

*A presentation of the Steenrod algebra using
Kristensen's operator*

Sandling, Robert

2011

MIMS EPrint: **2011.100**

Manchester Institute for Mathematical Sciences
School of Mathematics

The University of Manchester

Reports available from: <http://eprints.maths.manchester.ac.uk/>

And by contacting: The MIMS Secretary
School of Mathematics
The University of Manchester
Manchester, M13 9PL, UK

ISSN 1749-9097

A presentation of the Steenrod algebra using Kristensen's operator

Robert Sandling
School of Mathematics
University of Manchester

The Steenrod algebra had its origin in operators on the cohomology rings of topological spaces. Here we develop it in an algebraic setting. Such an exposition can be made in a form mimicking the historical origin, for example, by using a polynomial ring in place of a cohomology ring. We do this here but make more explicit the role of the tensor algebra.

An alternative approach is to present the Steenrod algebra by generators and relations. We also take this path but attempt to smooth it by providing motivation for the relations, called the Adem or Adem-Wu relations. This is accomplished by the use of an operator on the tensor algebra due to Kristensen [Kr63,65; Gr75] in a recursive argument relying on nothing more than Pascal's triangle.

We demonstrate that the two definitions give the same object, namely, by identifying the ideal in the tensor algebra which is generated by the relators with that which is the kernel of the representation as operators. The core of the presentation is the theorem of Serre in this setting; it provides what is known as the admissible basis of the Steenrod algebra. Only the characteristic 2 case is addressed. For a reader unfamiliar with the Steenrod algebra, the material is elementary and reasonably self-contained.

1. An action of the tensor algebra

In this section we present the Steenrod algebra A as the quotient of the free or tensor algebra $T(V)$, where V is a graded \mathbf{F} -vector space of countable dimension, each of whose components in positive grading is of dimension 1 and whose components in non-positive grading are 0, by the kernel of an action which it has on polynomials (here $\mathbf{F} := \mathbf{F}_2$, the prime field in characteristic 2). The homogeneity convention is adopted: elements of $T(V)$ are taken to be homogeneous. We interpret the tensor algebra as the polynomial algebra $\mathbf{F}[\mathbf{X}]$, where $\mathbf{X} = \{X_1, X_2, \dots\}$ is a countable set of non-commuting indeterminates with the grading of X_m being m . We write monomials in the notation $X_{\mathbf{m}} = X_{m_1} X_{m_2} \dots$, where $\mathbf{m} = (m_1, m_2, \dots)$ is a countable vector of integers having only finitely many positive entries (by convention $X_0 = 1$ and $X_m = 0$, $m < 0$). We usually take such vectors \mathbf{m} in this context to be *normalised*, i.e., $m_1 \geq 0$ and, for $i > 1$, $m_i > 0$ implies $m_{i-1} > 0$; we write $\ell(\mathbf{m}) := \max\{i \mid m_i \neq 0\}$ for normalised \mathbf{m} , $\mathbf{m} \neq \mathbf{0}$, with $\ell(1) = \ell(X_0) := 0$, and often write $\mathbf{m} = (m_1, m_2, \dots, m_{\ell(\mathbf{m})})$ and $\mathbf{0} = (0, 0, \dots, 0)$. We use the notation $|\mathbf{m}| = |X_{\mathbf{m}}| = \sum m_i$ for the *grading*, or *degree*, of $X_{\mathbf{m}}$.

To define an action of $\mathbf{F}[\mathbf{X}]$ on the graded polynomial algebra $\mathbf{F}[\mathbf{x}]$, where $\mathbf{x} = \{x_1, x_2, \dots\}$ is a countable set of commuting indeterminates with the grading of x_n being 1 for all n , it suffices by the universal mapping property of the tensor algebra to define the image of each generator \mathbf{X}_m , $m \geq 1$, on a basis for $\mathbf{F}[\mathbf{x}]$ (we adopt the same convention as before: $x_0 = 1$ and $x_n = 0$, $n < 0$). We use the standard monomial basis for the commuting polynomial algebra with the notation

