HIT POLYNOMIALS AND CONJUGATION IN THE STEENROD ALGEBRA AND ITS DUAL

JUDITH H. SILVERMAN

ABSTRACT. Let $\mathcal{A}^*$ be the mod-2 Steenrod algebra of cohomology operations and χ its canonical antiautomorphism. For all positive integers f and k, we show that the excess of the element $\chi[Sq(2^{k-1}f)\cdot Sq(2^{k-2}f)\cdots Sq(2f)\cdot Sq(f)]$ is $(2^k-1)\mu(f)$, where $\mu(f)$ denotes the minimal number of summands in any representation of f as a sum of numbers of the form 2^i-1. We also interpret this result in purely combinatorial terms. In so doing, we express the Milnor basis representation of the products $Sq(a_1)\dots Sq(a_n)$ and $\chi[Sq(a_1)\dots Sq(a_n)]$ in terms of the cardinalities of certain sets of matrices.

For $s\geq 1$, let $\mathbb{P}_s=\mathbb{F}_2[x_1,\dots,x_s]$ be the mod-2 cohomology of the s-fold product of $\mathbb{R}P^\infty$ with itself, with its usual structure as an $\mathcal{A}^*$-module. A polynomial $P\in\mathbb{P}_s$ is *hit* if it is in the image of the action $\overline{\mathcal{A}^*}\otimes\mathbb{P}_s\longrightarrow\mathbb{P}_s$, where $\overline{\mathcal{A}^*}$ is the augmentation ideal of $\mathcal{A}^*$. We prove that if the integers e, f, and k satisfy $e<(2^k-1)\mu(f)$, then for any polynomials E and F of degrees e and f respectively, the product $E\cdot F^{2^k}$ is hit. This generalizes a result of Wood, conjectured by Peterson, and proves a conjecture of Singer and Silverman.

1. INTRODUCTION

Let $\mathcal{A}^*$ be the mod-2 Steenrod algebra of cohomology operations, multiplicatively generated by the Steenrod squares $Sq(f)$, $1\leq f<\infty$. If $\Theta\in\mathcal{A}^*$, we denote by $\hat{\Theta}$ the image of Θ under the canonical antiautomorphism χ of $\mathcal{A}^*$.

For $s\geq 1$, let $\mathbb{P}_s=\mathbb{F}_2[x_1,\dots,x_s]$ be the mod-2 cohomology of the s-fold product of $\mathbb{R}P^\infty$ with itself, with its usual structure as an $\mathcal{A}^*$-module. The *excess* of an element $\Theta\in\mathcal{A}^*$ is given by $\mathrm{ex}(\Theta)=\min\{s\,:\,\Theta(x_1x_2\cdots x_s)\neq 0\in\mathbb{P}_s\}$. In particular, for $f\geq 1$ we have $\mathrm{ex}(Sq(f))=f$ and $\mathrm{ex}(\hat{Sq}(f))=\mu(f)$ [Kra71], where $\mu(f)$ denotes the minimal number of summands in any representation of f as a sum of numbers of the form (2^i-1).

Given positive integers f and k, define $S[k;f]=Sq(2^{k-1}f)\cdot Sq(2^{k-2}f)\cdots Sq(2f)\cdot Sq(f)$ as in [Sil95b]. Our first theorem generalizes the above expression for the excess of $\hat{Sq}(f)=\hat{S}[1;f]$ to general k. The theorem is stated in full in Section 9; a simplified version reads:

Theorem 1.1. *Let f be a positive integer. Then for all positive integers k we have*

$$ex(\hat{S}[k;f]) \;=\; (2^k-1)\mu(f).$$

This result was conjectured in [Sil93] and proven for special cases there and in [Sil95b]. The present proof relies on the "stripping" action of the dual Steenrod algebra $\mathcal{A}_*$ on $\mathcal{A}^*$ and requires a formula for the conjugates of certain products in $\mathcal{A}_*$. A purely combinatorial interpretation of Theorem 1.1 is given in Section 3.

Date: October 11, 1996.

1991 *Mathematics Subject Classification.* Primary 55S05, 55S10; Secondary 57T05, 11P81.

Key words and phrases. Steenrod algebra, antiautomorphism, stripping, excess, hit polynomials.

A polynomial $P \in \mathbb{P}_s$ is *hit* if it is in the image of the action $\overline{\mathcal{A}^*} \otimes \mathbb{P}_s \longrightarrow \mathbb{P}_s$, where $\overline{\mathcal{A}^*}$ is the augmentation ideal of $\mathcal{A}^*$. The set of hit polynomials is linked to various important questions in algebraic topology and representation theory. For example, it is related to the group $\mathrm{Ext}^s_{\mathcal{A}^*}(\mathbb{F}_2, \mathbb{F}_2)$, which appears in the E_2–term of the Adams spectral sequence for the stable homotopy of spheres [Sin89]. Information about hit polynomials also yields information about $H^*(BO)$, $H^*(MO)$, and the cobordism classes of closed manifolds [Pet89]. Moreover, the set of hit polynomials is related to the study of simple representations of the general linear group $GL(s, \mathbb{F}_2)$. The action of this group on $\mathbb{P}_s$ is compatible with the $\mathcal{A}^*$ action, so that the set of hit elements is closed under the action of $GL(s, \mathbb{F}_2)$. A theorem in modular representation theory says that all simple representations of $GL(s, \mathbb{F}_2)$ occur in composition series of $\mathbb{P}_s$ in some degree [Mit85], and Wood has shown that the same is true after dividing out the $\mathcal{A}^*$—action [Woo89a].

In his elegant proof of a conjecture of Peterson [Pet87] concerning hit polynomials, Wood makes use of the excess formula for $\hat{Sq}(f)$ which is the case $k = 1$ of Theorem 1.1 [Woo89b]. Fueled by the general case of the theorem, his argument implies the following result, which was conjectured in [SS95]:

Theorem 1.2. *Let P be a polynomial of the form $E \cdot F^{2^k}$ for some polynomials E and F of degrees e and f respectively, and suppose that $e < (2^k - 1)\mu(f)$. Then P is hit.*

The reader is referred to [Sin91], [Mon94] and [Sil95a] for further discussion of results related to Theorem 1.2.

2. Preliminaries

2.1. **Sequences.** Denote by $\mathcal{S}$ the monoid of sequences of non–negative integers almost all of which are 0; write $0_{\mathcal{S}}$ for the trivial element and $B(j)$ for the sequence whose i-th coordinate is $\delta_{i,j}$. If $T = (t_1, t_2, \dots)$ has $t_i = 0$ for $i > L$, we identify T with $(t_1, \dots, t_L)$. Suppose that $T = (t_1, \dots, t_L)$ with $t_L \neq 0$. Then the *length* of T is $\mathrm{len}(T) = L$, the *degree* of T is $|T| = \sum_{i=1}^{L}(2^i - 1)t_i$, and the *excess* of T is $\mathrm{ex}(T) = \sum_{i=1}^{L} t_i$. Following [Mon], we say that $T \succ_R T'$ if T is greater than T' in the right–lexicographic order.

It will be convenient to write $\mathcal{S}^* = \mathcal{S} \cup \{*\}$, where $*$ satisfies $x + * = *$ for all $x \in \mathcal{S}^*$.

2.2. **Milnor bases.** The Milnor bases of the Steenrod algebra $\mathcal{A}^*$ and its dual $\mathcal{A}_*$ are both indexed by elements of $\mathcal{S}$ [Mil58]. For $T \in \mathcal{S}$, we write $M[T]$ for the corresponding basis element of $\mathcal{A}^*$ and $\xi[T]$ for the basis element of $\mathcal{A}_*$. In particular, $M[(t)]$ is the Steenrod square $Sq(t)$ for $t \geq 0$. The dual algebra is in fact a polynomial algebra on the generators $\xi_j = \xi[B(j)]$, so that if $T = (t_1, \dots, t_L)$, then $\xi[T] = \prod_{i=1}^{L} \xi_i^{t_i}$. We further set $M[*] = 0$ and $\xi[*] = 0$, and adopt the convention that $M[R] = 0$ and $\xi[R] = 0$ if R is a finite sequence of integers at least one of which is negative.

The definitions of length, degree, excess, and right–lexicographic order for elements of $\mathcal{S}$ induce similar definitions for elements of both bases and hence for all homogeneous elements of $\mathcal{A}^*$ and $\mathcal{A}_*$, where the length (resp. excess) of a sum of basis elements is determined by the maximal length (resp. minimal excess) of the summands, and sums are compared in the right–lexicographic order by comparing their terms in descending order. It is readily

verified that this definition of excess in $\mathcal{A}^*$ agrees with the definition of Section 1 based on the $\mathcal{A}^*$-action on polynomial rings (cf. [Kra71]).

Example. [Kra71] The function $\mu(f)$ of Section 1 may now be described as the least excess of any Milnor basis element of degree f or alternatively as the excess of the sum of all Milnor basis elements of that degree. Since this sum is equal to $\hat{Sq}(f)$ [Mil58], we find that $\mathrm{ex}(\hat{Sq}(f)) = \mu(f)$ as stated in Section 1.

