take $R = \mathbb{F}_p$ for simplicity (in fact any ring of characteristic p works here). We prove the following special case of the NGT for recurrences with empty middle (i.e., whose companion polynomials have no maximal-degree cross terms) and whose order is a power of p.

Theorem 2 (toy case of NGT). *Let $q = p^k$ for some $k \geq 1$. Suppose $T : \mathbb{F}_p[y] \to \mathbb{F}_p[y]$ is a degree-lowering linear operator so that the sequence $\{T(y^n)\}_n$ satisfies an $\mathbb{F}_p[y]$-linear recursion with companion polynomial*

$$P = X^q + (\textit{terms of total degree} < q) + ay^q \in \mathbb{F}_p[y][X]$$

for some $a \in \mathbb{F}_p$. Then $N_T(y^n) \ll n^{\log(q-1)/\log q}$.

Most of the main features of the proof of the NGT (Theorem 1) are already present in the proof of this toy case. We include the toy case here because the proof is technically much simpler; understanding it may suffice for all but the most curious readers.

4A. *The content function.* For $q = 3$, following Bellaïche (see the appendix of [Bellaïche and Khare 2015]), we define a function $c : \mathbb{N} \to \mathbb{N}$ depending on q as follows. Given an integer n, we write it in base q as $n = \sum_i n_i q^i$ with $0 \leq n_i < q$, only finitely many of which are nonzero, and define the q-*content of n* as $c(n) := c_q(n) := \sum_i n_i (q-1)^i$. For example, since $71 = [2\,4\,1]_5$ in base 5, the 5-content of 71 is $2 \cdot 4^2 + 4 \cdot 4 + 1 = 49$.

The following properties of the content function are easy to check. See also Section 7A, where the content function and variations are discussed in detail.

Proposition 4.1. (1) $c(n) \ll n^{\log_q(q-1)}$.

(2) $c(q^k n) = (q-1)^k c(n)$ for all $k \geq 0$.

(3) If $0 \leq n < q$, then $c(n) = n$.

(4) If i is a digit base q and $n \geq i$ has no more than 2 digits base q, then $c(n) - c(n-i)$ is either i or $i - 1$.

(5) If $q \leq n < q^2$, then $c(n-q) = c(n) - q + 1$.

4B. *Setup of the proof.* We now define the q-content of a polynomial $f \in \mathbb{F}_p[y]$ through the q-content of its degree. More precisely, if $0 \neq f = \sum a_n y^n$, let $\tilde{c}(f) := \max\{c(n) : a_n \neq 0\}$. Also set $\tilde{c}(0) := -\infty$. For example, the 3-content of $2y^9 + y^7 + y^2$ is $\max\{c(9), c(7), c(2)\} = 5$.

Now let $T : \mathbb{F}_p[y] \to \mathbb{F}_p[y]$ be a degree-lowering recursion operator whose companion polynomial

$$P = X^q + a_1(y)X^{q-1} + \cdots + a_q(y) \in \mathbb{F}_p[y][X]$$

satisfies $\deg a_i(y) < i$ for $1 \leq i < q$ and $\deg a_q(y) \leq q$.

To prove Theorem 2, we will show that T lowers the q-content of any $f \in \mathbb{F}_p[y]$; that is, that $\tilde{c}(Tf) < \tilde{c}(f)$. Since $\tilde{c}(f) < 0$ only if $f = 0$, the fact that T lowers q-content implies that $N_T(f) \leq c(f)$. Proposition 4.1(1) will then imply $N_T(f) \ll (\deg f)^{\log(q-1)/\log q}$, as desired.

It suffices to prove that $\tilde{c}(Tf) < \tilde{c}(f)$ for $f = y^n$. We will proceed by strong induction on n, each time using a deeper recursion of order q^{k+1} corresponding to P^{q^k}, with k chosen so that $q^{k+1} \leq n < q^{k+2}$. We learned this technique from Gerbelli-Gauthier's proof [2016] of the key technical lemmas of Nicolas and Serre [2012a]. Using deeper recurrences with induction allows us to compare n to $n - iq^k$, which has the same last k digits base q, rather than to $n - i$, whose base-q expansion may look very different.

4C. *The induction.* The base case is $n < q$, in which case being q-content-lowering is the same thing as being degree-lowering (Proposition 4.1(3)).

For $n \geq q$, we must show that $\tilde{c}(T(y^n)) < c(n)$ assuming that $\tilde{c}(T(y^m)) < c(m)$ for all $m < n$. As above, choose $k \geq 0$ with $q^{k+1} \leq n < q^{k+2}$. By Corollary 2.1, the sequence $\{T(y^n)\}$ satisfies the order-q^{k+1} recurrence

$$T(y^n) = a_1(y)^{q^k} T(y^{n-q^k}) + a_2(y)^{q^k} T(y^{n-2q^k}) + \cdots + a_q(y)^{q^k} T(y^{n-q^{k+1}}).$$

Pick a term y^m appearing in $T(y^n)$ with nonzero coefficient; we want to show that $c(m) < c(n)$. From the recursion, y^m appears with nonzero coefficient in $a_i(y)^{q^k} T(y^{n-iq^k})$ for some i. More precisely, y^m appears in $y^{jq^k} T(y^{n-iq^k})$ for some y^j appearing in $a_i(y)$, so that either $j < i$ or $i = j = q$. Then y^{m-jq^k} appears in $T(y^{n-iq^k})$, and by induction we know that $c(m - jq^k) < c(n - iq^k)$. To conclude that $c(m) < c(n)$, it would suffice to show that

$$c(n) - c(m) \geq c(n - iq^k) - c(m - jq^k),$$

or, equivalently, that

$$c(n) - c(n - iq^k) \geq c(m) - c(m - jq^k).$$

Since subtracting multiples of q^k leaves the last k digits of n base q untouched, we may replace n and m by $q^k \lfloor \frac{n}{q^k} \rfloor$ and $q^k \lfloor \frac{m}{q^k} \rfloor$, respectively, and then use Proposition 4.1(2) to cancel out a factor of $(q-1)^k$. In other words, we must show that

$$c(n) - c(n - i) \geq c(m) - c(m - j)$$

for n, m, i and j satisfying $i \le n < q^2$ and $j \le m < n$ and either $j < i$ or $i = j = q$. But this is an easy consequence of Proposition 4.1(4)–(5): For $j < i$, we know that $c(n) - c(n-i)$ is at least $i - 1$ and $c(m) - c(m-j)$ is at most $j \le i - 1$. And for $i = j = q$ both sides equal $q - 1$.

This completes the proof of Theorem 2.

4D. *Toy case versus general case.* The proof of the full NGT (Theorem 1) proceeds by first reducing to the empty-middle case over a finite field (Theorem 4 below), of which Theorem 2 is a special case where the order of the recursion is a power of p. Apart from the reduction step, most of the difficulty in generalizing from Theorem 2 to Theorem 4 comes from working with more general versions of the content function to accommodate any recursion order. Namely, we will extend the content function to rational numbers and prove sufficiently strong analogues of Proposition 4.1 for the induction to proceed, see Sections 7–9.

5. Applications to mod p Hecke algebras

This section gives an indication of how the NGT can give information about lower bounds of mod p Hecke algebras, the author's main motivation for proving the theorem. More precisely, in this section we will use Theorem 2 to complete the proof of Theorem 24 of [Bellaïche and Khare 2015, Appendix], which establishes the structure of the mod 3 Hecke algebra of level one.[(ii)] Simultaneously and using the same methods, we will give an alternate proof of the main result of Nicolas and Serre [2012b], the structure of the mod 2 Hecke algebra in level one. See Theorem 3 below.

More generally, the NGT can be used to obtain lower bounds on Krull dimensions of local components of big mod p Hecke algebras acting on forms of level N in the case where $X_0(Np)$ has genus zero, for this is precisely the condition for the algebra of modular forms of level N mod p to be a polynomial algebra over $\mathbb{F}_p$. For more details, see [Medvedovsky 2015] (for level one) or [Medvedovsky $\ge$ 2018]. To generalize the nilpotence method to all (p, N), one must generalize the NGT to all rings of S-integers in characteristic-p global fields, with the max-pole-order filtration generalizing degree.

We work in level one with $p \in \{2, 3\}$. Let $M = M(1, \mathbb{F}_p) \subset \mathbb{F}_p[\![q]\!]$ be the space of modular forms of level one modulo p in the sense of Swinnerton-Dyer and Serre (that is, reductions of integral q-expansions). For $p = 2, 3$ Swinnerton-Dyer observes [1973] that $M = \mathbb{F}_p[\Delta]$, where $\Delta = \prod_{i=1}^{n}(1 - q^n)^{24} \in \mathbb{F}_p[\![q]\!]$. Standard dimension formulas show that $M_k := \mathbb{F}_p[\Delta]_k$, the polynomials in Δ of degree bounded by k, coincides with the space of mod p reductions of q-expansions of forms of weight $12k$, and hence is Hecke invariant. Further, one can show that $K := \langle \Delta^n : p \nmid n \rangle_{\mathbb{F}_p} \subset M$ is the kernel of the operator U_p, which implies that K and the finite-dimensional subspaces $K_k := M_k \cap K$ are all Hecke invariant.

Let $A_k \subset \mathrm{End}_{\mathbb{F}_p}(K_k)$ be the algebra generated by the action of the Hecke operators T_ℓ with ℓ prime and $\ell \ne p$. Since $K_k \hookrightarrow K_{k+1}$, we have $A_{k+1} \twoheadrightarrow A_k$. Let $A := \varprojlim_k A_k$. Then A is a profinite ring

---

[(ii)]More precisely, we prove a weaker version of [Bellaïche and Khare 2015, Proposition 35]. In the notation of Section 7B here, we show for $f \in \mathbb{F}_3[\Delta]$ that $c_{3,2}(T_2 f) \le c_{3,2}(f) - 1$ and $c_{9,6}(T_7' f) \le c_{9,6}(f) - 3$. This suffices to complete the proof of [loc. cit., Theorem 24], but we do not prove the stronger claim that $c_{3,2}(T_7' f) \le c_{3,2}(f) - 2$.

embedding into $\mathrm{End}(K)$; it is the shallow Hecke algebra acting on forms of level one mod p. The standard pairing $A \times K \to \mathbb{F}_p$ given by $\langle T, f \rangle \mapsto a_1(Tf)$ is nondegenerate on both sides and continuous in the profinite topology on A. Therefore A is in continuous duality with K. By work of Tate [1994] and Serre [1986, p. 710, Note 229.2] we know that Δ is the only Hecke eigenform in $K \otimes \bar{\mathbb{F}}_p$.[(iii)] This implies that A is a local $\mathbb{F}_p$-algebra with maximal ideal $\mathfrak{m}$ and residue field $\mathbb{F}_p$ generated by the modified Hecke operators $T'_\ell := T_\ell - a_\ell(\Delta)$, acting locally nilpotently on $M = \mathbb{F}_p[\Delta]$. Using deformation theory of Chenevier pseudorepresentations, one can deduce:

Proposition 5.1. *There is a surjection* $\mathbb{F}_p[\![x, y]\!] \twoheadrightarrow A$ *given by* $\begin{cases} x \mapsto T_3, \ y \mapsto T_5 & \text{if } p = 2, \\ x \mapsto T_2, \ y \mapsto T'_7 & \text{if } p = 3. \end{cases}$

For $p = 2$, the fact that A is generated by T_3 and T_5 was first proved without deformation theory in [Nicolas and Serre 2012a]. For $p = 3$, Proposition 5.1 is stated [Bellaïche and Khare 2015, Appendix], using deformation theory of reducible Rouquier pseudocharacters developed in [Bellaïche 2012]. See [Medvedovsky 2015, Chapter 7] for detailed computations of tangent spaces to reducible local components of mod p Hecke algebras in level one.