$$x^{\mathbf{v}} := \prod x_i^{v_i};$$

here \mathbf{v} is a countable vector as before but, in this context, we do not assume that it is normalised and take $\ell(x^{\mathbf{v}}) := \max\{i \mid v_i \neq 0\}$ for $\mathbf{v} \neq \mathbf{0}$ with $\ell(\mathbf{1}) = \ell(x^{\mathbf{0}}) := 0$; the grading of $x^{\mathbf{v}}$ is written as $|x^{\mathbf{v}}| = |\mathbf{v}| = \sum v_i$. For $f \in \mathbf{F}[\mathbf{x}]$, there is a finite set S for which $f = \sum_{\mathbf{v} \in S} x^{\mathbf{v}}$; write $\text{supp}(f) = \{i \mid v_i \neq 0 \text{ for some } \mathbf{v} \in S\}$. We make use of one further item of notation. For two such vectors \mathbf{v} and \mathbf{r} , write

$$\binom{\mathbf{v}}{\mathbf{r}} := \prod_i \binom{v_i}{r_i}.$$

We next define the desired (graded) representation of the tensor algebra on $\mathbf{F}[\mathbf{x}]$, which we denote $\rho : \mathbf{F}[\mathbf{X}] \rightarrow \text{End}(\mathbf{F}[\mathbf{x}])$, the algebra of graded linear transformations from $\mathbf{F}[\mathbf{x}]$ to itself. As noted it suffices to define ρ on the generators of $\mathbf{F}[\mathbf{X}]$. The definition of the Steenrod algebra as $A := \mathbf{F}[\mathbf{X}]/\text{Ker } \rho$ is close in spirit to its original definition. A bibliography of the original work on the Steenrod algebra may be found in Wood's survey [Wo98], and an analysis of its early history in [Ma99]. Here the focus is on fundamental results of Cartan [Ca50] and of Serre [Se53].

Definition. For $m \geq 0$ and $x^{\mathbf{v}}$ a monomial in $\mathbf{F}[\mathbf{x}]$, define $\rho(\mathbf{X}_m)$ via

$$\rho(\mathbf{X}_m)(x^{\mathbf{v}}) := \mathbf{X}_m(x^{\mathbf{v}}) = \sum_{\mathbf{r}} \binom{\mathbf{v}}{\mathbf{r}} x^{\mathbf{v}+\mathbf{r}}$$

with the sum taken over all vectors \mathbf{r} for which $|\mathbf{r}| = m$.

As illustrations, note that $\mathbf{X}_{|\mathbf{v}|}(x^{\mathbf{v}}) = x^{2\mathbf{v}}$ and so $\mathbf{X}_d(f) = f^2$ for a polynomial f of degree d (whence the name "squaring operation"), and that $\mathbf{X}_m(x^{\mathbf{v}}) = 0$ if $m > |\mathbf{v}|$.

The presence of the binomial coefficients makes it most feasible to apply the definition when the polynomial being acted upon is a product of distinct variables x_i . Indeed we conclude the paper with the criterion of Serre for a element of $\mathbf{F}[\mathbf{X}]$ to act trivially which is formulated in terms of such elements. Our notation for them is as follows. Let I be a set of positive integers. Then $\mathbf{1}_I$ denotes the vector whose i th entry is $\delta_{i,I}$, i.e., 1 if $i \in I$ and 0 otherwise. That is, if $I = \{i_1, \dots, i_\ell\}$, then

$$x^{\mathbf{1}_I} = x_{i_1} \cdots x_{i_\ell}.$$

We abbreviate $\mathbf{1}_{[1,h]}$ to $\mathbf{1}_h$, where $[1, h]$ is the interval $\{1, 2, \dots, h\}$. Note that $\mathbf{1}_0 = \mathbf{0}$.

The formula for the action of a general monomial is given as follows. For $\mathbf{m} = (m_1, m_2, \dots, m_\ell)$, $\ell \geq 1$, $\mathbf{X}_{\mathbf{m}}(x^{\mathbf{v}})$ is the sum of all terms

$$\text{BC}(\mathbf{v}, \mathbf{r}_\ell, \dots, \mathbf{r}_1) x^{\mathbf{v} + \mathbf{r}_\ell + \dots + \mathbf{r}_1},$$

where the sum is taken over all sequences of vectors $\mathbf{r}_\ell, \dots, \mathbf{r}_1$ for which $|\mathbf{r}_i| = m_i$ and where the coefficient is defined as

$$\binom{\mathbf{v}}{\mathbf{r}_\ell} \binom{\mathbf{v} + \mathbf{r}_\ell}{\mathbf{r}_{\ell-1}} \dots \binom{\mathbf{v} + \mathbf{r}_\ell + \dots + \mathbf{r}_{i+1}}{\mathbf{r}_i} \dots \binom{\mathbf{v} + \mathbf{r}_\ell + \dots + \mathbf{r}_2}{\mathbf{r}_1}.$$