Let $T \in \mathcal{S}$ be a sequence, $\Theta \in \mathcal{A}^*$. We say $M[T] \in \Theta$ if $M[T]$ occurs with non–trivial coefficient in the Milnor basis representation of Θ. In other words, $M[T] \in \Theta \iff \langle \xi[T],\, \Theta\rangle = 1$, where $\langle\ ,\ \rangle$ is the pairing relative to the Milnor bases of $\mathcal{A}_*$ and $\mathcal{A}^*$. In what follows, we use this criterion in the form

$$M[T] \in \hat{\Theta} \iff \langle \hat{\xi}[T],\, \Theta\rangle = 1, \tag{1}$$

where $\hat{\xi}[T]$ denotes the image of $\xi[T]$ under the canonical antiautomorphism χ of $\mathcal{A}_*$.

2.3. **Stripping.** Let Δ^* be the diagonal map of the Hopf algebra $\mathcal{A}^*$. There is a natural action of the dual Hopf algebra $\mathcal{A}_*$ on $\mathcal{A}^*$, in which each $\xi \in \mathcal{A}_*$ acts via

$$D(\xi)\ :\ \mathcal{A}^* \xrightarrow{\Delta^*} \mathcal{A}^* \otimes \mathcal{A}^* \xrightarrow{1\otimes\langle \xi,\ \rangle} \mathcal{A}^*.$$

Alternatively, $D(\xi)$ satisfies

$$\langle \xi\cdot\psi,\, \Theta\rangle \ \ =\ \ \langle \psi,\, D(\xi)\,\Theta\rangle \tag{2}$$

for all $\psi \in \mathcal{A}_*$ and $\Theta \in \mathcal{A}^*$. We refer to $D(\xi)$ as the operation of *stripping by* ξ. For typographical convenience, we follow [WW96] in expressing $D(\xi)$ in cap–product notation:

$$\xi \cap \Theta \ \stackrel{\mathrm{def}}{=}\ D(\xi)\Theta$$

for all $\xi \in \mathcal{A}_*$ and $\Theta \in \mathcal{A}^*$. Thus (2) becomes $\langle \xi\cdot\psi,\, \Theta\rangle = \langle \xi,\, \psi\cap\Theta\rangle$, an identity we use in the form

$$\langle \xi[U+V],\, \hat{\Theta}\rangle = \ \langle \hat{\xi}[U+V],\, \Theta\rangle \ = \langle \hat{\xi}[U],\, \hat{\xi}[V]\cap\Theta\rangle \tag{3}$$

for all $U, V \in \mathcal{S}$ and $\Theta \in \mathcal{A}^*$. In particular, if $|T| = |\Theta|$ we have $\langle \hat{\xi}[T],\, \Theta\rangle = \langle 1,\, \hat{\xi}[T]\cap\Theta\rangle$, so that

$$M[T] \in \hat{\Theta} \iff \hat{\xi}[T]\cap\Theta = 1. \tag{4}$$

The following results are proven in [CW94] and [Sil95b], respectively.

Proposition 2.1. 1. *Let $\Theta \in \mathcal{A}^*$. If $j > \mathit{len}(\Theta)$, then $\xi_j \cap \Theta = 0$.*
2. *For all $k, f \geq 1$, we have $\xi_k \cap S[k; f] = S[k; f-1]$.*

Proposition 2.2. *For all $j,\ k,\ f \geq 1$, we have*
1. $\hat{\xi}_j \cap Sq(f) = Sq(f-(2^j-1))$.
2. $\hat{\xi}_j \cap S[k; f] = S[k-1; 2f]\cdot[\hat{\xi}_j \cap Sq(f)] = S[k-1; 2f]\cdot Sq(f-(2^j-1))$.

Observation 2.3. Part 1 of Proposition 2.2, which may be restated as "$D(\hat{\xi}_j) = D(\xi_1^{2^j-1})$ when restricted to elements of $\mathcal{A}^*$ of length 1", follows from Part 1 of Proposition 2.1 and the fact that $\hat{\xi}_j \equiv \xi_1^{2^j-1}$ modulo elements of $\mathcal{A}_*$ of length > 1.

2.4. **Application.** As an illustration of the technique of stripping, we give a new proof of a theorem from [Sil96] which we will need in the proof of Theorem 1.1.

Theorem 2.4. [Sil96] Let f, L, and k be positive integers such that $f < 2^L - 1$. Then any sequence T with $M[T] \in \hat{S}[k; f]$ has length $< L$.

Proof. Suppose T is a sequence of length $\geq L$, so that $t_j \geq 1$ for some $j \geq L$. By (3) and Part 2 of Proposition 2.2, we have

$$\begin{aligned} \langle \hat{\xi}[T],\, S[k; f] \rangle &= \langle \hat{\xi}[T - B(j)],\, \hat{\xi}_j \cap S[k; f] \rangle \\ &= \langle \hat{\xi}[T - B(j)],\, S[k-1; f] \cdot [\hat{\xi}_j \cap Sq(f)] \rangle. \end{aligned}$$

But $f < 2^j - 1$, so $\hat{\xi}_j \cap Sq(f) = 0$ by Part 1 of Proposition 2.2. Consequently the inner product $\langle \hat{\xi}[T],\, S[k; f] \rangle = 0$, and we conclude from (1) that $M[T] \notin \hat{S}[k; f]$. ∎

3. Computing products in $\mathcal{A}^*$

In this section we use the language of stripping to express the question of whether a Milnor basis element appears in $\chi[Sq(a_1) \cdot \ldots \cdot Sq(a_n)]$ in terms of the cardinality of a set of matrices (cf. Milnor's product formula for $M(R) \cdot M(S)$ [Mil58]). We then interpret Theorem 1.1 in combinatorial terms involving matrices and partitions. The rest of the paper does not draw upon the material in this section.

We begin by giving an explicit description of the effect of stripping products of Steenrod squares by $\hat{\xi}_i$. The following property of stripping operators is discussed in [Sil95b]: Let $\xi \in \mathcal{A}_*$ and write $\phi_* \xi = \sum \xi' \otimes \xi''$, where ϕ_* is the diagonal map of $\mathcal{A}_*$. Then

$$\hat{\xi} \cap (\theta_1 \theta_2) \;=\; \sum (\hat{\xi}'' \cap \theta_1) \cdot (\hat{\xi}' \cap \theta_2) \tag{5}$$

for all $\theta_1,\ \theta_2 \in \mathcal{A}^*$.

Definition. Fix integers $i, n > 1$. A $\hat{\xi}_i$*–strip vector of length* n is a sequence $(v_1, \ldots, v_n)$ in which the nonzero elements form exactly the sequence

$$(2^{e_1} - 1, 2^{e_2} - 2^{e_1}, \ldots, 2^i - 2^{e_m})$$

for some m and some sequence $1 < e_1 < e_2 \ldots < e_m < i$. For example, the $\hat{\xi}_2$–strip vectors of length 3 are (3,0,0), (0,3,0), (0,0,3), (1,2,0), (1,0,2), and (0,1,2). Write $\hat{\mathcal{V}}_{n,i}$ for the set of $\hat{\xi}_i$–strip vectors of length n. An inductive argument based on (5) and Part 1 of Proposition 2.2 yields the following:

Proposition 3.1. *Let i and $a_1, \ldots, a_n$ be non–negative integers. Then*

$$\hat{\xi}_i \cap [Sq(a_1) \cdot \ldots \cdot Sq(a_n)] \;=\; \sum_{V \in \hat{\mathcal{V}}_{n,i}} Sq(a_1 - v_1) \cdot \ldots \cdot Sq(a_n - v_n). \tag{6}$$

Note that the product on the left is not required to be admissible, nor are those on the right claimed to be.

We proceed to interpret $\langle \hat{\xi}[T],\, Sq(a_1) \cdot \ldots \cdot Sq(a_n) \rangle$ as the parity of the cardinality of a certain set of matrices.

Definition. An $(\mathrm{ex}(T) \times n)$-matrix $R = (r_{ij})$ is a $\{\xi[T];\ Sq(a_1) \cdot \ldots \cdot Sq(a_n)\}$–matrix if the entries satisfy the following conditions:

- For $1 \leq i \leq L$, the rows numbered $(\sum_{m=1}^{i-1} t_m) + 1$ through $\sum_{m=1}^{i} t_m$ are $\hat{\xi}_i$–strip vectors.
- For $1 \leq j \leq n$, the column sum $\sum_{i=1}^{L} r_{ij}$ is equal to a_j.

For example,

$$\begin{pmatrix} 0 & 1 \\ 3 & 4 \\ 7 & 0 \end{pmatrix}$$

is a $\{\hat{\xi}[(1,0,2)];\ Sq(10) \cdot Sq(5)\}$–matrix.

The following characterization of the summands of $\chi[Sq(a_1) \cdot \ldots \cdot Sq(a_n)]$ follows from (3), (4) and Proposition 3.1 by induction on $\mathrm{ex}(T)$:

Proposition 3.2. *Let $a_1, \ldots, a_n$ be positive integers, and suppose that $T \in \mathcal{S}$ with $|T| = \sum a_j$. Then $M[T]$ appears in $\chi[Sq(a_1) \cdot \ldots \cdot Sq(a_n)] \Longleftrightarrow$ there are an odd number of $\{\hat{\xi}[T];\ Sq(a_1) \cdot \ldots \cdot Sq(a_n)\}$–matrices.*

A similar characterization of the summands of $Sq(a_1) \cdot \ldots \cdot Sq(a_n)$ may be obtained by considering matrices whose rows are ξ_i–strip vectors as described in [CW94] or [Sil95b].