The main result of this section is the following:

Theorem 3. *The surjection* $\mathbb{F}_p[\![x, y]\!] \twoheadrightarrow A$ *of Proposition 5.1 is an isomorphism.*

The key input will be Theorem 2, as well as the following observation: if T is any Hecke operator and f is a modular form in a Hecke invariant algebra M, then the sequence $\{T(f^n)\}_n$ satisfies an M-linear recursion. For more details on the Hecke recursion, see [Medvedovsky 2015, Chapter 6] or [Medvedovsky $\geq$ 2018], but here we will only need some special cases for $f = \Delta$ already given in [Nicolas and Serre 2012a; Bellaïche and Khare 2015, Appendix]. For $p = 2$, we have, as in [Nicolas and Serre 2012a, (13)–(14)],

$$\begin{aligned} T_3(\Delta^n) &= \Delta\, T_3(\Delta^{n-3}) + \Delta^4\, T_3(\Delta^{n-4}), \quad n \geq 3, \\ T_5(\Delta^n) &= \Delta^2\, T_5(\Delta^{n-2}) + \Delta^4\, T_5(\Delta^{n-4}) + \Delta\, T_5(\Delta^{n-5}) + \Delta^6\, T_5(\Delta^{n-6}), \quad n \geq 6, \end{aligned} \tag{5-1}$$

with companion polynomials $P_3 = X^4 + \Delta X + \Delta^4$ and $P_5 = X^6 + \Delta^2 X^4 + \Delta^4 X^2 + \Delta X + \Delta^6$. Note that $\{T_5(\Delta^n)\}_n$ also satisfies the recursion defined by $P_5^* = P_5(X^2 + \Delta^2) = X^8 + \Delta X^3 + \Delta^3 X + \Delta^8$. And for $p = 3$, the recursions satisfied by $\{T_2(\Delta^n)\}$ and $\{T'_7(\Delta^n)\}$ have companion polynomials

$$P_2 = X^3 - \Delta X + \Delta^3 \quad \text{and} \quad P'_7 = X^9 - \Delta X^5 - \Delta^2 X^4 + (\Delta^4 - \Delta)X^2 + (\Delta^5 + \Delta^2)X - \Delta^9. \tag{5-2}$$

See Lemma 33 in [Bellaïche and Khare 2015, Appendix].[(iv)]

Proof of Theorem 3. Let T and S be the generators of A from Proposition 5.1. Then T, S are filtered and degree-lowering recursion operators on $\mathbb{F}_p[\Delta]$, each satisfying the conditions of Theorem 2. In other words, there exists an $\alpha < 1$ so that $N(\Delta^n) := N_T(\Delta^n) + N_S(\Delta^n) \ll n^\alpha$.

[(iii)] Alternatively, one can use an observation of Serre to conclude that any Hecke eigenform in $K \otimes \bar{\mathbb{F}}_p$ is in fact defined over $\mathbb{F}_p$, reducing the eigenform search to a finite computation. See [Bellaïche and Khare 2015, Section 1.2 footnote].

[(iv)] Lemma 33 in [Bellaïche and Khare 2015, Appendix] gives the degree-8 recurrence satisfied by $\{T_7(\Delta^n)\}_n$; the recurrence satisfied by $\{\Delta^n + T_7(\Delta^n)\}_n$ has an extra factor of $X - \Delta$.

We now claim that the Hilbert–Samuel function of A grows *faster than linearly*, so that the Krull dimension of A is at least 2. Indeed, the Hilbert–Samuel function of A sends a positive integer k to

$$\dim_{\mathbb{F}_p} A/\mathfrak{m}^k = \dim_{\mathbb{F}_p} K[\mathfrak{m}^k] \geq \#\{n : \Delta^n \in K, \ \mathfrak{m}^k \Delta^n = 0\} = \#\{n \text{ prime to } p : N(\Delta^n) \geq k\} \gg k^{1/\alpha},$$

which is certainly faster than linear, since $\frac{1}{\alpha} > 1$. Therefore it grows at least quadratically, and the Krull dimension of A is at least 2. By the Hauptidealsatz, the kernel of the surjection $\mathbb{F}_p[\![x, y]\!] \twoheadrightarrow A$ from Proposition 5.1 is trivial. $\qquad\square$

Using the more precise bounds on α from Theorem 4, we can conclude that, for $p = 2$ we have $\alpha = \max\{\log_4 2, \log_8 4\} = \frac{2}{3}$ and for $p = 3$ we have $\alpha = \max\{\log_3 2, \log_9 6\} \approx 0.815$. Compare to $\alpha = \frac{1}{2}$ obtained for $p = 2$ by Nicolas and Serre [2012a, §4.1]. Computations suggest that $\alpha = \frac{1}{2}$ also holds for $p = 3$, but we have not been able to prove this.

6. The proof of the NGT begins

We now begin the proof of Theorem 1.

6A. *Overview of the proof.* The proof proceeds as follows.

(1) *Reduce the NGT to the case where R is a finite field*: See Section 6B below.

(2) *Reduce to the empty-middle case*: The NGT over a finite field is implied by a special empty-middle case (Theorem 4), where the companion polynomial has no terms of maximal total degree except for X^d and y^d (i.e., the highest-degree homogeneous part has an empty middle). Note that Theorem 4 holds over any ring of characteristic p. See Section 6C below for the statement of Theorem 4 and the reduction step.

(3) *Prove Theorem 4*: The main idea of the proof is as follows. Given an operator T satisfying the conditions of Theorem 4, we define a function $c_T : \mathbb{N} \to \mathbb{N}$ that grows like n^α for some $\alpha < 1$. We extend this function to polynomials in $R[y]$ via the degree. Finally, we use strong induction to prove that applying T strictly lowers the c_T-value of any polynomial in $R[y]$. Therefore, $N_T(y^n)$ is bounded by $c_T(n) \asymp n^\alpha$.

The key features of this kind of proof are already present in the proof of Theorem 2 in Section 4.

6B. *Reduction to the case where R is a finite field.*

Proposition 6.1. *If Theorem 1 is true whenever R is a finite field, then Theorem 1 is true.*

Proof. First, suppose R is a finite artinian local ring with maximal ideal $\mathfrak{m}$ and finite residue field $\mathbb{F}$. Let ℓ be the least positive integer so that $\mathfrak{m}^\ell = 0$.

Let $T : R[y] \to R[y]$ be the operator in the statement of the theorem, and write $\overline{T} : \mathbb{F}[y] \to \mathbb{F}[y]$ for the operator obtained by tensoring with the quotient map $R \twoheadrightarrow \mathbb{F}$. Theorem 1 for $\mathbb{F}$ guarantees that $N_{\overline{T}}(y^n) \ll n^\alpha$ for some $\alpha < 1$. Let

$$g(n) := \max_{n' \leq n}\{N_{\overline{T}}(y^{n'}) + 1\} \ll n^\alpha,$$

so that g is nondecreasing, integer-valued, and satisfies $\overline{T}^{g(\deg f)} f = 0$ for every f in $\mathbb{F}[y]$. Lifting back to R, we get that $T^{g(\deg f)} f$ is in $\mathfrak{m}[y]$ for all $f \in R[y]$.[V] More generally, if f is in $\mathfrak{m}^i[y]$, then $T^{g(\deg f)}$ sends f to $\mathfrak{m}^{i+1}[y]$. Since $\mathfrak{m}^\ell = 0$, we have $T^{\ell g(\deg f)} f = 0$ for every $f \in R[y]$, so that $N_T(y^n) \leq \ell g(n) - 1 \ll n^\alpha$.

In the general case, R is a finite product of finite artinian local rings R_i, and an R-linear operator $T : R[y] \to R[y]$ decomposes as $\sum T_i$ where $T_i : R_i[y] \to R_i[y]$ is the R_i-linear restriction of T to R_i. From the paragraph above, we can choose $\alpha_i < 1$ so that $N_{T_i}(y^n) \ll n^{\alpha_i}$ for all n. Then $\alpha = \max_i \{\alpha_i\}$ works for T. $\square$

6C. Reduction to the empty-middle NGT. From now on we fix a prime p. Theorem 1 over a finite field of characteristic p is implied by the following special case in which the shape of the recursion satisfied by T is restricted. Note that the statement below has no finiteness restrictions on the base ring, and no restriction on the coefficient of y^d.

Theorem 4 (empty-middle NGT). *Let R be a ring of characteristic p, and suppose that T is a degree-lowering linear operator on $R[y]$ so that the sequence $\{T(y^n)\}_n$ satisfies a linear recursion whose companion polynomial has the shape*

$$X^d + ay^d + (\text{terms of total degree} \leq d - D)$$

for some $D \geq 1$ and some constant $a \in R$. Let $b \geq d$ be a power of p, and suppose that either $b - d \leq 1$ or that $D \leq \frac{b}{2}$. Then

$$N_T(y^n) \ll n^\alpha \quad \text{for } \alpha = \frac{\log(b - D)}{\log b}.$$

The case $d = b$ and $D = 1$ has already been established in Theorem 2; an analogous argument extends to $d = b$ and any D with $1 \leq D \leq b - 1$.

Proposition 6.2 (empty-middle NGT implies NGT). *Theorem 4 implies Theorem 1.*

Proof. By Proposition 6.1, we may assume that we are working over a finite field $\mathbb{F}$. Let

$$P = X^d + a_1 X^{d-1} + \cdots + a_d \in \mathbb{F}[y][X]$$

be the filtered recursion satisfied by the sequence $\{T(y^n)\}_n$ as in the setup of Theorem 1; recall that we insist that $\deg a_d = d$. We will show that P divides a polynomial of the form

$$X^e - y^e + (\text{terms of total degree} < e)$$

for $e = q^m(q - 1)$, where q is a power of p and $m \geq 0$. Then the sequence $\{T(y^n)\}_n$ will also satisfy the recursion associated to a polynomial whose shape fits the requirements of Theorem 4.