Because of the binomial coefficients, we need consider only sequences for which

$$0 \leq r_{ij} \leq v_j + r_{\ell j} + \dots + r_{(i+1)j}$$

for all i, j , $\ell \geq i \geq 1$, $1 \leq j \leq \ell(\mathbf{v})$. For the same reason, $\text{supp}(F(f)) \subseteq \text{supp}(f)$ for $F \in \mathbf{F}[\mathbf{X}]$ and $f \in \mathbf{F}[\mathbf{x}]$.

An important method for determining the action, especially in induction settings, is a recursive one attributed to Cartan.

Theorem. [Cartan's formula.] For $f, g \in \mathbf{F}[\mathbf{x}]$ and $\mathbf{m} \geq \mathbf{0}$,

$$\mathbf{X}_{\mathbf{m}}(fg) = \sum_{\mathbf{k} + \mathbf{\ell} = \mathbf{m}} \mathbf{X}_{\mathbf{k}}(f) \mathbf{X}_{\mathbf{\ell}}(g).$$

Proof. We may assume that $\mathbf{m} \neq \mathbf{0}$. By induction on the length of \mathbf{m} , we may take \mathbf{m} to be of length 1. By linearity we may assume that f, g are monomials.

Suppose then that $m > 0$, $f = x^{\mathbf{u}}$ and $g = x^{\mathbf{v}}$. Then

$$\mathbf{X}_m(fg) = \sum_{|\mathbf{r}|=m} \binom{\mathbf{u} + \mathbf{v}}{\mathbf{r}} x^{\mathbf{u} + \mathbf{v} + \mathbf{r}}.$$

On the other hand,

$$\begin{aligned} \sum_{k+\ell=m} \mathbf{X}_k(x^{\mathbf{u}}) \mathbf{X}_\ell(x^{\mathbf{v}}) &= \sum_{k+\ell=m} \sum_{|\mathbf{p}|=k} \binom{\mathbf{u}}{\mathbf{p}} x^{\mathbf{u} + \mathbf{p}} \sum_{|\mathbf{q}|=\ell} \binom{\mathbf{v}}{\mathbf{q}} x^{\mathbf{v} + \mathbf{q}} = \\ &= \sum_{k+\ell=m} \sum \binom{\mathbf{u}}{\mathbf{p}} \binom{\mathbf{v}}{\mathbf{q}} x^{\mathbf{u} + \mathbf{v} + \mathbf{p} + \mathbf{q}} = \sum_{|\mathbf{r}|=m} \left(\sum_{\mathbf{p} + \mathbf{q} = \mathbf{r}} \binom{\mathbf{u}}{\mathbf{p}} \binom{\mathbf{v}}{\mathbf{q}} \right) x^{\mathbf{u} + \mathbf{v} + \mathbf{r}}. \end{aligned}$$

The conclusion now follows as

$$\binom{\mathbf{u} + \mathbf{v}}{\mathbf{r}} = \sum_{\mathbf{p} + \mathbf{q} = \mathbf{r}} \binom{\mathbf{u}}{\mathbf{p}} \binom{\mathbf{v}}{\mathbf{q}}$$

by the Vandermonde identity [AS72, 24.1.1] applied to each vector coordinate. \square

Cartan's formula in the length 1 case and the additional hypotheses that, for all i , $X_1(x_i) = x_i^2$ and $X_m(x_i) = 0$ if $m > 1$ characterise the representation ρ in the sense that any linear representation of $\mathbf{F}[X]$ on $\mathbf{F}[\mathbf{x}]$ which satisfies these hypotheses must be ρ . See [Wo98, Lemma 1.3].