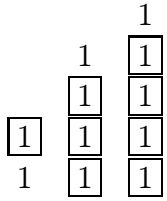
Proposition 3.2 makes possible the following combinatorial interpretation of Theorem 1.1 as stated in Section 1:

Theorem 1.1, combinatorial version. *Let f and k be positive integers, and suppose that $T \in \mathcal{S}$ with $|T| = (2^k - 1)f$. If $\sum t_i < (2^k - 1)\mu(f)$, then there are an even number of $\{\hat{\xi}[T];\ S[k; f]\}$–matrices.*

Theorem 2.4 of Section 2.4 may be restated as follows:

Theorem 2.4, combinatorial version. *With notation as above, suppose that T is of length L. If $f < 2^L - 1$, then there are an even number of $\{\hat{\xi}[T];\ S[k; f]\}$–matrices.*

In the case $k = 2$, the restatements of Theorems 1.1 and 2.4 may in turn be restated as follows: Given $f \geq 1$ and an L–tuple $G = (g_1, \ldots, g_L)$ of positive integers such that $\sum_{i=1}^{L}(2^{g_i} - 1) = 3f$, define $\mathcal{H}(G)$ to be the set of L–tuples $H = (h_1, \ldots, h_L)$ for which $0 \leq h_i \leq g_i$ for all i and $\sum_{i=1}^{L}(2^{h_i} - 1) = 2f$. Elements of $\mathcal{H}(G)$ may be regarded as illustrated below in the case $f = 7$, $G = (1, 2, 2, 3, 3)$, and $H = (0, 1, 2, 3, 2)$. Here the rows of 1's represent the numbers $2^{g_i} - 1$ in binary notation, and the boxed portions represent the corresponding numbers $2^{h_i} - 1$.



Theorems 1.1 and 2.4, case $k = 2$. *Let $f \geq 1$. Then with notation as above, the cardinality of $\mathcal{H}(G)$ is even if $L < 3\mu(f)$ or if $2^{g_m} - 1 > f$ for some m.*

4. Motivation for the Proof of Theorem 1.1

Recall from Sections 1 and 2.2 that $\mu(f)$ is the minimal number of summands in any representation of f as a sum of numbers of the form $(2^i - 1)$ or, equivalently, the least excess of any Milnor basis element of degree f. The main result of [Sil95b], quoted below as Theorem 9.1, implies that $\mathrm{ex}(\hat{S}[k; f]) \leq (2^k - 1)\mu(f)$ for all k and f. Consequently to prove the simplified version of Theorem 1.1 as stated in Section 1, it suffices to establish the reverse inequality. In order to motivate the computations to follow, we now give a stripping proof of the obvious inequality $\mathrm{ex}(\hat{Sq}(f)) \geq \mu(f)$, for the appreciation of which the reader is advised to ignore the second interpretation of $\mu(f)$ and to regard this function simply as one satisfying $\mu(0) = 0$, $\mu(1) = 1$, and

$$\mu(f) \;\leq\; \mu(f - (2^i - 1)) + 1 \tag{7}$$

whenever $f \geq 2^i - 1$. We then indicate the difficulties in extending the argument to show that $\mathrm{ex}(\hat{S}[k; f]) \geq (2^k - 1)\mu(f)$ for $k > 1$. This exercise is intended to motivate and illustrate techniques rather than to sketch the path actually followed in proving Theorem 1.1.

Proof. Since $\hat{Sq}(1) = Sq(1)$, the claim is true for $f = 1$. Now fix $f > 1$ and assume inductively that the result is true for $f' < f$. Let $M[T]$ be a Milnor basis element $\in \hat{Sq}(f)$ and choose i for which $t_i \geq 1$. By (3) and Part 1 of Proposition 2.2, we have

$$\begin{aligned} 1 = \langle \hat{\xi}[T],\, Sq(f) \rangle &= \langle \hat{\xi}[T - B(i)],\, \hat{\xi}_i \cap Sq(f) \rangle \\ &= \langle \hat{\xi}[T - B(i)],\, Sq(f - (2^i - 1)) \rangle, \end{aligned}$$

and consequently $M[T - B(i)] \in \hat{Sq}(f - (2^i - 1))$. By the inductive hypothesis,

$$(\ \mathrm{ex}(T) - 1 = \ )\quad \mathrm{ex}(T - B(i)) \;\geq \mu(f - (2^i - 1)),$$

and so $\mathrm{ex}(T) \geq \mu(f)$ by (7). ∎

This argument rests on the unremarkable fact that if $M[T] \in \hat{Sq}(f)$, then $t_i \geq 1$ for some i. The analogue of this starting point for general k, namely that $t_i \geq 2^k - 1$ for some i, is not the case for arbitrary summands of $\hat{S}[k; f]$. As we show in Proposition 8.2, however, this condition does hold for the sequence W whose associated Milnor basis element $M[W]$ is maximal in right–lexicographic order among Milnor basis elements of minimal excess appearing in $\hat{S}[k; f]$. We therefore have

$$1 = \langle \hat{\xi}[W],\, S[k; f] \rangle \;=\; \langle \hat{\xi}[W - (2^k - 1)B(i)],\, \hat{\xi}_i^{2^k - 1} \cap S[k; f] \rangle. \tag{8}$$

In order to proceed with the argument, one needs an analogue of Part 1 of Proposition 2.2 to describe the effect of stripping Steenrod operations of length k by $\hat{\xi}_i^{2^k - 1}$ (cf. Observation 2.3). We accomplish this in Section 5 by finding a formula for $\hat{\xi}_i^{2^k - 1}$ modulo elements of $\mathcal{A}_*$ of length $> k$. This formula involves $\xi_k^{2^i - 1}$ along with certain error terms E_n. In Section 7, we further show that the defining property of W implies that the error terms

$\langle \hat{\xi}[W-(2^k-1)B(i)],\, E_n \cap S[k;f]\rangle$ arising from substituting this expression for $\hat{\xi}_i^{2^k-1}$ in (8) all vanish. Consequently, we find that

$$\begin{aligned} 1 &= \langle \hat{\xi}[W-(2^k-1)B(i)],\, \xi_k^{2^i-1} \cap S[k;f]\rangle \\ &= \langle \hat{\xi}[W-(2^k-1)B(i)],\, S[k;f-(2^i-1)]\rangle, \end{aligned}$$

the latter equality holding by Part 2 of Proposition 2.1. Therefore $M[W-(2^k-1)B(i)]$ appears in $\hat{S}[k;f-(2^i-1)]$, and the inductive argument goes through as before; we conclude that $\mathrm{ex}(\hat{S}[k;f]) \geq (2^k-1)\mu(f)$. As observed at the beginning of this section, this inequality confirms the simplified version of Theorem 1.1.

We include this argument to motivate the conjugation formulae of Section 5 and to illustrate the technique of manipulating these conjugation formulae while stripping, which is central to our proof of Proposition 8.2 in Sections 7 and 8 below. Once established, however, this proposition makes possible a more straightforward proof of Theorem 1.1 than the one indicated above. In Section 9, where the theorem is stated in its entirety and proved, we invoke Proposition 8.2 in its full strength to establish not only that W has an entry $\geq 2^k-1$ but that each entry of W is divisible by 2^k-1. This along with Theorem 9.1 enables us to bypass the inductive argument on f and, for fixed f, to prove the result for general k directly from the result for $k=1$.

5. Conjugation formula

5.1. **Formula for $\hat{\xi}_i$.** Let $\mathcal{S}(k)$ be the symmetric group with identity Id^k acting on $\{0,1,\ldots,k-1\}$. For $\tau \in \mathcal{S}(k)$ and $i \geq 0$, define

$$Z_i(k;\tau) = \begin{cases} \sum_{j=0}^{k-1} 2^j B(i+\tau(j)-j) & i+\tau(j)-j \geq 0 \text{ for all } j \\ * & \text{else} \end{cases} \tag{9}$$

and

$$X_i(k;\tau) = \xi[Z_i(k;\tau)] = \begin{cases} \prod_{j=0}^{k-1} \xi_{i+\tau(j)-j}^{2^j} & i+\tau(j)-j \geq 0 \text{ for all } j \\ 0 & \text{else.} \end{cases}$$

Observation 5.1.
- $Z_i(k;\mathrm{Id}^k) = (2^k-1)B(i)$, so that $X_i(k;\mathrm{Id}^k) = \xi_i^{2^k-1}$.
- If $X_i(k;\tau) \neq 0$, then $|X_i(k;\tau)| = (2^k-1)(2^i-1)$ independently of τ.
- Any non-$*$ $Z_i(k;\tau)$ with $\tau \neq \mathrm{Id}^k$ is strictly $\succ_R$ and no greater in excess than $(2^k-1)B(i) = Z_i(k;\mathrm{Id}^k)$.