Let H be the degree-d homogeneous part of P, so that $P = H + (\text{terms of total degree} < d)$. We claim that there exists a homogeneous polynomial $S \in \mathbb{F}[y, X]$ so that $H \cdot S = X^e - y^e$ for some positive

---

[V] If $\mathfrak{a} \subset R$ is an ideal, write $\mathfrak{a}[y] \subset R[y]$ for the ideal of polynomials all of whose coefficients are in $\mathfrak{a}$.

integer e of required form. Once we find such an S, we know that $P \cdot S$ will have the desired shape $X^e - y^e + (\text{terms of total degree} < e)$.

To find S, we dehomogenize the problem by setting $y = 1$. Let $h(X) := H(1, X) \in \mathbb{F}[X]$, a monic polynomial of degree d and nonzero constant coefficient. Let $\mathbb{F}'$ be the splitting field of $h(X)$; under our assumptions on a_d, all the roots of $h(X)$ are nonzero. Let q be the cardinality of $\mathbb{F}'$. Every nonzero element $\alpha \in \mathbb{F}'$, and hence every root of $h(X)$, satisfies $\alpha^{q-1} = 1$.

Finally, let q^m be a power of q not less than any multiplicity of any root of $h(X)$. Since every root of h satisfies the polynomial $X^{q-1} - 1$, we know that $h(X)$ divides the polynomial $(X^{q-1} - 1)^{q^m} = X^{q^m(q-1)} - 1$. Set $e = q^m(q-1)$, and let $s(X)$ be the polynomial in $\mathbb{F}[X]$ satisfying $h(X)s(X) = X^e - 1$.

Now we finally "rehomogenize" again; if $S \in \mathbb{F}[y, X]$ is the homogenization of $s(X)$, then $Q \cdot S = X^e - y^e$, so that S is the homogeneous scaling factor for P that we seek. $\qquad\square$

6D. *The main induction for the proof the empty-middle NGT.* From now on, having already fixed p, we will always assume that R is a ring of characteristic p, not necessarily finite.

Definition. If $T : R[y] \to R[y]$ an R-linear operator, we will call T a (d, D)-*NRO*, for *nilpotent recursion operator*, if T satisfies the conditions of Theorem 4; that is, T lowers degrees, and $\{T(y^n)\}$ satisfies an $R[y]$-linear recursion with companion polynomial

$$X^d + ay^d + (\text{terms of total degree} \le d - D)$$

for some $d \ge 1$ and some $D \ge 1$. Note that any (d, D)-NRO is a (d, D')-NRO for any $1 \le D' \le D$.

The proof of Proposition 6.2 shows that, if R is a finite field, then any T satisfying the conditions of Theorem 1 is in fact a (d, D)-NRO for some d and D.

On the other hand, we make the following definition, for any triple (b, d, D) with $b \ge d \ge D \ge 1$:

Definition. A function $c : \mathbb{Q}_{\ge 0} \to \mathbb{Q}_{\ge 0}$ is a (b, d, D)-*nilgrowth witness* if it satisfies the following properties:

(1) *Discreteness*: $c(\mathbb{N})$ is contained in a lattice of $\mathbb{Q}$ (that is, $\exists M \in \mathbb{N}$ with $Mc(\mathbb{N}) \subset \mathbb{N}$).

(2) *Growth property*: $c(n) \asymp n^{\log(b-D)/\log b}$ as $n \to \infty$.

(3) *Base property*: $0 = c(0) < c(1) < \cdots < c(d-1)$ and $c(d-D) < c(d)$.

(4) *Step property*: For any $k \ge 0$, and any pair $(i, j) \in \{0, 1, \ldots, d\}^2$ with either $(i, j) = (d, d)$ or $i - j \ge D$, and any integers n, m satisfying $db^k \le n < db^{k+1}$ and $jb^k \le m$ we have

$$c(n) - c(n - ib^k) \ge c(m) - c(m - jb^k).$$

In this section, we prove, using strong induction, that if b is a power of p, then the growth of the nilpotence index of a (d, D)-NRO is bounded by the growth of a (b, d, D)-nilgrowth witness.

Proposition 6.3. *Let T be a (d, D)-NRO, and $b \ge d$ a power of p. If c is a (b, d, D)-witness, then $N_T(y^n) \le c(n)$.*

Before proving Proposition 6.3, we record a corollary using the growth property above.

Corollary 6.4. *Suppose that T is a (d, D)-NRO, and let $b = p^{\lceil \log_p d \rceil}$. If there exists a (b, d, D)-witness, then $N_T(y^n) \ll n^{\log(b-D)/\log b}$.*

In other words, in this section we reduce the proof of Theorem 4 to establishing the existence of a $(p^{\lceil \log_p d \rceil}, d, D)$-witness for the (d, D)-NRO T, provided that D is not too big.

Proof of Proposition 6.3. Given a (b, d, D)-witness c, we define a new function $\tilde{c} : R[y] \to \mathbb{N} \cup \{-\infty\}$ via

$$\tilde{c}\left(\sum a_n y^n\right) := \max\{c(n) : a_n \neq 0\} \quad \text{and} \quad \tilde{c}(0) := -\infty.$$

We will show that T lowers the $\tilde{c}$-value of polynomials in $R[y]$: that is, that for any nonzero $f \in R[y]$, we have $\tilde{c}(T(f)) < \tilde{c}(f)$.

It suffices to show this for $f = y^n$.

Write x_n for $T(y^n)$. We will use strong induction to show that $\tilde{c}(x_n) < c(n)$.

The base case is all n with $0 \leq n < d$. Since $\deg x_n < n$, the statement $\tilde{c}(x_n) < c(n)$ for $n < d$ is implied by the statement that c is strictly increasing on $\{0, 1, \ldots, d-1\}$. This is the base property above.

For $n > d$, let $k \geq 0$ be the integer so that $d \cdot b^k \leq n < d \cdot b^{k+1}$. Let $P(X) \in \mathbb{F}[y][X]$ be the companion polynomial of the given recursion satisfied by the sequence $\{x_n\}$. Let

$$\mathcal{I} := \{(i, j) : 0 \leq j < j + D \leq i \leq d\} \cup \{(d, d)\}.$$

By assumption, P has the form

$$P = X^d + \sum_{(i,j) \in \mathcal{I}} a_{i,j} y^j X^{d-i}$$

for some $a_{i,j} \in R$. By Corollary 2.1, the sequence $\{x_n\}$ also satisfies the order-db^k recursion corresponding to P^{b^k}: namely, for all $n \geq db^k$, we have

$$x_n = - \sum_{(i,j) \in \mathcal{I}} a_{i,j} y^{jb^k} x_{n-ib^k}.$$

We will show that, if y^m appears with nonzero coefficient in one of the terms on the right-hand side above, then $c(m) < c(n)$. Since $\tilde{c}(x_n)$ is equal to one of these $c(m)$s, this will imply our claim. So suppose that y^m appears with nonzero coefficient in the (i, j)-term on the right-hand side. That is, y^m appears in $y^{jb^k} x_{n-ib^k}$ for some $(i, j) \in \mathcal{I}$. That means that y^{m-jb^k} appears in x_{n-ib^k}. Note that $i \geq D$, so that $n - ib^k < n$, and the induction assumption applies; since y^{m-jb^k} appears in x_{n-ib^k}, we can assume that $c(m - jb^k) < c(n - ib^k)$.

To show that $c(m) < c(n)$, it therefore suffices to show that

$$c(n) - c(m) \geq c(n - ib^k) - c(m - jb^k),$$

since the latter is assumed to be strictly positive. But this, slightly rearranged, is just the step property from the definition of a (b, d, D)-witness above. $\square$

We now aim to construct a (b, d, D)-witness if $b - d \leq 1$ or if $D \leq \frac{b}{2}$. This will occupy the next three sections. In Section 7 we investigate the properties of a *content* function, which writes numbers in one base and reads them in another. In Section 8, we establish some inequalities about the content of rational numbers. In Section 9, we use the content function to construct a (b, d, D)-nilgrowth witness, completing the proof of Theorem 4.

7. The content function and its properties

In this section we will introduce a function $c : \mathbb{Q}_{\geq 0} \to \mathbb{Q}_{\geq 0}$ that will serve as a nilgrowth witness in the proof of Theorem 4. This type of function was first introduced in the appendix to [Bellaïche and Khare 2015], loosely inspired by the "code" of [Nicolas and Serre 2012a].

7A. *Base-b representation of numbers.* We fix an integer $b \geq 2$ to be the *base*.

Let $\mathcal{D}(b) = \{0, \ldots, b-1\}$ be the alphabet of digits base b, and $\mathcal{D}(b)^*$ the set of finite words on $\mathcal{D}(b)$, including the empty word ϵ. The number-of-letters function for a word $x \in \mathcal{D}(b)^*$ will be denoted by $\ell(x)$, for *length*.

Let $\mathcal{R}(b)$ be the set of all bi-infinite pointed words

$$\underline{x} = \ldots x_2 x_1 x_0 . x_{-1} x_{-2} \ldots$$

on $\mathcal{D}(b)$ that start with $^\infty 0$ (the digit 0 repeated infinitely to the left). The set of finite words $\mathcal{D}(b)^*$ naturally embeds into $\mathcal{R}(b)$ via $x \mapsto (^\infty 0) x . (0^\infty)$, where 0^∞ is the digit 0 repeated infinitely to the right. More generally, any pointed right-infinite word will be viewed as an element of $\mathcal{R}(b)$ by appending $^\infty 0$ on the left. For $\underline{x} \in \mathcal{R}(b)$, we can define the real number $\pi_b(\underline{x}) \in \mathbb{R}_{\geq 0}$ by reading it as a sequence of digits base b, via $\pi_b(\underline{x}) := \sum_i x_i b^i$. Since $x_i = 0$ for $i \gg 0$, this sum converges. The map π_b is not injective; indeed, $\sum_{i<k}(b-1)b^i = b^k$, so that for any finite word w and digit $x \neq b-1$, and any radix point placement, we have $\pi_b(wx(b-1)^\infty) = \pi_b(w(x+1)0^\infty)$. But we can choose a section of π_b by restricting the domain: Let $\mathcal{R}'(b) \subset \mathcal{R}(b)$ the subset of those that do not end with $(b-1)^\infty$. Then the reading-base-b function $\pi_b : \mathcal{R}'(b) \to \mathbb{R}_{\geq 0}$ is a bijection, and the inverse map $\rho_b : \mathbb{R}_{\geq 0} \to \mathcal{R}'(b)$ takes a nonnegative real number q to its *normal* (that is, not ending in $(b-1)^\infty$) base-b representation $\underline{x} = \rho_b(q)$ satisfying $\pi_b(\underline{x}) = q$.