Definition. The vector \mathbf{a} is *admissible* if $a_i \geq 2a_{i+1}$ for all i , $i \geq 1$, and a monomial $X_{\mathbf{a}}$ is *admissible* in $\mathbf{F}[X]$ if \mathbf{a} is. Note that $1 = X_{\mathbf{0}}$ is admissible. The quantity

$$e = e_{\mathbf{a}} := 2a_1 - \deg \mathbf{a} = a_1 - (a_2 + \cdots + a_{\ell(\mathbf{a})}) = a_1 - (a_2 + \cdots)$$

is called the *excess* of \mathbf{a} (i.e., $e = (a_1 - 2a_2) + \cdots + (a_{\ell-1} - 2a_{\ell}) + (a_{\ell} - 2a_{\ell+1}) + \cdots$). The degree of an admissible is greater than or equal to its excess.

Serre showed that the images of the admissibles form a basis of the Steenrod algebra. In our proof of the independence of the admissible basis we make use of the following order relation on vectors. It is used in the subsequent proposition to produce, for a given non-empty set of admissibles, a certain polynomial which their sum does not send to 0.

Definition. The *left lexicographic order* on the set of countable vectors of integers is as follows. Let $\mathbf{m} = (m_1, m_2, \cdots)$ and $\mathbf{n} = (n_1, n_2, \cdots)$. Then $\mathbf{m} < \mathbf{n}$ if there is j such that, for $1 \leq i < j$, $m_i = n_i$ and $m_j < n_j$. The same phrase is used for the order induced on the monomials $X_{\mathbf{m}}$ of $\mathbf{F}[X]$.

Note that, in the formula for the action of a general monomial, we may assume that \mathbf{r}_i is less than or equal to $\mathbf{v} + \mathbf{r}_{\ell} + \cdots + \mathbf{r}_{i+1}$ in each coordinate so that, for $\ell \geq i \geq 1$,

$$\mathbf{r}_i \leq \mathbf{v} + \mathbf{r}_{\ell} + \cdots + \mathbf{r}_{i+1}$$

in the left lexicographic order.

Lemma. *Let \mathbf{a} be an admissible. If $h \geq e_{\mathbf{a}}$, then $X_{\mathbf{a}}(x^{1^h}) \neq 0$.*

Proof. Let $\mathbf{a} = (a_1, a_2, \cdots, a_{\ell})$, $a_{\ell} > 0$. Then $X_{\mathbf{a}}(x^{1^h})$ is the sum of all terms

$$\text{BC}(\mathbf{1}_h, \mathbf{r}_{\ell}, \cdots, \mathbf{r}_1) x^{1^h + \mathbf{r}_{\ell} + \cdots + \mathbf{r}_1}$$

taken over all sequences of suitable vectors $\mathbf{r}_{\ell}, \cdots, \mathbf{r}_1$ as described above. Such a sequence may be presented as a matrix with rows $\mathbf{1}_h, \mathbf{r}_{\ell}, \cdots, \mathbf{r}_1$. The column sums give the vector $\mathbf{r}_0 := \mathbf{1}_h + \mathbf{r}_{\ell} + \cdots + \mathbf{r}_1$, and the row sums the vector $(h, a_{\ell}, a_{\ell-1}, \cdots, a_1)$.

As an illustration we give the matrix $M_{\mathbf{a}}$ which provides the maximal, according to the left lexicographic order, among all vectors obtainable from such sequences, namely,

$$\begin{array}{cccccccccccccccc} 1 & \cdots & 1 & 1 & \cdots & 1 & \cdots & 1 & \cdots & 1 & 1 & \cdots & 1 & 0 & \cdots \\ 1 & \cdots & 1 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots \\ 2 & \cdots & 2 & 1 & \cdots & 1 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots \\ & \cdots & & & \cdots & & \cdots & & \cdots & & & \cdots & & & \cdots \\ & \cdots & & & \cdots & & \cdots & & \cdots & & & \cdots & & & \cdots \\ 2^{\ell-1} & \cdots & 2^{\ell-1} & 2^{\ell-2} & \cdots & 2^{\ell-2} & \cdots & 1 & \cdots & 1 & 0 & \cdots & 0 & 0 & \cdots \end{array}$$