Finally, define

$$\mathcal{X}_i(k) = \sum_{\tau \in \mathcal{S}(k)} X_i(k;\tau)$$

and note that

$$\mathcal{X}_i(1) = X_i(1;\mathrm{Id}^1) = \xi_i. \tag{10}$$

Lemma 5.2. *For $k \geq 1$, we have $\hat{\xi}_k = \mathcal{X}_1(k)$.*

Proof. The proof is by induction on k and makes use of Milnor's recursive formula for the canonical antiautomorphism [Mil58]: $\hat{\xi}_1 = \xi_1$ and $\hat{\xi}_k = \sum_{l=0}^{k-1} \xi_{k-l}^{2^l} \hat{\xi}_l$. For $k = 1$, we have $\mathcal{X}_1(1) = X_1(1; \mathrm{Id}^1) = \xi_1 = \hat{\xi}_1$, and for $k = 2$ we have $\mathcal{X}_1(2) = X_1(2; \mathrm{Id}^2) + X_1(2; (0,1)) = \xi_1^3 + \xi_2^{2^0}\xi_0^{2^1} = \hat{\xi}_2$. Suppose now that $k \geq 3$ and that the formula holds for $k' < k$. Note that if $X_1(k;\tau) \neq 0$, we must have $\tau(j) \geq j-1$ for all j. Consequently if $X_1(k;\tau) \neq 0$ and l satisfies $\tau(l) = k-1$, we must have $\tau(k-1) = k-2$, $\tau(k-2) = k-3, \ldots,$ $\tau(l+1) = l$, so that τ has cycle–decomposition of the form $(k-1, k-2, \ldots, l)\sigma$ for some $\sigma \in \mathcal{S}(l)$. In this case, we have $X_1(k;\tau) = \xi_{k-l}^{2^l} X_1(k;\sigma)$.

Let $\mathcal{S}_l(k) = \{\tau \in \mathcal{S}(k) \ : \ \tau(l) = k-1\}$. Then

$$\begin{aligned} \mathcal{X}_1(k) &= \sum_{l=0}^{k-1} \sum_{\tau \in \mathcal{S}_l(k)} X_1(k;\tau) = \sum_{l=0}^{k-1} \xi_{k-l}^{2^l} \sum_{\sigma \in \mathcal{S}(l)} X_1(k;\sigma) \\ &\overset{\text{ind}}{=} \sum_{l=0}^{k-1} \xi_{k-l}^{2^l} \hat{\xi}_l = \hat{\xi}_k. \end{aligned}$$

∎

5.2. **General formula.** Equation 10 and Lemma 5.2 suggest the following generalization, which may be verified for $i = 2$ and all k by an argument similar to the proof of Proposition 5.5 below:

Conjecture 5.3. *$\hat{\mathcal{X}}_k(i) = \mathcal{X}_i(k)$ for all $i, k \geq 1$.*

For our purposes, however, we require a different generalization of Lemma 5.2.

Definition. For $k \geq 1$, let $\mathcal{I}(k)$ denote the set of non–decreasing sequences $(i_0, i_1, \ldots, i_{k-1})$ of positive integers. Now for $\tau \in \mathcal{S}(k)$ and $I \in \mathcal{I}(k)$, define

$$Z_I(k;\tau) = \begin{cases} \sum_{j=0}^{k-1} 2^j B(i_{\tau(j)} + \tau(j) - j) & i_{\tau(j)} + \tau(j) - j \geq 0 \text{ for all } j \\ * & \text{else} \end{cases}$$

$$X_I(k;\tau) = \xi[Z_I(k;\tau)] = \begin{cases} \prod_{j=0}^{k-1} \xi_{i_{\tau(j)}+\tau(j)-j}^{2^j} & i_{\tau(j)} + \tau(j) - j \geq 0 \text{ for all } j \\ 0 & \text{else} \end{cases}$$

$$\mathcal{X}_I(k) = \sum_{\tau \in \mathcal{S}(k)} \xi[Z_I(k;\tau)] = \sum_{\tau \in \mathcal{S}(k)} X_I(k,\tau)\,.$$

We further define

$$P_I(k;\tau) = \begin{cases} \sum_{j=0}^{k-1} 2^{j+i_0} B(i_{\tau(j)} + \tau(j) - (j+i_0)) & i_{\tau(j)} + \tau(j) - (j+i_0) \geq 0 \text{ for all } j \\ * & \text{else} \end{cases}$$

$$R_I(k;\tau) = \xi[P_I(k;\tau)] = \begin{cases} \prod_{j=0}^{k-1} \xi_{i_{\tau(j)}+\tau(j)-(i_0+j)}^{2^{i_0+j}} & i_{\tau(j)} + \tau(j) - (j+i_0) \geq 0 \text{ for all } j \\ 0 & \text{else} \end{cases}$$

$$\mathcal{R}_I(k) = \sum_{\tau \in \mathcal{S}(k)} \xi[P_I(k;\tau)] = \sum_{\tau \in \mathcal{S}(k)} R_I(k;\tau)\,.$$

Observation 5.4.
- If $I = (i, \dots, i) \in \mathcal{I}(k)$ is a constant sequence, then $Z_I(k;\tau) = Z_i(k;\tau)$ as defined in (9). Moreover $P_I(k;\tau) = *$ for $\tau \neq \mathrm{Id}^k$, and consequently $\mathcal{R}_I(k) = R_I(k;\mathrm{Id}^k) = 1$.
- For general I, $Z_I(k;\mathrm{Id}^k) = \sum_{j=0}^{k-1} 2^j B(i_{\tau(j)})$.
- Any non-$*$ $Z_I(k;\tau)$ with $\tau \neq \mathrm{Id}^k$ is strictly $\succ_R$ and no greater in excess than $\sum_{j=0}^{k-1} 2^j B(i_{\tau(j)}) = Z_I(k;\mathrm{Id}^k)$.
- The degrees of the non–$*$ $Z_I(k;\tau)$ and $P_I(k;\tau)$ depend only on I and not on τ.
- If $I = (i_0, i_1, \dots, i_{k-1}) \in \mathcal{I}(k)$ and $i_0 > 1$, write $I[-1]$ for the sequence $(i_0 - 1, i_1 - 1, \dots, i_{k-1} - 1)$. Then $\mathcal{R}_I(k) = [\mathcal{R}_{I[-1]}(k)]^2$.

Proposition 5.5. *With notation as above, $\hat{\mathcal{X}}_I(k) \equiv \xi_k^{2^{i_0}-1} \cdot \hat{\mathcal{R}}_I(k)$ modulo elements of length $> k$.*

Proof. Fix k. The proof is by induction on i_0.

When $i_0 = 1$, we have

$$\xi_k \cdot \hat{\mathcal{R}}_I(k) \;=\; \sum_{\tau} \xi_k \hat{\xi}^{2^k}_{i_{\tau(k-1)}+\tau(k-1)-k} \prod_{j\neq k-1} \hat{\xi}^{2^{j+1}}_{i_{\tau(j)}+\tau(j)-(j+1)} \cdot \tag{11}$$

On the other hand, we have

$$\hat{\mathcal{X}}_I(k) \;=\; \sum_{\rho} \hat{\xi}^{2^0}_{i_{\rho(0)}+\rho(0)-0} \prod_{j\neq 0} \hat{\xi}^{2^j}_{i_{\rho(j)}+\rho(j)-j} \cdot$$

For each $\rho \in \mathcal{S}(k)$, define ρ' by

$$\rho'(l) \;=\; \begin{cases} \rho(0) & l = k-1 \\ \rho(l+1) & 0 \leq l \leq k-2. \end{cases} \tag{12}$$

Then

$$\hat{\mathcal{X}}_I(k) \;=\; \sum_{\rho'} \hat{\xi}_{i_{\rho'(k-1)}+\rho'(k-1)} \prod_{l\neq k-1} \hat{\xi}^{2^{l+1}}_{i_{\rho'(l)}+\rho'(l)-(l+1)} \cdot \tag{13}$$

Observe that by Milnor's formula,

$$\hat{\xi}_{i_{\rho'(k-1)}+\rho'(k-1)} \;\equiv\; \sum_{n=1}^{k} \xi_n \hat{\xi}^{2^n}_{i_{\rho'(k-1)}+\rho'(k-1)-n} \quad \text{modulo terms of length } > k. \tag{14}$$

From (13), we find that

$$\hat{\mathcal{X}}_I(k) \;\equiv\; \sum_{\rho'} \left(\sum_{n=1}^{k} \xi_n \hat{\xi}^{2^n}_{i_{\rho'(k-1)}+\rho'(k-1)-n} \right) \prod_{l\neq k-1} \hat{\xi}^{2^{l+1}}_{i_{\rho'(l)}+\rho'(l)-(l+1)} \cdot \tag{15}$$

We claim now that the sums in (11) and (15) agree. Indeed, their difference is

$$\sum_{\rho'}\sum_{n=1}^{k-1}\xi_n\hat{\xi}^{2^n}_{i_{\rho'(k-1)}+\rho'(k-1)-n}\prod_{l\neq k-1}\hat{\xi}^{2^{l+1}}_{i_{\rho'(l)}+\rho'(l)-(l+1)} =$$

$$\sum_{n=1}^{k-1}\sum_{\rho'}\hat{\xi}_n\hat{\xi}^{2^n}_{i_{\rho'(k-1)}+\rho'(k-1)-n}\hat{\xi}^{2^{(n-1)+1}}_{i_{\rho'(n-1)}+\rho'(n-1)-[(n-1)+1]}\prod_{l\neq n-1,\ k-1}\hat{\xi}^{2^{l+1}}_{i_{\rho'(l)}+\rho'(l)-(l+1)}\,, \qquad (16)$$

and it is easily verified that the summand in (16) associated to n and ρ' equals the term associated to n and ρ'', where

$$\rho''(l) \;=\; \begin{cases} \rho'(l) & l\neq n-1,\ k-1 \\ \rho'(n-1) & l=k-1 \\ \rho'(k-1) & l=n-1. \end{cases}$$

Consequently each term in (16) appears twice. We conclude that the sums in (11) and (15) are equal, and hence that $\xi_k\cdot\hat{\mathcal{R}}_I(k)$ and $\hat{\mathcal{X}}_I(k)$ are congruent modulo elements of length $>k$. This establishes the proposition for the case $i_0=1$.