The base-b representation $\rho_b(q)$ is *eventually periodic* (that is, ends with z^∞ for some finite word z) if and only if $q \in \mathbb{Q}_{\geq 0}$. For $q \in \mathbb{Q}_{\geq 0}$, then, we know that

$$\rho_b(q) = x . y z^\infty,$$

where x, y, z are in $\mathcal{D}(b)$. If we insist that x does not start with 0 and that first y and then z have minimal length among such representations, then x, y, and z are defined uniquely. We will assume this minimality from now on. Note that by construction x and y may be empty words, but z has length at least 1.

We define, then, three constants associated to $q \in \mathbb{Q}_{\geq 0}$:

$$\ell(q) = \ell_b(q) := \ell(x) = \max\{0, \lfloor \log_b q \rfloor + 1\},$$

$$s(q) = s_b(q) := \ell(y) = \min\{k \geq 0 : \text{denominator of } b^k q \text{ is prime to } b\},$$

$$t(q) = t_b(q) := \ell(z) = \min\{k \geq 1 : \text{denominator of } b^{s(q)} q \text{ divides } b^k - 1\}.$$

In particular, we know that, for $q \in \mathbb{Q}_{\geq 0}$, we have

$$q = n + \frac{u}{b^{s(q)}} + \frac{m}{b^{s(q)}(b^{t(q)} - 1)}, \tag{7-1}$$

where n, u, and m are all integers with $n = \lfloor q \rfloor = \pi_b(x)$, $u = \pi_b(y)$, and $m = \pi_b(z)$.

We will need the following very simple lemma.

Lemma 7.1. *Given $q \in \mathbb{Q}_{\geq 0}$ and a base b, if $q' \in q\mathbb{N}$, then*

(1) $s_b(q') \leq s_b(q)$,

(2) $t_b(q')$ *divides* $t_b(q)$.

The proof follows from the fact that the denominator of q' is a divisor of the denominator of q. Alternatively, one can consider the effect of multiplication by integers on base-b expansions.

7B. *The content function.* Now let $b, \beta \geq 2$ be bases. Define the (b, β)-*content* of any q in $\mathbb{R}_{\geq 0}$ to be the result of reading the normal base-b representation of q in base β,

$$c_{b,\beta}(q) := \pi_\beta(\rho_b(q)).$$

Note that π_β makes sense as a function $\mathcal{R}(b) \to \mathbb{R}_{\geq 0}$; the series $\sum_{i \leq k} x_i b^i$ always converges if the x_i are bounded.

Examples. (1) Since $\rho_5(196) = 1241$, we have $c_{5,3}(196) = 1 \cdot 3^3 + 2 \cdot 3^2 + 4 \cdot 3 + 1 = 58$.

(2) We have $\rho_7\left(\frac{1}{3}\right) = 0.(2)^\infty$. Therefore $c_{7,5}(n) = 2\sum_{i \geq 1} 5^{-i} = \frac{1}{2}$.

(3) $c_{8,3}\left(\frac{1}{6}\right) = \pi_3(0.1(25)^\infty) = \frac{1}{3} + \left(\frac{2}{3^2} + \frac{5}{3^3}\right)\sum_{i \geq 0} 3^{-2i} = \frac{1}{3} + \frac{11}{27} \cdot \frac{9}{8} = \frac{19}{24}$.

The following lemma, which will be used frequently, is an easy computation.

Lemma 7.2. *For $q \in \mathbb{Q}_{\geq 0}$, let $s = s_b(q)$ and $t = t_b(q)$. Then if $q = n + \dfrac{u}{b^s} + \dfrac{m}{b^s(b^t - 1)}$ as in* (7-1) *above, we have*

$$c_{b,\beta}(q) = c_{b,\beta}(n) + \frac{c_{b,\beta}(u)}{\beta^s} + \frac{c_{b,\beta}(m)}{\beta^s(\beta^t - 1)}.$$

We will also use the following growth estimate:

Lemma 7.3. *We have $c_{b,\beta}(n) \asymp n^{\log_b \beta}$. More precisely, for $n \geq 1$, we have*

$$\beta^{-1} n^{\log_b \beta} < c_{b,\beta}(n) < \frac{\beta(b-1)}{\beta - 1} n^{\log_b \beta}.$$

Proof. Let $\ell = \ell_b(n)$, so that for $n \geq 1$ we have $\ell = 1 + \lfloor \log_b n \rfloor$, or $\log_b n < \ell \leq 1 + \log_b n$. Then, on one hand, the (b, β)-content of n is bounded above by π_β of the infinite pointed word $(b-1)^\ell.(b-1)^\infty$:

$$c_{b,\beta}(n) < \sum_{k < \ell} (b-1)\beta^k = \frac{b-1}{\beta-1}\beta^\ell \leq \frac{\beta(b-1)}{\beta-1}\beta^{\log_b n} = \frac{\beta(b-1)}{\beta-1}n^{\log_b \beta}.$$

On the other hand, the (b, β)-content of n is at least π_β of the pointed word $1(0^{\ell-1}).0^\infty$:

$$c_{b,\beta}(n) \geq \beta^{\ell-1} > \beta^{-1}n^{\log_b \beta}. \qquad \square$$

In particular, if $\beta < b$, then $c_{b,\beta}$ grows slower than linearly in n.

Remarks. (1) If $\beta < b$, then $c_{b,\beta}$ is never monotonically increasing, even on the integers. Indeed,

$$c_{b,\beta}(b^k - 1) = \frac{b-1}{\beta-1}(\beta^k - 1),$$

which is greater than $c_{b,\beta}(b^k) = \beta^k$ as soon as $\beta^k > \frac{\beta-1}{b-2}$.

(2) The sequence $\{c_{b,\beta}(n)\}_{n \in \mathbb{N}}$ is *b-regular* in the sense of Allouche and Shallit [1992]. A sequence $\{s_n\} \in \mathbb{Q}^{\mathbb{N}}$ is said to be b-regular if there exists a $\mathbb{Q}$-vector space V, endomorphisms $M_0, \ldots, M_{b-1}$ of V, a vector $v \in V$, and a functional $\lambda : V \to \mathbb{Q}$ so that $s_{\pi_b(n_\ell \ldots n_0)} = \lambda(M_{n_\ell} \cdots M_{n_0} v)$. For the sequence $\{c_{b,\beta}(n)\}$, we can take $V = \mathbb{Q}^2$, $M_i = \begin{bmatrix} 1 & i \\ 0 & \beta \end{bmatrix}$, $v = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ and $\lambda = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Indeed, b-regularity appears to be lurking in many places in this theory, but we have not yet been able to make use of it.[vi]

Finally, we record a trivial scaling property of $c_{b,\beta}$:

Lemma 7.4. *For $n \in \mathbb{R}_{\geq 0}$, and any $k \in \mathbb{Z}$, we have $c_{b,\beta}(b^k n) = \beta^k c_{b,\beta}(n)$.*

7C. *Content and the carry-digit word.* Even though the content function is not monotonic, it behaves well under addition. For $m, n \in \mathbb{R}_{\geq 0}$, we define $r_b(m, n) \in \mathcal{R}(2)$ to be the word of carry digits when the sum $s = m + n$ is computed in base b. More precisely, let $\rho_b(m) = \underline{m}$, $\rho_b(n) = \underline{n}$, and $\rho_b(s) = \underline{s}$, and let $r_b(m, n) := \underline{r}$ satisfying

$$m_i + n_i + r_{i-1} = s_i + br_i \quad \text{for all } i \text{ in } \mathbb{Z}. \tag{7-2}$$

Since m_i, n_i, and s_i are all 0 for $i > \ell_b(s)$, and since $r_i \in \{0, 1\}$, the set of equations above defines r_i uniquely, inductively down from $i = \ell_b(s)$.

Examples. (1) We compute $r_3(77, 11)$ starting with $\rho_3(77) = 2212$ and $\rho_3(11) = 102$. When the addends have finite base-b expansions, the carry digits appear as a byproduct of the base-b addition algorithm. Below, we read off that $r_3(77, 5) = 1101$:

$$
\begin{array}{r}
{\scriptstyle 1\,1\,0\,1} \\
2212 \\
+102 \\
\hline
10021.
\end{array}
$$

[vi]For example, the Nicolas–Serre code sequence $h(n)$ defined in [Nicolas and Serre 2012a, §4.1] is 2-regular, as are its constituent parts $n_3(n)$ and $n_5(n)$.

Note the shift one space to the right in our indexing the conventional carry digit notation.

(2) We compute $r_5\left(\frac{53}{60}, \frac{23}{100}\right)$. We have $\rho_5\left(\frac{53}{60}\right) = 0.4(20)^\infty$ and $\rho_5\left(\frac{23}{100}\right) = 0.10(3)^\infty$. Their sum is $\frac{167}{150} = \pi_5(1.02(40)^\infty)$. Comparing the base-5 expansions of the two addends with the expansion of the sum allows us to compute the carry digits left to right.

$$\begin{array}{r} 1\ 00101010...\\ 0.42020202\ldots\\ +0.10333333\ldots\\ \hline 1.02404040\ldots \end{array}$$

Therefore $r_5\left(\frac{53}{60}, \frac{23}{100}\right) = 0.10(01)^\infty$. In this case, we will get the same infinite carry-digit word if we take the "limit" of the finite carry-digit words obtained by truncating the expansions of the two addends.

(3) Note that $r_{10}\left(\frac{1}{3}, \frac{2}{3}\right) = 0.1^\infty$, even though any finite truncation of the decimal expansions of $\frac{1}{3}$ and $\frac{2}{3}$ would yield no carry digits in the sum. If the addends are not in $\mathbb{Z}\left[\frac{1}{b}\right]$ but the sum is, one computes the expansion of the sum before computing the carry-digit word.

The carry digit word exactly keeps track of the difference between values of $c_{b,\beta}$:

Lemma 7.5. *For m, n in $\mathbb{R}_{\geq 0}$, we have*

$$c_{b,\beta}(m) + c_{b,\beta}(n) = c_{b,\beta}(m+n) + (b-\beta)\pi_\beta(r_b(m,n)).$$

Proof. Let $s = m + n$, and let $\underline{m}$, $\underline{n}$, $\underline{s}$ be the corresponding base-b expansions and $\underline{r}$ the carry-digit word. Scaling (7-2) by β^i and summing up over all i gives us

$$\sum m_i \beta^i + \sum n_i \beta^i + \sum r_{i-1}\beta^i = \sum s_i \beta^i + b \sum r_i \beta^i$$

or, equivalently,

$$c_{b,\beta}(m) + c_{b,\beta}(n) + \beta\pi_\beta(\underline{r}) = c_{b,\beta}(m+n) + b\pi_\beta(\underline{r}). \qquad \square$$

We will typically use Lemma 7.5 when comparing $c_{b,\beta}(m)$ and $c_{b,\beta}(n)$ by analyzing $c_{b,\beta}(m-n)$ and $r_b(m-n, n)$.