Here the initial row, is $\mathbf{1}_h$ while, if $\ell \geq i \geq 1$, the $(\ell - i + 2)$ th row, \mathbf{m}_i , is defined as

$$\mathbf{1}_{a_i - (a_{i+1} + \dots)} + \mathbf{m}_\ell + \mathbf{m}_{\ell-1} + \dots + \mathbf{m}_{i+1}.$$

For this sequence $\text{BC}(\mathbf{1}_h, \mathbf{m}_\ell, \dots, \mathbf{m}_1) = 1$ since, if $\ell \geq i \geq 1$,

$$\begin{pmatrix} \mathbf{1}_h + \mathbf{m}_\ell + \dots + \mathbf{m}_{i+1} \\ \mathbf{m}_i \end{pmatrix} = \begin{pmatrix} \mathbf{m}_i \\ \mathbf{m}_i \end{pmatrix} \begin{pmatrix} \mathbf{1}_{[1+(a_i - (a_{i+1} + \dots)), h]} \\ \mathbf{0} \end{pmatrix}.$$

For $M_{\mathbf{a}}$ the column sums give the vector

$$\mathbf{m}_0 = (2^\ell, \dots, 2^\ell, 2^{\ell-1}, \dots, 2^{\ell-1}, \dots, 1, \dots, 1, 0, \dots)$$

with a_ℓ entries 2^ℓ , $a_{\ell-1} - 2a_\ell$ entries $2^{\ell-1}$, \dots , $a_k - 2a_{k+1}$ entries 2^k , \dots until $a_1 - 2a_2$ entries 2 and lastly $h - e_{\mathbf{a}}$ entries 1. For a different such sequence $\mathbf{r}_\ell, \dots, \mathbf{r}_1$ an induction on i , $\ell \geq i \geq 1$, shows that $\mathbf{m}_i \geq \mathbf{r}_i$. At least one of the inequalities must be strict and so $\mathbf{m}_0 > \mathbf{r}_0$. Thus \mathbf{m}_0 is maximal among vectors arising from such sequences, and arises from a unique such sequence. It follows that $X_{\mathbf{a}}(x^{\mathbf{1}_h}) \neq 0$. \square

We use the notation $\mathbf{v}_{\mathbf{a}}$ for \mathbf{m}_0 . Note that \mathbf{a} may be recovered from $\mathbf{v}_{\mathbf{a}}$.

Proposition. *A non-zero sum of admissibles acts non-trivially.*

Proof. Let A be a non-empty set of admissible vectors and $F = \sum_{\mathbf{a} \in A} X_{\mathbf{a}}$. Let d be the common degree of the vectors of A . Thus, $d \geq e_{\mathbf{a}}$ for all $\mathbf{a} \in A$ so that the previous lemma applies. We show that, if $h \geq d$, $F(x^{\mathbf{1}_h}) \neq 0$.

Using the notation of the previous proof, let \mathbf{a}_{\max} be that unique element of A for which $\mathbf{v}_{\mathbf{a}_{\max}}$ is maximal among the $\mathbf{v}_{\mathbf{a}}$, $\mathbf{a} \in A$, in the left lexicographic ordering. Taking $\mathbf{a} \in A$, let $\ell = \ell(\mathbf{a})$ be its length and, for a suitable sequence $\mathbf{r}_\ell, \dots, \mathbf{r}_1$ as described above, let

$$\mathbf{v} = \mathbf{1}_h + \mathbf{r}_\ell + \dots + \mathbf{r}_1$$

so that $x^{\mathbf{v}}$ arises as a summand of $X_{\mathbf{a}}(x^{\mathbf{1}_h})$ as above. Then $\mathbf{v}_{\mathbf{a}_{\max}}$ is strictly greater than \mathbf{v} if $\mathbf{a} \neq \mathbf{a}_{\max}$ since

$$\mathbf{v}_{\mathbf{a}_{\max}} > \mathbf{v}_{\mathbf{a}} \geq \mathbf{v},$$

as well as if $\mathbf{a} = \mathbf{a}_{\max}$ unless the sequence $\mathbf{r}_1, \dots, \mathbf{r}_\ell$ for \mathbf{v} is that giving rise to $\mathbf{v}_{\mathbf{a}_{\max}}$ itself. Thus $F(x^{\mathbf{1}_h}) \neq 0$. \square