The proof for general I is similar. Suppose that the proposition is true for the non–decreasing sequence $(i_0-1, i_1-1,\ldots,i_k-1)\stackrel{\text{def}}{=}I[-1]$. For typographical convenience, let $i=i_0$. By Observation 5.4, we have

$$\begin{aligned}
\xi_k^{2^i-1}\cdot\hat{\mathcal{R}}_I(k) &= [\xi_k^{2^{i-1}-1}\cdot\hat{\mathcal{R}}_{I[-1]}(k)]^2\cdot\xi_k \\
&\stackrel{\text{ind}}{\equiv} [\hat{\mathcal{X}}_{I[-1]}(k)]^2\cdot\xi_k \qquad\qquad \text{(modulo length } >k) \\
&= \xi_k\,[\sum_{\tau}\prod_{j=0}^{k-1}\hat{\xi}^{2^{j+1}}_{(i_{\tau(j)}-1)+\tau(j)-j}] \\
&= \sum_{\tau}\xi_k\,\hat{\xi}^{2^k}_{(i_{\tau(k-1)}-1)+\tau(k-1)-(k-1)}\prod_{j\neq k-1}\hat{\xi}^{2^{j+1}}_{(i_{\tau(j)}-1)+\tau(j)-j} \\
&= \sum_{\tau}\xi_k\,\hat{\xi}^{2^k}_{i_{\tau(k-1)}+\tau(k-1)-k}\prod_{j\neq k-1}\hat{\xi}^{2^{j+1}}_{i_{\tau(j)}+\tau(j)-(j+1)}. \qquad (17)
\end{aligned}$$

On the other hand,

$$\begin{aligned}
\hat{\mathcal{X}}_I(k) &= \sum_{\rho}\prod_{j=0}^{k-1}\hat{\xi}^{2^j}_{i_{\rho(j)}+\rho(j)-j} \\
&= \sum_{\rho}\hat{\xi}^{2^0}_{i_{\rho(0)}+\rho(0)}\prod_{j\neq 0}\hat{\xi}^{2^j}_{i_{\rho(j)}+\rho(j)-j}. \qquad (18)
\end{aligned}$$

One can now define ρ' as in (12) and use (14) as in the case $i_0=1$ to establish the congruence of the sums in (17) and (18) modulo elements of length $>k$. This establishes the inductive step and proves the proposition. ∎

Define

$$\hat{\mathcal{X}}'_I(k) \;=\; \sum_{\mathrm{Id}^k\neq\tau\in\mathcal{S}(k)}\hat{X}_i(k;\tau).$$

We can then rewrite the conclusion of Proposition 5.5 as

$$\hat{X}_I(k;\mathrm{Id}^k) \equiv \hat{\mathcal{X}}'_I(k) + \xi_k^{2^{i_0}-1}\cdot\hat{\mathcal{R}}_I(k)$$

modulo elements of length $> k$. When $I = (i,\dots,i)$ is a constant sequence, we find by Observation 5.4 that

$$\hat{\xi}_i^{2^k-1} \equiv \hat{\mathcal{X}}'_I(k) + \xi_k^{2^i-1};$$

this is the formula referred to in Section 4.

Recall from Part 1 of Proposition 2.1 that $\xi_j \cap \Theta = 0$ if $\mathrm{len}(\Theta) < j$, and from Part (2) of the same proposition that $\xi_k \cap S[k;f] = S[k;f-1]$. Proposition 5.5 therefore implies:

Corollary 5.6. 1. *If* $\mathit{len}(\Theta) < k$, *then* $\hat{\mathcal{X}}_I(k)\cap\Theta = 0$ *for all* $I \in \mathcal{I}(k)$.

2. *If* $\mathit{len}(\Theta) = k$, *then* $\hat{\mathcal{X}}_I(k)\cap\Theta = \hat{\mathcal{R}}_I(k)\cap(\xi_k^{2^{i_0}-1}\cap\Theta)$.

3. *In particular,* $\hat{\mathcal{X}}_I(k)\cap S[k;f] = \hat{\mathcal{R}}_I(k)\cap S[k;f-(2^{i_0}-1)]$.

In what follows, we will use Part 3 of the corollary in the form

$$\hat{X}_I(k;\mathrm{Id}^k)\cap S[k;f] = (\hat{\mathcal{X}}'_I(k)\cap S[k;f]) + (\hat{\mathcal{R}}_I(k)\cap S[k;f-(2^{i_0}-1)]). \tag{19}$$

6. k–REDUCTIONS

In this section, we identify a constraint on the Milnor basis elements appearing in $\hat{S}[k;f]$. Here and in Section 8, we make frequent use of the inequality

$$2(2^{\Lambda-1}-1) < 2^\Lambda - 1. \tag{20}$$

Definition. Let $T \in \mathcal{S}$, and suppose $I(1),\dots,I(n) \in \mathcal{I}(k)$ are sequences such that all coefficients of $T - \sum_{r=1}^n Z_{I(r)}(k;\mathrm{Id}^k)$ are non–negative. Then T is *k–reducible via* $\overline{I} = \{I(1),\dots,I(n)\}$. Suppose also that $\tau_1,\dots,\tau_n \in \mathcal{S}(k)$ satisfy $P_{I(r)}(k;\tau_r) \neq *$ for $1 \le r \le n$. Set $U = T - \sum_{r=1}^n Z_{I(r)}(k;\mathrm{Id}^k)$ and $P = \sum_r P_{I(r)}(k;\tau_r)$. Then $[U;\,P] \in \mathcal{S}\times\mathcal{S}$ is the *k–reduction of T via $\overline{I}$ and $\overline{\tau} = \{\tau_1,\dots,\tau_n\}$.* We define the degree $|[U;\,P]|$ to be the sum $|U| + |P|$; it is easy to check that

Fact 6.1. With notation as above, the degree $|[U;\,P]| = |T| - \sum_r (2^k-1)(2^{i_0(r)}-1)$, where $i_0(r)$ is the 0-th coordinate of $I(r)$. In particular, any k–reduction of T has degree $\equiv |T|$ modulo (2^k-1).

Observation 6.2.
- If T is k–reducible via $\{I(1),\dots,I(n)\} \subset \mathcal{I}(k)$, then one may obtain a k–reduction of T by taking $\tau_r = \mathrm{Id}^k$ for all r. Indeed, if the $I(r)$ are constant sequences, then by Observation 5.4 this is the only corresponding k–reduction of T.
- If $t_i \ge 2^k-1$ for some i, then $[T - (2^k-1)B(i);0_{\mathcal{S}}]$ is a k–reduction of T via the constant sequence $(i,\dots,i) \in \mathcal{I}(k)$ and the permutation $\tau = \mathrm{Id}^k$. (Here $0_{\mathcal{S}}$ denotes the trivial element of $\mathcal{S}$ as in Section 2.1.) Consequently, every T has a k–reduction $[U;\,0_{\mathcal{S}}]$ with $u_i \le 2^k-2$ for all i.

Definition. A sequence $T \in \mathcal{S}$ is *k–irreducible* if it fails to be k–reducible via $\{I\}$ for any $I \in \mathcal{I}(k)$.