Examples. (1) We have $c_{3,2}(77) = \pi_2(2212) = 28$, and $c_{3,2}(88) = \pi_2(10021) = 21$. The difference is accounted for by

$$c_{3,2}(88 - 77) = \pi_2(102) = 6 \quad \text{and} \quad \pi_2(r_3(77, 11)) = \pi_2(1101) = 13.$$

Since $28 + 6 = 21 + 13$, we are consistent with Lemma 7.5.

(2) Let $(b, \beta) = (5, 3)$, and consider the $\frac{53}{60} + \frac{23}{100}$ example from above. Using Lemma 7.2, we find that

$$c_{5,3}\left(\tfrac{53}{60}\right) = \pi_3(0.4(20)^\infty) = \frac{\pi_3(4)}{3} + \frac{\pi_3(20)}{3(3^2 - 1)} = \frac{19}{12}$$

$$c_{5,3}\left(\tfrac{23}{100}\right) = \pi_3(0.10(3)^\infty) = \frac{\pi_3(10)}{3^2} + \frac{\pi_3(3)}{3^2(3 - 1)} = \frac{1}{2}$$

$$c_{5,3}\left(\tfrac{167}{150}\right) = \pi_3(1.02(40)^\infty) = \pi_3(1) + \frac{\pi_3(02)}{3^2} + \frac{\pi_3(40)}{3^2(3^2 - 1)} = \frac{25}{18}$$

$$\pi_3\left(r_5\left(\tfrac{53}{60}, \tfrac{23}{100}\right)\right) = \pi_3(0.10(01)^\infty) = \frac{\pi_3(10)}{3^2} + \frac{\pi_3(01)}{3^2(3^2 - 1)} = \frac{25}{72}.$$

As expected from Lemma 7.5, we have $\frac{19}{12} + \frac{1}{2} = \frac{25}{12} = \frac{25}{18} + 2 \cdot \frac{25}{72}$.

(3) Finally, let $b = 10$ and return to the addition equation $\frac{1}{3} + \frac{2}{3} = 1$. For $a \in \mathcal{D}(9)$, we have $\pi_\beta(0.a^\infty) = \frac{a}{\beta - 1}$. Therefore the two sides of the Lemma 7.5 equation agree:

$$\text{(LHS)} \qquad\qquad c_{10,\beta}\left(\tfrac{1}{3}\right) + c_{10,\beta}\left(\tfrac{2}{3}\right) = \pi_\beta(0.3^\infty) + \pi_\beta(0.6^\infty) = \frac{9}{\beta - 1}$$

$$\text{(RHS)} \quad c_{10,\beta}(1) + (10 - \beta)\pi_\beta\left(r_{10}\left(\tfrac{1}{3}, \tfrac{2}{3}\right)\right) = 1 + (10 - \beta)\pi_\beta(0.1^\infty) = \frac{9}{\beta - 1}.$$

8. Content of some proper fractions

In this section, we will prove inequalities about $(b, b - D)$-content of some proper fractions that we will use in Section 9 to produce a (b, d, D)-nilgrowth witness.

8A. *Unit fractions in base b.* Let $b \geq 2$ be a base, and fix a denominator d with $1 < d \leq b$. To motivate the discussion, we note that

$$\frac{1}{d} = \frac{\frac{1}{b}}{1 - \frac{b-d}{b}} = \sum_{k \geq 1}(b - d)^{k-1}b^{-k}.$$

Therefore, the base-b expansion of $\frac{1}{d}$ "wants" to be $0.1(b - d)(b - d)^2(b - d)^3 \ldots$. Of course, unless $b - d \leq 1$, this is not possible; $(b - d)^k$ is not a digit base b for k large enough. However, letting $\rho_b\left(\frac{1}{d}\right) = 0.a_1 a_2 a_3 \ldots$, we can say the following:

Lemma 8.1. *For $k \geq 1$, we have $a_i = (b - d)^{i-1}$ for $i = 1, \ldots, k$ if and only if $(b - d)^k < d$.*

Proof. We will establish this claim by induction on k. The a_k can be defined recursively via

$$a_k = \left\lfloor \frac{b^k}{d} - \sum_{i=1}^{k-1} a_i b^{k-i} \right\rfloor = \left\lfloor \frac{b^k}{d} \right\rfloor - \sum_{i=1}^{k-1} a_i b^{k-i}. \qquad (8\text{-}1)$$

For $k = 1$, we have $a_1 = \left\lfloor \frac{b}{d} \right\rfloor \geq 1$. Therefore $a_1 = 1$ if and only if $\frac{b}{d} < 2$, which is equivalent to $(b - d)^1 < d$. So the claim for $k = 1$ is true.

Now suppose we already know that $a_i = (b-d)^{i-1}$ for $i < k$. Then $a_k = (b-d)^{k-1}$ if and only if

$$
\begin{aligned}
1 &> \left(\frac{b^k}{d} - \sum_{i=1}^{k-1} a_i b^{k-i} \right) - (b-d)^{k-1} \\
&= \frac{b^k}{d} - (b^{k-1} + (b-d)b^{k-2} + \cdots + (b-d)^{k-2}b + (b-d)^{k-1}) \\
&= \frac{b^k}{d} - \frac{b^k - (b-d)^k}{b - (b-d)} = \frac{(b-d)^k}{d},
\end{aligned}
$$

as desired. $\square$

The same argument also implies the immediate:

Corollary 8.2. *If $a_i = (b-d)^{i-1}$ for $i < k$, then $a_k = (b-d)^{k-1} + \left\lfloor \frac{1}{d}(b-d)^k \right\rfloor$.*

We can now delineate what $\frac{1}{d}$ must look like in base b.

Lemma 8.3. *For $d \leq b$, the base-b expansion of $\frac{1}{d}$ falls into one of five mutually exclusive cases.*

(1) $\rho_b\left(\frac{1}{d}\right) = 0.2^+ \ldots$ *(in other words, $a_1 \geq 2$) if and only if $d \leq \frac{b}{2}$.*

(2) $\rho_b\left(\frac{1}{d}\right) = 0.13^+ \ldots$ *(i.e., $a_1 = 1$ and $a_2 \geq 3$) if and only if $\frac{b}{2} < d \leq \frac{b^2}{b+3} = b - 3 + \frac{9}{b+3}$.*

(3) $\rho_b\left(\frac{1}{d}\right) = 0.124^+ \ldots$ *(i.e., $a_1 = 1$, $a_2 = 2$, and $a_3 \geq 4$) if and only if $b > 6$ and $d = b - 2$.*

(4) $\rho_b\left(\frac{1}{d}\right) = 0.1^\infty$ *if and only if $d = b - 1$.*

(5) $\rho_b\left(\frac{1}{d}\right) = 0.10^\infty$ *if and only if $d = b$.*

Proof. If $b - d = 0$ or $b - d = 1$, then $a_i = (b-d)^{i-1}$ for all i (Lemma 8.1). Assume $b - d \geq 2$. If $d \leq \frac{b}{2}$, then $\frac{1}{d} \geq \frac{2}{b}$, so that $a_1 \geq 2$, as claimed. Otherwise, we must have $(b-d)^1 < d$, so that $a_1 = 1$ and $a_2 \geq b - d$. This means that $a_2 \geq 3$ unless *both* $b - d = 2$ and $(b-d)^2 < d$, in which case we have $d > 4$ (and hence $b > 6$), and $a_2 \geq (b-d)^2 = 4$. $\square$

8B. *The carry-digit word for a proper fraction in base b.* Keeping the notation b and d, we additionally fix a D with $1 \leq D < d$ and investigate $\rho_b\left(\frac{D}{d}\right) =: 0.e_1 e_2 e_3 \ldots$. The e_k satisfy the same type of recursion as the a_k, namely,

$$
e_k = \left\lfloor \frac{b^k D}{d} \right\rfloor - \sum_{i=1}^{k-1} e_i b^{k-i}.
$$

In particular, $e_1 = \left\lfloor \frac{Db}{d} \right\rfloor$, and the following generalization of Lemma 8.1 and Corollary 8.2 holds:

Lemma 8.4. *For $k \geq 0$, we have $e_i = D(b-d)^{i-1}$ for $i = 1, \ldots, k$ if and only if $D(b-d)^k < d$. Moreover, if $D(b-d)^k < d$, then $e_{k+1} = D(b-d)^k + \left\lfloor \frac{1}{d} D(b-d)^{k+1} \right\rfloor$.*

In order to understand the relationship between the a_k and the e_k we define the carry-digit word $\underline{r} = 0.r_1 r_2 r_3 \ldots$ for the addition problem $\sum_{i=1}^{D} \frac{1}{d} = \frac{D}{d}$. (See Section 7C for definitions.) In other words, set $\underline{r}^{(i)} := 0.r_1^{(i)} r_2^{(i)} r_3^{(i)} \cdots := r_b\left(\frac{i}{d}, \frac{1}{d}\right)$ for $1 \leq i \leq D - 1$, and then set $r_j := \sum_{i=1}^{d-1} r_j^{(i)}$.

Then $0 \leq r_j \leq D-1$ for every i; since $\frac{D}{d} < 1$, we have $r_1 = 0$. Putting together equations (7-2) applied to $\frac{i}{d} + \frac{1}{d}$ for $1 \leq i < D$ gives us the precise relationship between the a_k, e_k, and r_k:

$$e_k = Da_k + r_{k+1} - br_k. \tag{8-2}$$

In fact, we have a closed formula for r_k:

Lemma 8.5. *For $k \geq 1$ we have*

$$r_k \;=\; \left\lfloor \frac{Db^{k-1}}{d} \right\rfloor - D \left\lfloor \frac{b^{k-1}}{d} \right\rfloor \;=\; \left\lfloor \frac{D(b-d)^{k-1}}{d} \right\rfloor - D \left\lfloor \frac{(b-d)^{k-1}}{d} \right\rfloor.$$

Proof. The formula is true for $k = 1$ (in our case, all quantities are 0). For $k \geq 1$, we will establish the formula for $k + 1$, starting with (8-2):

$$r_{k+1} = e_k - Da_k + br_k = \left\lfloor \frac{Db^k}{d} \right\rfloor - \sum_{i=1}^{k-1} e_i b^{k-i} - D \left\lfloor \frac{b^k}{d} \right\rfloor + D \sum_{i=1}^{k-1} a_i b^{k-i} + br_k$$

$$= \left\lfloor \frac{Db^k}{d} \right\rfloor - D \left\lfloor \frac{b^k}{d} \right\rfloor + \sum_{i=1}^{k-1} b^{k-i}(br_i - r_{i+1}) + br_k = \left\lfloor \frac{Db^k}{d} \right\rfloor - D \left\lfloor \frac{b^k}{d} \right\rfloor.$$