Theorem. [Serre.] *The admissible monomials are independent modulo $\text{Ker } \rho$.*

2. Just inadmissible monomials

Among the simplest examples of admissible elements are the monomials $X_{2b}X_b$, $b \geq 1$. As these are only just admissible, then good candidates for *just inadmissible* monomials are those of the form $X_{2b-1}X_b$, $b \geq 1$. In this section we

show that these latter elements act trivially. In subsequent sections we see that, from them, a set of ideal generators for $\text{Ker}\rho$ can be obtained. We prove that $\mathbf{X}_{2b-1}\mathbf{X}_b \in \text{Ker}\rho$ by showing that $\mathbf{X}_{2b-1}\mathbf{X}_b(x^{1^h}) = 0$ for all $h, h \geq 0$.

To see that this case suffices we use a result which is of independent interest, the fact that the action of $\mathbf{F}[\mathbf{X}]$ commutes with functions of the variables x_1, x_2, \dots .

Proposition. *The action of $\mathbf{F}[\mathbf{X}]$ on $\mathbf{F}[\mathbf{x}]$ commutes with endomorphisms induced by functions from \mathbf{x} to itself.*

Proof. A function from \mathbf{x} to itself sends each indeterminate x_i to another $x_{\tau(i)}$, where τ is a function from the set \mathbf{P} of positive integers to itself. We use the same symbol τ for the induced endomorphism of $\mathbf{F}[\mathbf{x}]$, i.e., for $f \in \mathbf{F}[\mathbf{x}]$,

$$\tau(f)(x_1, x_2, \dots) = f(x_{\tau(1)}, x_{\tau(2)}, \dots).$$

This defines a left action of the monoid of functions from the set \mathbf{P} to itself on the ring $\mathbf{F}[\mathbf{x}]$. For a monomial f , e.g., $f(x_1, x_2, \dots) = x_1^{v_1} x_2^{v_2} \dots = x^{\mathbf{v}}$, we write $\tau(x^{\mathbf{v}})$ for $\tau(f)(x_1, x_2, \dots)$; we use similar notation in more complicated settings as well. If we write $\tau(\mathbf{v})$ for the vector $(\sum\{v_i \mid \tau(i) = 1\}, \sum\{v_i \mid \tau(i) = 2\}, \dots)$, then $\tau(x^{\mathbf{v}}) = x^{\tau(\mathbf{v})}$.

We must show that, for each $F \in \mathbf{F}[\mathbf{X}]$, $\tau(F(x^{\mathbf{v}})) = F(x^{\tau(\mathbf{v})})$. It suffices to prove the special case that, for each $m, m \geq 0$, and for each vector \mathbf{v} ,

$$\tau(\mathbf{X}_m(x^{\mathbf{v}})) = \mathbf{X}_m(x^{\tau(\mathbf{v})}) = \mathbf{X}_m(\tau(x^{\mathbf{v}})).$$

We initially deal with it for a permutation τ of a finite number of integers. As these are generated by transpositions of consecutive integers, we may assume that $\tau = (i, i+1)$ for some i , a case for which a proof is straightforward.

Working with a given \mathbf{v} so that $\text{supp}(x^{\mathbf{v}})$ is a finite set, we may assume that τ moves only a finite number of integers. If τ is injective, then it is a permutation, a case already settled. If τ is not injective, then we may assume by use of permutations that there are non-negative integers k_1, k_2, \dots such that, for all i ,

$$\tau^{-1}(i) = \{1 + (k_1 + \dots + k_{i-1}), \dots, k_i + (k_1 + \dots + k_{i-1})\}.$$