Lemma 6.3. 1. *If U is a k–irreducible sequence of length Λ, then*

$$|U| \leq |(0, \dots, 0, \underset{\Lambda}{2^k - 2})| = (2^k - 2)(2^\Lambda - 1).$$

2. *Suppose that $j \leq k-1$. With U as above, let $U[m] = (0, \dots, 0, u_m, \dots, u_\Lambda)$ and suppose that $U[m_j]$ is j–reducible but $U[m_j + 1]$ is j–irreducible. Then*

$$|U| \leq |(0, \dots, 0, \underset{m_j}{2^k - 2}, 0, \dots, 0, \underset{\Lambda}{2^j - 2})| = (2^k - 2)(2^{m_j} - 1) + (2^j - 2)(2^\Lambda - 1).$$

Proof. Part 1 is obvious for either $\Lambda = 1$ or $k = 1$. Suppose it is true for $\Lambda - 1$ and $k' \leq k$, and also for Λ and $k' \leq k - 1$. Let $U' = (u_1, \dots, u_{\Lambda-1})$; observe that U' is also k–irreducible. Suppose first that $u_\Lambda \leq 2^{k-1} - 1$. Then by (20) and the inductive hypothesis, $|U| < (2^k - 2)(2^\Lambda - 1)$ as claimed. Suppose now that $u_\Lambda = 2^k - 2^j + \varepsilon$ for some $1 \leq j \leq k-1$ and some $0 \leq \varepsilon \leq 2^{j-1} - 1$. This time U' must be j–irreducible, and again (20) and the inductive hypothesis give the required inequality for $|U|$. The proof of Part 2 is similar. ∎

Corollary 6.4. *Fix positive integers k and f, and let $T \in \mathcal{S}$ satisfy $M[T] \in \hat{S}[k; f]$. Then T is k–reducible.*

Proof. Let $\Lambda = \mathrm{len}(T)$. If T is k–irreducible, then by Lemma 6.3, $|T| \leq (2^k - 2)(2^\Lambda - 1)$. Since $|T| = (2^k - 1)f$, we find that $f < 2^\Lambda - 1$. Theorem 2.4 then implies that $\mathrm{len}(T) < \Lambda$. This contradiction proves the corollary. ∎

Given positive integers k and f, denote by $W \in \mathcal{S}$ the sequence for which $M[W]$ is maximal in right–lexicographical order among Milnor summands of minimal excess of $\hat{S}[k; f]$. Recall from Section 4 that our goal is to show that $w_i \geq 2^k - 1$ for some i; that is, that W is not only k–reducible but k–reducible via a constant sequence. This we accomplish in the next two sections by exploiting the conjugation formula of Section 5.

7. k–Reductions and stripping

In what follows, we assume for typographical convenience that $\overline{I}$ and $\overline{\tau}$ consist respectively of the single sequence $I \in \mathcal{I}(k)$ and the single permutation $\tau \in \mathcal{S}(k)$, but in fact the same reasoning applies for arbitrary $\overline{I}$ and $\overline{\tau}$.

Let $T \in \mathcal{S}$ and $\Theta \in \mathcal{A}^*$, and suppose that T is k–reducible via $I \in \mathcal{I}(k)$. Then by (3) we have

$$\begin{aligned} \langle \xi[T], \hat{\Theta} \rangle = \langle \hat{\xi}[T], \Theta \rangle &= \langle \hat{\xi}[T - Z_I(k; \mathrm{Id}^k)], \hat{\xi}[Z_I(k; \mathrm{Id}^k)] \cap \Theta \rangle \\ &= \langle \hat{\xi}[T - Z_I(k; \mathrm{Id}^k)], \hat{X}_I(k; \mathrm{Id}^k) \cap \Theta \rangle. \end{aligned} \tag{21}$$

Fix positive integers k and f, and set $\Theta = S[k; f]$. Henceforth we denote by $W \in \mathcal{S}$ the sequence for which $M[W]$ is maximal in the right–lexicographical order among all summands of minimal excess in the Milnor basis representation of $\hat{S}[k; f]$. If W is k–reducible via I, we have from (1), (19), and (21) that

$$\begin{aligned} 1 = \langle \hat{\xi}[W], S[k; f] \rangle &= \langle \hat{\xi}[W - Z_I(k; \mathrm{Id}^k)], \hat{\mathcal{X}}'_I(k) \cap S[k; f] \rangle \\ &\quad + \langle \hat{\xi}[W - Z_I(k; \mathrm{Id}^k)], \hat{\mathcal{R}}_I(k) \cap S[k; f - (2^{i_0} - 1)] \rangle, \end{aligned}$$

so

$$1 = \sum_{\substack{\mathrm{Id}^k \neq \tau \in \mathcal{S}(k) \\ Z_I(k;\tau) \neq *}} \langle \hat{\xi}[W - Z_I(k;\mathrm{Id}^k) + Z_I(k;\tau)],\, S[k;f]\rangle \tag{22}$$

$$+ \sum_{P_I(k;\tau)\neq *} \langle \hat{\xi}[\{W - Z_I(k;\mathrm{Id}^k)\} + P_I(k;\tau)],\, S[k;f-(2^{i_0}-1)]\rangle. \tag{23}$$

By Observation (5.4), any non-$*$ $Z_I(k;\tau)$ for $\tau \neq \mathrm{Id}^k$ is $\succ_R$ and no greater in excess than $Z_I(k;\mathrm{Id}^k)$, and so the term $W - Z_I(k;\mathrm{Id}^k) + Z_I(k;\tau)$ of (22) is $\succ_R$ and no greater in excess than W. By definition of W, then, $W - Z_I(k;\mathrm{Id}^k) + Z_I(k;\tau)$ does not appear in $\hat{S}[k;f]$. This implies that each summand in (22) vanishes, and consequently that the summation in (23) is non-0. But this summation is equal to

$$\sum \langle \hat{\xi}[U+P],\, S[k;f-(2^{i_0}-1)]\rangle,$$

where the summation is over all k–reductions $[U;\, P]$ via the sequence I and some $\tau \in \mathcal{S}(k)$. We therefore conclude by (1) and (22)-(23) that $M[U+P]$ appears in $\hat{S}[k;f-(2^{i_0}-1)]$ for some k–reduction $[U;\, P]$ of W via I.

Observe that if $w_i \geq 2^k - 1$ for some i, then I may be taken to be the constant sequence $(i,\dots,i)$. By Observation 5.4 the only k–reduction of W via I in this case corresponds to $\tau = \mathrm{Id}^k$. Since $Z_I(k;\mathrm{Id}^k) = (2^k-1)B(i)$ and $P_I(k;\mathrm{Id}^k) = 0_{\mathcal{S}}$, we find from (23) that $M[W-(2^k-1)B(i)] \in \hat{S}[k;f-(2^i-1)]$. This is the conclusion we reached when previewing the argument for the simplified version of Theorem 1.1 in Section 4 under the assumption that $w_i \geq 2^k-1$ for some i. In Section 8 below, we prove the stronger result that each entry of W is divisible by 2^k-1, and this result paves the way for the proof of Theorem 1.1 itself in Section 9.

In the meantime, we note that if W is reducible by any set of sequences $\overline{I} = \{I(1),\dots,I(n)\}$, then the above argument implies that $M[\overline{U}+\overline{P}]$ appears in $\hat{S}[k;f-\sum_r(2^{i_0(r)}-1)]$, where $M[\overline{U}+\overline{P}]$ is a k–reduction of W via $\overline{I}$ and some $\overline{\tau} \subset \mathcal{S}(k)$, and $i_0(r)$ denotes the 0-th coordinate of $I(r)$. We summarize this discussion in the following proposition, in which we deal separately with the constant and non–constant sequences in $\overline{I}$. First observe that if $T \in \mathcal{S}$ is written in the form $T = (2^k-1)G + H$ for some $G, H \in \mathcal{S}$, then by Observation 6.2, $(H;\, 0_{\mathcal{S}})$ is the unique k–reduction of T via constant sequences corresponding to the entries of G. Moreover, if $M[T] \in \hat{S}[k;f]$ (so that $|T| = (2^k-1)f$), then $|H| = (2^k-1)\eta$ for some integer η by Fact 6.1.

Proposition 7.1. *Let $M[W]$ be maximal in the right–lexicographical order among all Milnor basis elements of minimal excess appearing in $\hat{S}[k;f]$, and write $W = (2^k-1)G + H$ for some $G, H \in \mathcal{S}$ with $|H| = (2^k-1)\eta$. Then $M[H]$ appears in $\hat{S}[k;\eta]$.*

Suppose moreover that H is k–reducible via $\overline{I} = \{I(1),\dots,I(n)\}$, and denote by $i_0(r)$ the 0-th coordinate of $I(r)$. Then $M[U+P]$ appears in $\hat{S}[k;\eta-\sum_r(2^{i_0(r)}-1)]$ for some k–reduction $[U;\, P]$ of H via $\overline{I}$.

A similar argument yields the analogous conclusion concerning the sequence $\tilde{W}$ whose associated Milnor basis element is maximal in right–lexicographical order among all summands of $\hat{S}[k; f]$, regardless of excess.

8. Degrees of k–reductions

It follows from Observation 6.2 that if each coordinate of the sequence T is divisible by 2^k-1, then $[0_{\mathcal{S}}; 0_{\mathcal{S}}]$ is a k–reduction of T via constant sequences. We now turn our attention to sequences which are not divisible by 2^k-1.

Lemma 8.1. *Let $T \in \mathcal{S}$, and let L be the largest index i for which t_i is not divisible by 2^k-1. Then either $|T| < (2^k-1)(2^L-1)$ or T has a k–reduction $[U; P]$ of degree $< (2^k-1)(2^L-1)$ with $u_L > 0$.*

Proof. By Observation 6.2, we may assume that $t_i \leq 2^k-2$ for all i, and consequently that T is of length L. Thus by Lemma 6.3, the lemma is true if T is k-irreducible. In what follows, we assume that T is k–reducible.