Here we used (8-2) in the form $-e_i + Da_i = br_i - r_{i+1}$ for $1 \leq i < k$ to pass from the first line to the second, and cancellation of a telescoping sum for the final equality. Finally, we note that

$$\left\lfloor \frac{Db^k}{d} \right\rfloor - D \left\lfloor \frac{b^k}{d} \right\rfloor = \left\lfloor \frac{D(b-d)^k}{d} \right\rfloor - D \left\lfloor \frac{(b-d)^k}{d} \right\rfloor$$

because the intervening terms $\frac{1}{d}\left(D \sum_{i=1}^{k} \binom{k}{i} b^{k-i}(-d)^i \right)$ are integers, and hence can pass through the greatest-integer function to cancel. $\square$

Corollary 8.6. (1) *If $(b-d)^k < d$ for some $k \geq 0$, then for every $i \leq k+1$, we have $r_i = \left\lfloor \frac{1}{d} D(b-d)^{i-1} \right\rfloor$.*

(2) *If $D(b-d)^k < d$ for some $k \geq 0$, then for every $i \leq k + 1$, we have $r_i = 0$.*

8C. (b, β)-*Content of proper fractions.* We fix a triple (b, d, D) with $1 \leq D \leq d \leq b$ subject to the conditions that $b \geq 2$, and further impose the condition that $\beta := b - D \geq 2$. Recall the content function $c_{b,\beta}$ from Section 7B, and let $c_{b,\beta}^d : \mathbb{Q}_{\geq 0} \to \mathbb{Q}_{\geq 0}$ be the function defined by

$$c_{b,\beta}^d(n) := c_{b,\beta}\left(\frac{n}{d} \right).$$

Whenever the triple (b, d, D) is understood, we write $c = c_{b,\beta}^d$, and let $\underline{r} = 0.r_1 r_2 \ldots$ be the carry-digit word for $\frac{1}{d} + \cdots + \frac{1}{d} = \frac{D}{d}$ as in Section 8B. Lemma 7.5 implies that

$$c(D) = Dc(1) - D\pi_\beta(\underline{r}). \tag{8-3}$$

In this section, we will establish some lower bounds on $c(1)$ and $c(D)$. First, we dispatch the cases $d = b$ and $d = b - 1$, which yield easy explicit formulas.

Lemma 8.7. *Suppose that $d = b$ or $d = b - 1$, and $0 \leq i < d$. Then $c(i) = \frac{i}{d-D}$.*

Proof. Computation, see Lemma 8.3. Note that $D \leq \min\{d, b-2\}$ precludes the possibility that $d = D$. $\square$

Proposition 8.8. *If $d \leq b - 2$ and $b > 6$, then $c(1) \geq \frac{\beta+1}{\beta(\beta-1)}$.*

Remark. It is a simple exercise to check that the only exceptions for $b \leq 6$ are in fact $(b, d, D) = (4, 2, 2)$, $(5, 3, 3)$, or $(6, 4, 4)$ by exhausting all cases.

Proof. We go through the first three cases of Lemma 8.3. Note that $\beta = 2$ implies that all of the inequalities in $2 \leq b - d \leq b - D = \beta$ are equalities, so that $d = D$ and $d = b - 2$. In particular, if $\beta = 2$ and $b > 6$, then we must be in the third case.

(1) If $\rho_b\left(\frac{1}{d}\right)$ starts with 0.2^+, then $c(1) \geq \pi_\beta(0.2) = \frac{2}{\beta}$. Since $\beta \geq 3$, we have $2 \geq \frac{\beta+1}{\beta-1}$, so that $c(1) \geq \frac{\beta+1}{\beta(\beta-1)}$, as desired.

(2) If $\rho_b\left(\frac{1}{d}\right)$ starts with 0.13^+, then we have $c(1) \geq \rho_\beta(0.13) = \frac{\beta+3}{\beta^2}$. This last is no less than $\frac{\beta+1}{\beta(\beta-1)}$ if and only if $\beta^2 + 2\beta - 3 = (\beta + 3)(\beta - 1) \geq (\beta + 1)\beta = \beta^2 + \beta$. Therefore $\beta \geq 3$ again implies $c(1) \geq \frac{\beta+1}{\beta(\beta-1)}$, as desired.

(3) If $\rho_b\left(\frac{1}{d}\right)$ starts with 0.124^+, then

$$c(1) \geq \pi_\beta(0.124) = \frac{\beta^2 + 2\beta + 4}{\beta^3} = \frac{\beta + 1}{\beta(\beta - 1)} + \frac{2\beta - 4}{\beta^4 - \beta^3}.$$

Since $\beta \geq 2$, our claim is established. $\square$

Corollary 8.9. *If $d \leq b - 2$ and $D \leq \frac{b}{2}$, then $c(1) \geq \frac{D}{\beta(\beta-1)}$.*

Proof. For $D \leq \frac{b}{2}$, we have $\beta = b - D \geq b - \frac{b}{2} = \frac{b}{2} \geq D$. Therefore Proposition 8.8 establishes the desired inequality, the exceptional cases $(b, d, D) = (4, 2, 2)$, $(5, 3, 3)$, or $(6, 4, 4)$ being easy to check explicitly. $\square$

In the next proposition we will show that $c(D)$ is not too small, provided that d is not too big relative to b, or, failing that, that D is not too big relative to b and d.

Proposition 8.10. *Suppose $D < d \leq b - 2$ and at least one of the following conditions is satisfied*:

(1) $d \leq \frac{b}{2}$.

(2) $D < d\left(1 - \frac{1}{b-d}\right)$.

Then $c(D) \geq \frac{D(\beta+1)}{\beta(\beta-1)}$.

Remark. Computationally, it appears that the optimal statement is as follows. If $D < d \leq b - 2$, then $c(D) \geq \frac{D(\beta+1)}{\beta(\beta-1)}$ if and only if at least one of the following is true: (1') $(b-d)^2 > b-1$ or (2) $D < d\left(1 - \frac{1}{b-d}\right)$. Note that Condition (1) above implies condition (1'), but this latter is strictly weaker. Here we only prove Proposition 8.10 as stated.

Before proving Proposition 8.10, some preparatory lemmas.

Lemma 8.11. *Under the assumption $d > \frac{b}{2}$, condition (2) from Proposition 8.10 is equivalent to the inequality $r_2 < b - d - 1$.*

Proof. Apply Corollary 8.6 for $k = 0$ to deduce that $r_2 = \left\lfloor \frac{1}{d} D(b - d) \right\rfloor$. Then $r_2 < b - d - 1$ if and only if $\frac{D(b-d)}{d} < b - d - 1$ if and only if $D < \frac{d(b-d-1)}{b-d}$, which is condition (2), as desired. $\qquad\square$

As before, let $\rho_b\left(\frac{1}{d}\right) = 0.a_1 a_2 \ldots$ and $\rho_b\left(\frac{D}{d}\right) = 0.e_1 e_2 \ldots$, and let $\underline{r} = 0.r_1 r_2 \ldots$ be the carry digits as in Section 8B. Using (8-2) and the partial-sum cutoffs $c(D) \geq \pi_\beta(0.e_1 \ldots e_k) = \sum_{i=1}^{k} e_i \beta^{-i}$, we get the following partial-sum versions of (8-3):

Lemma 8.12. *For any $k \geq 1$, the quantity $c(D)$ satisfies the following inequality*:

$$c(D) \geq \frac{D \sum_{i=1}^{k}(a_i - r_i)\beta^{k-i} + r_{k+1}}{\beta^k}.$$

Corollary 8.13. *Any of the following conditions are sufficient to guarantee $c(D) \geq \frac{D(\beta+1)}{\beta(\beta-1)}$:*

(1) $\beta \geq 3$ *and* $a_1 \geq 2$;

(2) $r_2 \geq \frac{2D}{\beta-1}$;

(3) $\beta \geq 3$ *and* $a_2 - r_2 \geq 3$;

(4) $a_2 - r_2 = 2$ *and* $r_3 \geq \frac{2D}{\beta-1}$;

(5) $\beta \geq 3$ *and* $a_2 - r_2 = 2$ *and* $a_3 - r_3 \geq 3$.

Proof. We use Lemma 8.12 for each specified k. Recall that $r_1 = 0$.

(1) $k = 1$, use estimate $r_2 \geq 0$. We have $c(D) \geq \frac{2D}{\beta} \geq \frac{D(\beta+1)}{\beta(\beta-1)}$, since $2 \geq \frac{\beta+1}{\beta-1}$ for $\beta \geq 3$.

(2) $k = 1$, use estimate $a_1 \geq 1$:

$$c(D) \geq \frac{D + r_2}{\beta} \geq \frac{D + \frac{2D}{\beta-1}}{\beta} = \frac{D(\beta + 1)}{\beta(\beta - 1)}.$$

(3) $k = 2$, use estimate $a_1 \geq 1$ and $r_3 \geq 0$:

$$c(D) \geq \frac{\beta(D + r_2) + (Da_2 - br_2)}{\beta^2} = \frac{D(\beta + a_2 - r_2)}{\beta^2}.$$

This last being greater than $\frac{D(\beta+1)}{\beta(\beta-1)}$ is equivalent to $(\beta + a_2 - r_2)(\beta - 1) \geq \beta(\beta + 1)$, or $a_2 - r_2 \geq \frac{2\beta}{\beta-1}$. For $\beta \geq 3$, this is guaranteed by $a_2 - r_2 \geq 3$.

(4) $k = 2$, use estimate $a_1 \geq 1$:

$$\frac{c(D)}{D} \geq \frac{\beta + (a_2 - r_2) + \frac{r_3}{D}}{\beta^2} \geq \frac{\beta + 2 + \frac{2}{\beta-1}}{\beta^2} = \frac{\beta + 1}{\beta(\beta - 1)}.$$

(5) $k = 3$, use estimate $a_1 \geq 1$ and $r_4 \geq 0$:

$$\frac{c(D)}{D} \geq \frac{\beta^2 + (a_2 - r_2)\beta + (a_3 - r_3)}{\beta^3} \geq \frac{\beta^2 + 2\beta + 3}{\beta^3} = \frac{\beta + 1}{\beta(\beta - 1)} + \frac{\beta - 3}{\beta^3(\beta - 1)}. \qquad\square$$

Proof of Proposition 8.10. Note that the assumptions $D < d \leq b - 2$ guarantee that $\beta \geq 3$.

If condition (1) holds, then $a_1 \geq 2$ (Lemma 8.3(1)) so that Corollary 8.13(1) gives what we want. If condition (1) fails, but condition (2) holds, then by Lemma 8.11, we have $r_2 \leq b - d - 2$. Moreover, from

Lemma 8.4 for $D = 1$ and $k = 1$, we know that $a_2 \geq b - d$. If either inequality is strict, Corollary 8.13(3) gives us the desired inequality. Therefore it remains to consider the case $r_2 = b - d - 2$ (so that $\frac{d(b-d-2)}{b-d} \leq D < \frac{d(b-d-1)}{b-d}$) and $a_2 = b - d$ (so that $(b - d)^2 < d^{(\text{vii})}$).