With $\ell = \ell(\tau(\mathbf{v}))$ and with

$$\mathbf{v}(i) := (v_{1+(k_1+\dots+k_{i-1})}, \dots, v_{k_i+(k_1+\dots+k_{i-1})})$$

so that the i th coordinate of $\tau(\mathbf{v})$ is $|\mathbf{v}(i)|$, it follows that

$$\begin{aligned} \tau(\mathbf{X}_m(x^{\mathbf{v}})) &= \sum_{|\mathbf{r}|=m} \binom{\mathbf{v}}{\mathbf{r}} x^{\tau(\mathbf{v}+\mathbf{r})} = \\ &= \sum_{|\mathbf{r}|=m} \binom{\mathbf{v}}{\mathbf{r}} x_1^{(v_1+r_1)+\dots+(v_{k_1+r_{k_1}})} x_2^{(v_{1+k_1+r_{1+k_1}})+\dots+(v_{k_2+k_1+r_{k_2+k_1}})} \dots = \end{aligned}$$

$$\begin{aligned}
&= \sum_{|\mathbf{r}|=m} \binom{\mathbf{v}}{\mathbf{r}} x^{\tau(\mathbf{v})+\tau(\mathbf{r})} = \sum_{|\mathbf{p}|=m} \left(\sum_{\mathbf{r}, \tau(\mathbf{r})=\mathbf{p}} \binom{\mathbf{v}}{\mathbf{r}} \right) x^{\tau(\mathbf{v})+\mathbf{p}} = \\
&= \sum_{|\mathbf{p}|=m} \left(\sum_{\tau(\mathbf{r})=\mathbf{p}} \prod_{1 \leq i \leq \ell} \binom{\mathbf{v}(i)}{\mathbf{r}(i)} \right) x^{\tau(\mathbf{v})+\mathbf{p}} = \\
&= \sum_{|\mathbf{p}|=m} \left(\prod_{1 \leq i \leq \ell} \left(\sum_{\mathbf{m}, |\mathbf{m}|=p_i} \binom{\mathbf{v}(i)}{\mathbf{m}} \right) \right) x^{\tau(\mathbf{v})+\mathbf{p}}.
\end{aligned}$$

But by the generalised Vandermonde identity the last expression is

$$\sum_{|\mathbf{p}|=m} \prod_i \binom{|\mathbf{v}(i)|}{p_i} x^{\tau(\mathbf{v})+\mathbf{p}},$$

which is thus

$$\sum_{|\mathbf{p}|=m} \binom{\tau(\mathbf{v})}{\mathbf{p}} x^{\tau(\mathbf{v})+\mathbf{p}} = \mathbf{X}_m(x^{\tau(\mathbf{v})})$$

as required. \square

The action of $\mathbf{F}[\mathbf{X}]$ on $\mathbf{F}[\mathbf{x}]$ does not commute in general with functions from \mathbf{x} to $\mathbf{x} \cup \{x_0 = 1\}$.

Corollary. *An element F of $\mathbf{F}[\mathbf{X}]$ belongs to $\text{Ker} \rho$ if and only if, for all non-negative h , $F(x_1 x_2 \cdots x_h) = 0$.*

Proof. We show that $F(f) = 0$ when f is a monomial, the case of principal interest. Write $f = x_{i_1} x_{i_2} \cdots x_{i_h}$. Define $\tau(j) = i_j$ if $1 \leq j \leq h$ and $\tau(j) = j$ otherwise. Thus $\tau(x^{\mathbf{1}^h}) = f$ and

$$F(f) = F(\tau(x^{\mathbf{1}^h})) = \tau(F(x^{\mathbf{1}^h})) = 0.$$

\square

To prove the main result of this section we use the following congruence of binomial coefficients.

Lemma. *For an integer n ,*

$$\binom{3n-1}{n} \equiv 0 \pmod{2}.$$

Proof. For $n \geq 1$, let n_i be the i th binary digit of n , i.e., $n = \sum_i n_i 2^i$, $n_i = 0, 1$, and let k be minimal such that $n_k = 1$. As

$$\binom{3n-1}{n} \equiv \prod_i \binom{\ell_i}{n_i} \pmod{2},$$

where ℓ_i be the i th binary digit of $3n - 1$ (see [AS72, 24.1.1]). If $k = 0$, then n is odd and $3n - 1$ is even so that $\binom{3n-1}{n}$ is even. If $k > 0$, then from $3n - 1 = 2n + (n - 1)$ we see that the k th binary digit of $3n - 1$ is 0 (viz., the k th binary digits both of $2n$ and of $n - 1$ are 0 while the $(k - 1)$ th binary digit of $2n$ is 0) and the k th binary digit of n is 1, from which the result follows. \square

Theorem. *Each monomial $X_{2b-1}X_b$ belongs to $\text{Ker}\rho$.*

Proof. It suffices to show that $X_{2b-1}X_b(x^{\mathbf{1}_h}) = 0$ for all h . We may assume that $h \geq b \geq 1$. Since $\binom{\mathbf{1}_h}{\mathbf{r}} = 1$ if and only if $\mathbf{r} = \mathbf{1}_R$ for a subset R of $\{1, \dots, h\}$, we see that

$$X_b(x^{\mathbf{1}_h}) = \sum x^{\mathbf{1}_h + \mathbf{1}_R},$$

where the sum is taken over all such R with $|R| = b$.