The lemma is easily verified for $L = 1$ and all k, and for $k = 1$ and all L. Suppose then that it is true for $L-1$ and $k' \leq k$, and for L and $k' \leq k-1$. In what follows, we write T' for $(t_1, \ldots, t_{L-1})$.

Case I: $1 \leq t_L \leq 2^{k-1}-1$. Let $[U'; P']$ be a k–reduction of T' of degree $< (2^k-1)(2^{L-1}-1)$, which exists by the inductive hypothesis. Then it is easily verified using (20) that $[U' + t_L B(L); P']$ is a k–reduction of T satisfying the conditions stated in the lemma.

Case II: $2^{k-1} \leq t_L \leq 2^k-1$. Write $t_L = 2^k - 2^j + \varepsilon$ for some $1 \leq j \leq k-1$ and $0 \leq \varepsilon \leq 2^{j-1}-1$. Since T is k–reducible and $t_L < 2^k - 2^{j-1}$, T' must be j–reducible. We distinguish the cases $\varepsilon > 0$ and $\varepsilon = 0$.

Case IIa: $\varepsilon > 0$. With T' as usual, define $T'[m_j]$ as in Lemma 6.3. By the lemma, we have

$$|T| = |T'| + t_L(2^L-1) \leq (2^k-2)(2^{m_j}-1) + (2^j-2)(2^{L-1}-1) + (2^k-2^j+\varepsilon)(2^L-1).$$

Now $t_L \geq 2^k - 2^j$ and $T'[m_j]$ is j–reducible, so the sequence T itself is k–reducible via some sequence $I \in \mathcal{I}(k)$ with $i_0 = m_j$, $i_{j-1} < L$, and $i_l = L$ for $j \leq l \leq k-1$. Let $[U; P]$ be any k–reduction via I, and observe that $u_L = t_L - (2^k - 2^j) = \varepsilon \neq 0$. It is easily verified using Fact 6.1 that the degree $|[U; P]|$ satisfies the inequality given in the lemma.

Case IIb: $\varepsilon = 0$. If T' is $(j+1)$–reducible, proceed as in Case IIa, replacing j by $j+1$ when invoking Lemma 6.3 and choosing I; this time $u_L = t_L - (2^k - 2^{j+1}) = 2^j$. If T' is $(j+1)$–irreducible, the inductive hypothesis along with (20) implies that $[T; 0_{\mathcal{S}}]$ satisfies the requirements of the lemma. ∎

Proposition 8.2. *Fix positive integers k and f, and suppose that $W \in \mathcal{S}$ is the sequence whose associated Milnor basis element is maximal in right–lexicographical order among all summands of $\hat{S}[k; f]$ of minimal excess. Then $W = (2^k-1)Q$ for some $Q \in \mathcal{S}$ of degree f. Similarly, if $\tilde{W} \in \mathcal{S}$ is the sequence whose associated Milnor basis element is maximal in right–lexicographical order among all summands of $\hat{S}[k; f]$ regardless of excess, then $\tilde{W} = (2^k-1)\tilde{Q}$ for some $\tilde{Q} \in \mathcal{S}$ of degree f.*

Proof. We give the proof for W; the argument for $\overline{W}$ is identical. Suppose that the corollary is not true, and let L be the largest index i for which w_i is not divisible by 2^k-1. Take $[U;\ P]$ to be a k–reduction of W satisfying the conclusion of Lemma 8.1, namely that $u_L > 0$ and $|[U;P]| < (2^k-1)(2^L-1)$. Since $|[U;\ P]| \equiv |W| \equiv 0 \pmod{2^k-1}$ by Fact 6.1, we can write $|[U;\ P]| = (2^k-1)\phi$ for some $\phi < 2^L-1$. By Proposition 7.1 and Part 4 of Observation 5.4 we may assume, perhaps after adjusting the parameter $\overline{\tau}$ of the k–reduction, that $M[U+P]$ appears in $\hat{S}[k;\phi]$. But Theorem 2.4 with $f=\phi$ then implies that $u_L+p_L = 0$, contradicting the assumption on $[U;\ P]$. We conclude that indeed $w_i \equiv 0 \pmod{2^k-1}$ for all i. ∎

9. Proof of Theorem 1.1

Before stating Theorem 1.1 in its entirety and giving a proof, we recall a result of [Sil95b]. Given a positive integer f, let $\Lambda(f) = \max\{\lambda\ :\ 2^\lambda - 1 \le f\}$, and define sequences $R_1(f)$ inductively by

$$R_1(f) \ = \ \begin{cases} B(1) & f=1 \\ B(\Lambda(f)) + R_1(f-(2^{\Lambda(f)}-1)) & f>1. \end{cases}$$

Fix f and write $R_1(f) = (r_1, \ldots, r_{\Lambda(f)})$. Evidently $|R_1(f)| = f$ and $r_i \le 1$ for all $i \le \Lambda(f)$, except that the first non-0 r_i is ≤ 2. One may verify inductively that $R_1(f)$ is both minimal in excess over all Milnor basis elements of degree f and maximal in the right–lexicographical order over all elements in that degree regardless of excess [Gal79]. In particular $\mathrm{ex}(R_1(f)) = \mu(f)$, so that $\mu(f)$ may be computed recursively via $\mu(1)=1$, $\mu(f) = 1+\mu(f-(2^{\Lambda(f)}-1))$. For future reference, we observe that writing $f = |R_1(f)| = \sum_j (2^j-1)r_j$ gives rise to a decomposition of f as the sum of $\mu(f)$ numbers of the form 2^i-1, all distinct except possibly the two smallest, in which the number of times each 2^i-1 appears is given by r_i.

For $k \ge 2$, define $R_k(f) = (2^k-1)R_1(f)$. Then

Theorem 9.1. [Sil95b] For all positive integers f and k, $R_k(f)$ is a summand of $\hat{S}[k;f]$, and so $\mathrm{ex}(\hat{S}[k;f]) \le (2^k-1)\mu(f)$.

We are now ready to prove:

Theorem 1.1 *Let f be a positive integer. Then for all positive integers k, the Milnor basis element $M[R_k(f)]$ is both minimal in excess and maximal in right–lexicographical order over all basis elements appearing in $\hat{S}[k;f]$. In particular, $ex(\hat{S}[k;f]) = (2^k-1)\mu(f)$.*

Proof. The theorem for $k=1$ follows from Milnor's formula

$$\hat{S}q(f) \ = \ \sum_{|T|=f} M[T]$$

and the remarks following the definition of $R_1(f)$. Suppose then that $k>1$, and let $\tilde{W}$ (resp. W) $\in \mathcal{S}$ be the sequence whose associated Milnor basis element is maximal in right–lexicographical order among all summands of $\hat{S}[k;f]$ (resp. among all summands of $\hat{S}[k;f]$ of minimal excess). By Proposition 8.2, we have $W = (2^k-1)Q$ and $\tilde{W} = (2^k-1)\tilde{Q}$ for

some sequences Q, $\tilde{Q}$ of degree f. But the relations "$\succeq_R$" and "$\geq$ in excess" are preserved by termwise multiplication by $2^k - 1$, and so it follows from Theorem 9.1 and Theorem 1.1 for $k = 1$ that $W = \tilde{W} = R_k(f)$. ∎

10. Proof of Theorem 1.2

The $\mathcal{A}^*$-action on the polynomial rings $\mathbb{P}_s$ is defined by the Cartan formula and the requirement that $Sq(1)x_i = x_i^2$ for all i. It follows that for all $f, k \geq 1$ and all polynomials F of degree f, we have

$$S[k; f]F \;=\; F^{2^k}. \tag{24}$$

Recall from Section 1 that the excess of an element $\Theta \in \mathcal{A}^*$ satisfies $\mathrm{ex}(\Theta) = \min\{s \;:\; \Theta(x_1 x_2 \cdots x_s) \neq 0 \in \mathbb{P}_s\}$. Since linear maps commute with the action of $\mathcal{A}^*$, it follows that $\Theta(P) = 0$ for any polynomial in any $\mathbf{P}_{s'}$ of degree $< \mathrm{ex}(\Theta)$.

Theorem 1.2 *Let P be a polynomial of the form $E \cdot F^{2^k}$ for some polynomials E and F of degrees e and f respectively, and suppose that $e < (2^k - 1)\mu(f)$. Then P is hit.*

Proof. As mentioned in Section 1, the proof mirrors exactly Reg Wood's proof of the case $k = 1$[Woo89b]; the main ingredient is the observation that if U and V are polynomials and $\Theta \in \mathcal{A}^*$, then $U \cdot \Theta V \equiv \hat{\Theta} U \cdot V$ modulo hit elements. By (24), we have

$$\begin{aligned} EF^{2^k} &= E \cdot S[k; f]F \\ &\equiv \hat{S}[k; f]E \cdot F \quad \text{(modulo hit elements).} \end{aligned}$$

But $\mathrm{ex}(\hat{S}[k; f]) = (2^k - 1)\mu(f)$ by Theorem 1.1, so $\hat{S}[k; f]E = 0$ by the definition of excess and the hypothesis on e, f, and k. We conclude that EF^{2^k} is indeed hit. ∎

A given monomial M can generally be written in the form $E \cdot F^{2^k}$ in several ways, especially if k is allowed to vary, and to each such decomposition the theorem assigns a different inequality to be checked. In what remains of this section, we show that it suffices to consider the obvious decompositions in which F is square–free and the degree in E of each variable x_i is $\leq 2^k - 1$. We begin by stating two immediate consequences of the discussion of $R_1(f)$ and $\mu(f)$ in Section 9.