We now estimate the third digits. By Corollary 8.6 and Lemma 8.4, we have $r_3 = \lfloor \frac{1}{d} D(b - d)^2 \rfloor$ and $a_3 \geq (b - d)^2$. Condition (2) implies that $r_3 \leq \frac{1}{d} D(b - d)^2 < (b - d)(b - d - 1)$, so that

$$a_3 - r_3 > (b - d)^2 - (b - d)(b - d - 1) = b - d \geq 3,$$

so that the desired inequality holds by Corollary 8.13(5). □

9. The nilgrowth witness: finishing the proof

Let $b \geq d \geq D \geq 1$ be integers subject to the conditions $b \geq 2$ and $\beta := b - D \geq 2$, as before. In this section we exhibit a (b, d, D)-nilgrowth witness and complete the proof of Theorem 4.

Recall from Section 8C that $c_{b,\beta}^d : \mathbb{Q}_{\geq 0} \to \mathbb{Q}_{\geq 0}$ is the function defined by

$$c_{b,\beta}^d(n) := c_{b,\beta}\left(\frac{n}{d}\right).$$

Also define the integer constant $M_{b,\beta}^d := \beta^{s_b(1/d)}(\beta^{t_b(1/d)} - 1)$. Here $c_{b,\beta}$ is the (b, β)-content function, first defined in Section 7B, and s_b and t_b count the number of digits after the decimal point of the preperiod and the period, respectively, of base-b expansions; see definition before (7-1).

The following theorem, combined with Corollary 6.4, will prove Theorem 4, completing in turn the proof of Theorem 1.

Theorem 5. *If $b - d \leq 1$, or if $D \leq \frac{b}{2}$, then the function $c_{b,b-D}^d$ is a (b, d, D)-nilgrowth witness.*

We begin the proof of Theorem 5. Recall from Section 6D that a (b, d, D)-nilgrowth witness must satisfy four properties: discreteness, growth, base, and step. We establish the first two immediately.

Lemma 9.1 (discreteness property). *For any $n \in \mathbb{N}$, we have $M_{b,\beta}^d c_{b,\beta}^d(n) \in \mathbb{N}$.*

Proof. It suffices to see that $\beta^{s_b(1/d)}(\beta^{t_b(1/d)} - 1)c_{b,\beta}(n)$ is an integer for $n \in \frac{1}{d}\mathbb{N}$. For $n = \frac{1}{d}$ this follows from Lemma 7.2, and for general $n \in \frac{1}{d}\mathbb{N}$ from Lemmas 7.2 and 7.1. □

Lemma 9.2 (growth property). *We have $c_{b,\beta}^d(n) \asymp n^{\log_b \beta}$.*

Proof. Lemma 7.3. □

It remains to establish the base property and the step property.

For $m, n \in \mathbb{Q}_{\geq 0}$ with $m \geq n$, set

$$R(m, n) := R_{b,\beta}^d(m, n) := D\pi_\beta r_b\left(\frac{m-n}{d}, \frac{n}{d}\right).$$

(vii)Incidentally this implies the failure of condition (1′), which should conjecturally replace condition (1) as noted in the remark after the statement of Proposition 8.10.

Here r_b is the carry-digit word, as in Section 7C. We then have, for m, n as above

$$c(n) - c(n-m) = c(m) - R(n,m). \tag{9-1}$$

This is just a restatement of Lemma 7.5, in the form in which we will use it below.

We now use the technical results of Section 8 to prove that our candidate nilgrowth witness satisfies the base property and the step property.

Lemma 9.3 (base property). *Suppose that $d = b$ or $d = b - 1$ or $D \leq \frac{b}{2}$. Then we have*:

(1) $c(d - D) \leq c(d)$;

(2) $0 = c(0) < c(1) < \cdots < c(d-1)$.

Remark. Part (1) is in fact true without the assumption $D \leq \frac{b}{2}$, but we do not need this greater generality. Part (2) above is not generally true if $D > \frac{b}{2}$. For example, for $(b, d, D) = (7, 5, 5)$, we have $c(2) = c(3) = 3$; and for $(b, d, D) = (11, 9, 7)$ we have $c(4) = \frac{334}{195} > \frac{316}{195} = c(5)$. The condition delineated here is certainly not optimal, however.

Proof. If $d = b$ or $d = b - 1$, then both statements are immediate from the formula in Lemma 8.7. (Note that $c(d)$ is always 1.) Assume therefore that $d \leq b - 2$.

(1) If $d = D$, then the inequality is trivial; so assume $D < d$. By (9-1), we have

$$c(d) - c(d - D) = c(D) - R(d, D).$$

Certainly $r_b\left(\frac{D}{d}, \frac{d-D}{d}\right)$ can be no greater than 0.1^∞. Therefore

$$R(d, D) = D\pi_\beta r_b\left(\tfrac{D}{d}, \tfrac{d-D}{d}\right) \leq D\pi_\beta(0.1^\infty) = \tfrac{D}{\beta-1}.$$

On the other hand, by Proposition 8.10, we know that $c(D) \geq \frac{D(\beta+1)}{\beta(\beta-1)} > \frac{D}{\beta-1}$. Therefore $c(d) > c(d - D)$ (and in fact the inequality is strict).

(2) It suffices to show that, for $0 < i < d$, we have $c(i) > c(i - 1)$. By (9-1) this is equivalent to the inequality $c(1) > R(i, 1)$. Since $i < d$, we know that

$$R(i, 1) = D\pi_\beta r_b\left(\tfrac{i-1}{d}, \tfrac{1}{d}\right) \leq D\pi_\beta(0.01^\infty) = \tfrac{D}{\beta(\beta-1)}.$$

Now Corollary 8.9 completes the claim. $\qquad\square$

Lemma 9.4 (step property). *Suppose $d = b$ or $d = b - 1$ or $D \leq \frac{b}{2}$. If $(i, j) \in \mathcal{I}$, and n, m are integers with $db^k \leq n < db^{k+1}$ and $jb^k \leq m$, then*

$$c(n) - c(n - ib^k) \geq c(m) - c(m - jb^k).$$

Here as before $\mathcal{I} = \{(i, j) : 0 \leq j < j + D \leq i \leq d\} \cup \{(d, d)\}$ is the set of pairs (i, j) so that $y^j X^{d-i}$ can appear in the companion polynomial of the recursion in question; see proof of Proposition 6.3.

Proof. We use Lemma 7.4 to divide each equation by β^k, and replace n and m by $\frac{n}{b^k}$ and $\frac{m}{b^k}$, respectively. We therefore seek to show that for $n, m \in \mathbb{Z}\left[\frac{1}{b}\right]_{\geq 0}$ satisfying $d \leq n < db$ and $m \geq j$, we have

$$c(n) - c(n-i) \geq c(m) - c(m-j).$$

Using (9-1) twice and rearranging terms, the desired statement is equivalent to

$$c(i) - c(j) \geq R(n, i) - R(m, j).$$

Since $R(m, j)$ is nonnegative, it suffices to show that

$$c(i) - c(j) \geq R(n, i).$$

We now take two cases. If $i = d$, then $R(n, i) = 0$, so that it suffices to show that $c(d) \geq c(j)$ for $(d, j) \in \mathcal{I}$. For $j = d$, this is clear; and for $j \leq i - D$, this follows from both parts of Lemma 9.3 above.

If, on the other hand, $i < d$, then (9-1) gives us $c(i) - c(j) = c(i - j) - R(i, j)$, which reduces the desired statement to

$$c(i - j) \geq R(n, i) + R(i, j). \tag{9-2}$$

If $d = b$ or $d = b - 1$, then $R(i, j) = 0$; since $c(i - j) = \frac{i-j}{d-D} \geq \frac{D}{d-D}$ (Lemma 8.7), it remains to show that $R(n, i) \leq \frac{D}{d-D}$. If $d = b$, then at most one digit is carried, so that $R(n, i) \leq D\pi_\beta(0.1) = \frac{D}{\beta}$. And if $d = b - 1$, then every digit may be carried, so that $R(n, i) \leq D\pi_\beta(0.1^\infty) = \frac{D}{\beta-1}$. In both cases, the desired inequality holds.

On the other hand if $d \leq b - 2$ $\left(\text{and so } D \leq \frac{b}{2}\right)$, then we reason as follows. The left-hand side of desired inequality (9-2) is bounded below by $c(D)$, and the right-hand side is bounded above by

$$D\pi_\beta(0.1^\infty) + D\pi_\beta(0.01^\infty) = \frac{D(\beta+1)}{\beta(\beta-1)}.$$

Therefore it suffices to show that $c(D) \geq \frac{D(\beta+1)}{\beta(\beta-1)}$ which is established in Proposition 8.10. $\square$

Lemmas 9.3 and 9.4 complete the proof of Theorem 5, which in turn completes the proof of Theorem 4, and hence of Theorem 1.

10. Complements

10A. *Refinement of Theorem 2.* We state a refinement of the toy version of the NGT. One can also obtain similar refinements of Theorem 4.

Theorem 6 (refined toy NGT). *Let $\mathbb{F}$ be a field of characteristic p and let $q = p^k$. Suppose that $T : \mathbb{F}[y] \to \mathbb{F}[y]$ is an $\mathbb{F}$-linear operator satisfying the following two conditions*:

(1) *For $f \in \mathbb{F}[y]$, we have $\deg T(f) \leq \deg f - E$ for some $E \geq 1$.*

(2) *The sequence $\{T(y^n)\}_n$ satisfies a linear recursion whose companion polynomial has the shape*

$$P = (X + cy)^d + (\text{terms of total degree} \leq d - D) \in \mathbb{F}[y][X]$$

for some $d \leq q$, $D \geq 1$, and $c \in \mathbb{F}$.

Then

$$N_T(f) \leq \frac{(q-D)(q-1)}{E(q-D-1)} (\deg f)^{\log(q-D)/\log q}.$$

Remark. Since the total degree of the companion polynomial of the recursion is the same as the order, it suffices to check the condition that $\deg T(f) \leq \deg f - E$ on $f = 1, y, \ldots, y^{d-1}$ only.

Proof. The sequence $\{T(y^n)\}_n$ also satisfies the linear recursion with companion polynomial

$$P' = (X + cy)^{q-d} P = X^q + cy^q + (\text{terms of total degree} \leq q - D).$$

Let $c = c_{q,q-D}$, define $\tilde{c} : \mathbb{F}[y] \to \mathbb{N} \cup \{-\infty\}$ from c as in the proof of Theorem 2, and follow the same inductive argument mutatis mutandis to show that $\tilde{c}(Tf) \leq \tilde{c}(f) - E$. (The main adjustment is in Proposition 4.1(4); if i is a digit base q and $n \geq i$ has no more than 2 digits base q, then $c(n) - c(n-i)$ is either i or $i - D$; see Lemma 7.5 for a conceptual explanation.)

We have therefore shown that $N_T(y^n) \leq \frac{c(n)}{E}$. Lemma 7.3 completes the proof. $\qquad\square$

10B. *Comments on α in Theorem 4.*

How optimal is the order of growth of the nilpotence index α from the empty-middle NGT?

To this end, if K is a field, and $T : K[y] \to K[y]$ a degree-lowering linear operator, let

$$\alpha(T) := \limsup_{n \to \infty} \frac{\log N_T(y^n)}{\log n},$$

and let

$$\alpha_K(d, D) = \sup_{T \in \mathcal{L}_K(d,D)} \{\alpha(T)\},$$

where $\mathcal{L}_K(d, D)$ is the set of degree-lowering operators $T : K[y] \to K[y]$ with $\{T(y^n)\}_n$ satisfying a recurrence with companion polynomial $X^d + cy^d + (\text{terms of total degree} \leq d - D)$ for some $c \in K$. Since $N_T(y^n) \leq n$, we know that $\alpha_K(d, D) \leq 1$. The following proposition clarifies that studying $\alpha_K(d, D)$ is only interesting in characteristic p.

Proposition 10.1. *If K has characteristic zero and $D < d$, then $\alpha_K(d, D) = 1$.*

Proof. Fix d, and consider the recursion operator $T : K[y] \to K[y]$ defined by the companion polynomial $P = X^d - y^d - y$, corresponding to the recurrence $T(y^n) = (y^d + y)T(y^{n-d})$, and initial values $\{T(y^n)\}_{n=0}^{d-1} = \{0, \ldots, 0, 1\}$. We will show that $N_T(y^{kd+d-1}) = \lfloor \frac{k}{d-1} \rfloor + 1$, which will establish that $\alpha_T = 1$.

Indeed, from the recurrence, we have $T(y^n) = 0$ if $n \not\equiv -1 \bmod d$, and $T(y^{kd+d-1}) = (y^d + y)^k$. For $f = \sum a_n y^n \in K[y]$, write $e(f)$ for the set $\{n : a_n \neq 0\}$ of exponents appearing in f. From above, we see that

$$e(T(y^{kd+d-1})) = \{k, k + (d-1), k + 2(d-1), \ldots, kd\}.$$

More generally, we can show by induction that the set $S_{m,k} := e(T^m(y^{kd+d-1}))$ is an arithmetic progression of common difference $d - 1$, greatest term $d(k - (m-1)(d-1))$, and length $k - (m-1)(d-1) + 1$, so long as $k \geq (m-1)(d-1)$; otherwise the set is empty and $T^m(y^{kd+d-1}) = 0$. Indeed, from the explicit

formulation of $T(y^n)$, we see that if $S_{m,k}$ is as claimed, then the greatest element of $S_{m+1,k} = N-(d-1)$, where N is the greatest element of $S_{m,k}$ congruent to $d-1$ modulo d. Since the maximum element of $S_{m,k}$ is congruent to 0 modulo d, and every successive smaller element is $d-1$ less, we see that N is the d-th greatest element of $S_{m,k}$. In other words,

$$N = d(k - (m-1)(d-1)) - (d-1)^2,$$

so that the greatest element of $S_{m+1,k}$ is $N - (d-1) = d(k - m(d-1))$, as desired. Since the relevant coefficients are positive and we are in infinite characteristic, no cancellation of intermediate terms is possible. Finally, since $T^m(y^{kd+d-1}) \neq 0$ if and only if $k \geq (m-1)(d-1)$, we have $N_T(y^{kd+d-1}) = \lfloor \frac{k}{d-1} \rfloor + 1$, as claimed. □

For $\mathbb{F}$ a field of characteristic p, let us confine our inquiry to the case where d can be taken to be a power of p, as in Theorem 6 above. Theorem 6 tells us that $\alpha_{\mathbb{F}}(p^k, D) \leq \log(p^k - D)/\log p^k$. How optimal is this estimate? Computationally, it appears that for $k = 1$ this inequality is optimal. A few examples for $D = 1$:

Examples. (1) $p = 3$: The recursion operator T with companion polynomial $X^3 + yX - y^3$ and initial values $\{0, 1, y\}$ appears to achieve $N_T(y^n) = c_{3,2}(n)$ infinitely often.

(2) $p = 5$: The recursion operator T with companion polynomial

$$X^5 + 3yX^3 + y^2X^2 + 3y^3X + 4y^5$$

and initial values $[0, 1, y, y^2, y^3]$ appears to achieve $N_T(y^n) = c_{5,4}(n)$ for "most" n; every counterexample n has 0s in its base-5 expansion.

(3) $p = 7$: The recursion operator T with companion polynomial

$$X^7 + 3y^2X^4 + 6y^3X^3 + 5y^4X^2 + 3y^5X + 6y^7$$

appears to achieve $N_T(y^n) = c_{7,6}(n)$ for most n. For $n < 1000$, there are only 36 counterexamples, and $c_{7,6}(n) - N_T(y^n) \leq 3$ for each one.

(4) $p = 11$. The recursion operator T with companion polynomial

$$P_T = X^{11} + 6yX^9 + 2y^2X^8 + 3y^3X^7 + 6y^4X^6 + 8y^6X^4 + y^8X^2 + 9y^9X + 10y^{11}$$

appears to achieve $N_T(y^n) = c_{11,10}(n)$ for most n. For $n < 1000$, there are only 8 counterexamples, and $N_T(y^n) = c_{11,10}(n) - 1$ for each one.

The estimate appears not to be optimal as soon as $k \geq 2$.

Acknowledgements

I would also like to thank Paul Monsky for many illuminating discussions on the topic of mod p modular forms and Hecke algebras, and Kiran Kedlaya for helpful comments on an earlier version of Theorem 1. Part of the writing of this document was carried out at the Max Planck Institute for Mathematics in Bonn, and I am grateful for their hospitality. Extensive computations related to this project were coded and carried out in SAGE [≥ 2018].

References

[Al Hajj Shehadeh et al. 2009] H. Al Hajj Shehadeh, S. Jaafar, and K. Khuri-Makdisi, "Generating functions for Hecke operators", *Int. J. Number Theory* **5**:1 (2009), 125–140. MR Zbl

[Allouche and Shallit 1992] J.-P. Allouche and J. Shallit, "The ring of k-regular sequences", *Theoret. Comput. Sci.* **98**:2 (1992), 163–197. MR Zbl

[Bellaïche 2012] J. Bellaïche, "Pseudodeformations", *Math. Z.* **270**:3-4 (2012), 1163–1180. MR Zbl

[Bellaïche and Khare 2015] J. Bellaïche and C. Khare, "Level 1 Hecke algebras of modular forms modulo p", *Compos. Math.* **151**:3 (2015), 397–415. MR Zbl

[Buzzard and Calegari 2005] K. Buzzard and F. Calegari, "Slopes of overconvergent 2-adic modular forms", *Compos. Math.* **141**:3 (2005), 591–604. MR Zbl

[Conrad 2016] K. Conrad, "Solving linear recursions over all fields", preprint, 2016, Available at https://tinyurl.com/kconrecu.

[Deo 2017] S. V. Deo, "Structure of Hecke algebras of modular forms modulo p", *Algebra Number Theory* **11**:1 (2017), 1–38. MR Zbl

[Emerton 2011] M. Emerton, "p-adic families of modular forms (after Hida, Coleman, and Mazur)", exposé 1013, pp. 31–61 in *Séminaire Bourbaki,* 2009/2010, Astérisque **339**, Soc. Math. France, Paris, 2011. MR Zbl

[Gerbelli-Gauthier 2016] M. Gerbelli-Gauthier, "The order of nilpotence of Hecke operators mod 2: a new proof", *Res. Number Theory* **2** (2016), art. id. 7. MR Zbl

[Gouvêa and Mazur 1998] F. Q. Gouvêa and B. Mazur, "On the density of modular representations", pp. 127–142 in *Computational perspectives on number theory* (Chicago, 1995), edited by D. A. Buell and J. T. Teitelbaum, AMS/IP Stud. Adv. Math. **7**, Amer. Math. Soc., Providence, RI, 1998. MR Zbl

[Jochnowitz 1982] N. Jochnowitz, "A study of the local components of the Hecke algebra mod l", *Trans. Amer. Math. Soc.* **270**:1 (1982), 253–267. MR Zbl

[Medvedovsky 2015] A. Medvedovsky, *Lower bounds on dimensions of mod-p Hecke algebras: the nilpotence method*, Ph.D. thesis, Brandeis University, 2015, Available at https://search.proquest.com/docview/1687761247.

[Medvedovsky ≥ 2018] A. Medvedovsky, "Lower bounds on dimensions of mod-p Hecke algebras in the genus-zero case", in preparation.

[Nicolas and Serre 2012a] J.-L. Nicolas and J.-P. Serre, "Formes modulaires modulo 2: l'ordre de nilpotence des opérateurs de Hecke", *C. R. Math. Acad. Sci. Paris* **350**:7-8 (2012), 343–348. MR Zbl

[Nicolas and Serre 2012b] J.-L. Nicolas and J.-P. Serre, "Formes modulaires modulo 2: structure de l'algèbre de Hecke", *C. R. Math. Acad. Sci. Paris* **350**:9-10 (2012), 449–454. MR Zbl

[SAGE ≥ 2018] SAGE, "SAGE software", Available at http://www.sagemath.org.

[Serre 1986] J.-P. Serre, *Œuvres, III: 1972-1984*, Springer, 1986. MR Zbl

[Swinnerton-Dyer 1973] H. P. F. Swinnerton-Dyer, "On l-adic representations and congruences for coefficients of modular forms", pp. 1–55 in *Modular functions of one variable, III* (Antwerp, 1972), edited by W. Kuyk and J.-P. Serre, Lecture Notes in Math. **350**, Springer, 1973. Correction in *Modular functions of one variable, IV*, Lecture Notes in Math. **476**, Springer, 1975, pp. 149. MR Zbl

[Tate 1994] J. Tate, "The non-existence of certain Galois extensions of $\mathbb{Q}$ unramified outside 2", pp. 153–156 in *Arithmetic geometry* (Tempe, AZ, 1993), edited by N. Childress and J. W. Jones, Contemp. Math. **174**, Amer. Math. Soc., Providence, RI, 1994. MR Zbl

medved@mpim-bonn.mpg.de   *Max-Planck-Institut für Mathematik, Bonn, Germany*

Algebra & Number Theory

msp.org/ant

Algebra & Number Theory

Volume 12   No. 3   2018