Consequence 10.1. 1. *$\mu(f + 1) = \mu(f) \pm 1$ accordingly as the first non-0 entry of $R_1(f)$ is 1 or 2.*
2. *If $A \leq 2^L - 1$ and $L \leq j_1 < \ldots < j_m$, then $\mu(A + \sum_{i=1}^m (2^{j_i} - 1)) = \mu(A) + m$.*

We shall need the following further properties of $\mu(f)$:

Lemma 10.2. 1. *For any positive integers g and h, we have $\mu(h) \leq \mu(g + h) + g$.*
2. *For any positive integer f, we have $\mu(f) \leq 2\mu(2f)$.*

Proof. Part 1 follows from the first consequence above by induction on g. To prove Part 2, write $f = \sum_{i=1}^{\mu(f)} (2^{j_i} - 1)$ with $j_1 \leq j_2 < j_3 < \ldots < j_{\mu(f)}$ as in the paragraph following the

definition of $R_1(f)$ in Section 9. Let $s = \left[\frac{\mu(f)+1}{2}\right]$. If $s \leq 1$, then $\mu(f) \leq 2 < 4 \leq 2\mu(2f)$. Suppose then that $s \geq 2$ and observe that

$$2\sum_{i=1}^{\mu(f)-s}(2^{j_i}-1) \geq 2(\mu(f)-s) \geq 2(s-1) \geq s.$$

Therefore we can write $2f$ in the form

$$\begin{aligned} 2f &= \sum_{i=\mu(f)-s+1}^{\mu(f)}[2(2^{j_i}-1)+1] + \left(2\sum_{i=1}^{\mu(f)-s}(2^{j_i}-1) - s\right) \\ &= \sum_{i=\mu(f)-s+1}^{\mu(f)}(2^{j_i+1}-1) + \left(2\sum_{i=1}^{\mu(f)-s}(2^{j_i}-1) - s\right). \end{aligned}$$

This decomposition may be seen to satisfy the conditions of the second consequence above, so we find that

$$\begin{aligned} \mu(2f) &= s + \mu\left(2\sum_{i=1}^{\mu(f)-s}(2^{j_i}-1) - s\right) \\ &\geq s \stackrel{\text{def}}{=} \left[\frac{\mu(f)+1}{2}\right], \end{aligned}$$

and hence $2\mu(2f) \geq \mu(f)$ as claimed. ∎

In what follows, we use uppercase letters such as E, F, and G to denote monomials and the corresponding lowercase letters to denote their respective degrees. If $e < (2^k-1)\mu(f)$, we shall say that the decomposition $M = E \cdot F^{2^k}$ satisfies the k-criterion for being hit.

Proposition 10.3. 1. *If the decomposition $(E \cdot G^{2^k}) \cdot H^{2^k}$ satisfies the k-criterion, then so does $E \cdot (FG)^{2^k}$.*
2. *If the decomposition $E \cdot (G^2)^{2^k}$ satisfies the k-criterion, then $E \cdot G^{2^{k+1}}$ satisfies the $(k+1)$-criterion.*

Proof. These statements follow easily from their counterparts in Lemma 10.2; we prove only the first. By assumption, we have $e + 2^k g < (2^k-1)\mu(h)$, so that $e < (2^k-1)\mu(h) - 2^k g$. But

$$\begin{aligned} (2^k-1)\mu(h) - 2^k g &< (2^k-1)(\mu(h)-g) \\ &\leq (2^k-1)\mu(g+h) \end{aligned}$$

by Part 1 of the lemma, so $e < (2^k-1)\mu(g+h)$ as claimed. ∎

If the monomial M is a perfect square, then we know from (24) and without recourse to Theorem 1.2 that it is hit. If M is not a square, then it can be uniquely written in the form $M = \prod_{j=0}^{n}(L_j)^{2^{k_j}}$, where each L_j is a non–trivial product of distinct x_i and $0 = k_0 < k_1 < \ldots < k_n$. For $1 \leq J \leq n$, define decompositions

$$D_J = D_J(M) = [\prod_{j<J}(L_j)^{2^{k_j}}] \cdot [\prod_{j \geq J}(L_j)^{2^{k_j - k_J}}]^{2^{k_J}}.$$

Proposition 10.3 implies that if any decomposition $E \cdot F^{2^k}$ of M satisfies the k-criterion for being hit, and if J is defined by $k_{J-1} < k \leq k_J$, then the decomposition D_J satisfies the k_J-criterion. This observation gives rise to the following

Hitness Test. *Let M be a monomial which is not a perfect square, and let D_j, $1 \leq J \leq n$ be its decompositions as above. Then to reap the full benefit of Theorem 1.2, it suffices to apply the k_J-criterion to D_J for $1 \leq J \leq n$.*

This test is best possible in the sense that for any pair $\kappa < \lambda$ of distinct positive integers, one can construct a pair of hit monomials M^1 and M^2 with the same exponent sequence $k_0 = 0,\ k_1 = \kappa,\ k_2 = \lambda$ such that for $i,\ j \in \{1,\ 2\}$, the decomposition $D_i(M^j)$ satisfies the k_i-criterion $\Longleftrightarrow i = j$. For example, given $\kappa = 1$ and $\lambda = 2$ one may take $M^1 = x_1^7 x_2^2$ and $M^2 = x_1^7 x_2^5 x_3$.

11. Acknowledgements

I thank David Carlisle, Grant Walker, and Reg Wood for many helpful conversations and for their hospitality during my visits to Manchester. I also thank the referee for his careful reading and detailed comments, and in particular for suggesting the "hitness test" at the end of Section 10.

References

[CW94] D.P. Carlisle and R.M.W. Wood. On an ideal conjecture in the Steenrod algebra. Preprint, 1994.

[Gal79] A. Gallant. Excess and conjugation in the Steenrod algebra. *Proc. Amer. Math. Soc.*, 76:161–166, 1979.

[Kra71] D. Kraines. On excess in the Milnor basis. *Bull. London Math. Soc.*, 3:363–365, 1971.

[Mil58] J. Milnor. The Steenrod algebra and its dual. *Ann. of Math.*, 67:150–171, 1958.

[Mit85] S. Mitchell. Finite complexes with $A(n)$–free cohomology. *Topology*, 24:227–248, 1985.

[Mon] K. G. Monks. Change of basis, monomial relations, and P_t^s bases for the Steenrod algebra. To appear in *Journal of Pure and Applied Algebra*.

[Mon94] K. G. Monks. Polynomial modules over the Steenrod algebra and conjugation in the Milnor basis. *Proc. Amer. Math. Soc.*, 122:625–634, 1994.

[Pet87] F.P. Peterson. Generators of $H^*(RP^\infty \wedge RP^\infty)$ as a module over the Steenrod algebra. *Abstracts of the Amer. Math. Soc.*, 833-55-89, 1987.

[Pet89] F.P. Peterson. A-generators for certain polynomial algebras. *Math. Proc. Cam. Phil. Soc.*, 105:311–312, 1989.

[Sil93] J. H. Silverman. Conjugation and excess in the Steenrod algebra. *Proc. Amer. Math. Soc.*, 119(2):657–661, 1993.

[Sil95a] J. H. Silverman. Hit polynomials and the canonical anti–automorphism. *Proc. Amer. Math. Soc.*, 123(2):627–637, 1995.

[Sil95b] J.H. Silverman. Stripping and conjugation in the Steenrod algebra. To appear in *Journal of Pure and Applied Algebra*, 1995.

[Sil96] J. H. Silverman. Multiplication and combinatorics in the Steenrod algebra. *Journal of Pure and Applied Algebra*, 111(3):303–323, 1996.

[Sin89] W. Singer. The transfer in homological algebra. *Math. Zeitschrift*, 1989.

[Sin91] W. Singer. On the action of Steenrod squares on polynomial algebras. *Proc. Amer. Math. Soc.*, 111(2):577–583, 1991.

[SS95] J.H. Silverman and W. Singer. On the action of Steenrod squares on polynomial algebras II. *Journal of Pure and Applied Algebra*, 98(1):95–103, 1995.

[Woo89a] R.M.W. Wood. Steenrod squares of polynomials. In *Advances in Homotopy Theory, London Math. Soc. Lecture Note Series 139*, pages 173–177. Cambridge University Press, 1989.

[Woo89b] R.M.W. Wood. Steenrod squares of polynomials and the Peterson conjecture. *Math. Proc. Cam. Phil. Soc.*, 105:307–309, 1989.

[WW96] G. Walker and R.M.W. Wood. The intersection of the admissible basis and the milnor basis of the mod–2 Steenrod algebra. preprint, 1996.

Indiana University–Purdue University at Columbus, 4601 Central Avenue, Columbus, IN 47203

E-mail address: judith@iu-math.math.indiana.edu