Next

$$X_{2b-1}(x^{\mathbf{1}_h + \mathbf{1}_R}) = \sum_{|\mathbf{r}|=2b-1} \binom{\mathbf{1}_h + \mathbf{1}_R}{\mathbf{r}} x^{\mathbf{1}_h + \mathbf{1}_R + \mathbf{r}}.$$

For such an \mathbf{r} , if $i > h$, we may assume that $r_i = 0$. Suppose that $i \leq h$; if $i \notin R$, then the i th factor of the coefficient is $\binom{1}{r_i}$, which is non-zero if and only if $r_i = 0, 1$; if $i \in R$, then the i th factor is $\binom{2}{r_i}$, which is non-zero modulo 2 if and only if $r_i = 0, 2$. We may thus assume that $\mathbf{r} = \mathbf{1}_S + 2\mathbf{1}_T$, where $S, T \subset \{1, \dots, h\}$, $S \cap R = \emptyset$, $T \subseteq R$ and $|S| + 2|T| = 2b - 1$. For such an \mathbf{r} ,

$$x^{\mathbf{1}_h + \mathbf{1}_R + \mathbf{1}_S + 2\mathbf{1}_T} = x^{\mathbf{1}_Q} x^{2\mathbf{1}_P} x^{3\mathbf{1}_T},$$

where $P := S \cup (R - T)$ and $Q := \{1, \dots, h\} - (P \cup T)$.

Lastly we count the number of times in which the term $x^{\mathbf{1}_Q} x^{2\mathbf{1}_P} x^{3\mathbf{1}_T}$ appears on the right-hand side above. We must show that it is even. It is the number of subsets R, S, T as above which give rise to the triple Q, P, T . For given T of size t and P of size p , there are $\binom{p}{b-t}$ suitable choices for R . Then S is determined as $P - (R - T)$ and has size $p - (b - t)$. But $2b - 1 = p - (b - t) + 2t$, whence $p = 3(b - t) - 1$. The conclusion follows as the binomial coefficient

$$\binom{p}{b-t} = \binom{3(b-t)-1}{b-t}$$

is even by the lemma. \square

3. Kristensen's operator

Kristensen [Kr63,65] observed that the Adem-Wu relations for the Steenrod algebra are obtainable from those of the form $\text{Sq}^{2b-1}\text{Sq}^b = 0$ by applying a linear operator which he referred to as a derivation. Subsequently his operator was interpreted in the context of bialgebras as an instance of an action of a commutative algebra on its dual algebra. This action has been called stripping [Wo98, Si97]. We give an exposition of Kristensen's observation (see [Gr75, p.

318]) in which the binomial coefficients modulo 2 which appear in the Adem-Wu relations emerge recursively via Pascal's triangle.

Our goal is to express each inadmissible $\text{Sq}^a \text{Sq}^b$, $0 < a < 2b$, as a linear combination of admissibles, which can be taken in the form $\text{Sq}^{a+b-k} \text{Sq}^k$, $0 \leq k$. Our recursive definition of the Adem-Wu relations has for its foundation the following relations among Steenrod squares:

$$\text{Sq}^{2b-1} \text{Sq}^b = 0, \quad \text{Sq}^{2b-2} \text{Sq}^b = \text{Sq}^{2b-1} \text{Sq}^{b-1}, \quad \text{Sq}^{2b-3} \text{Sq}^b = \text{Sq}^{2b-1} \text{Sq}^{b-2}$$

for all b , $b \geq 1$. They are reflected in the following lemma and built upon in the subsequent proposition.

Lemma. *For integers a, b, k , $0 < a < 2b$, $0 \leq k$, the numbers $\binom{b-1-k}{a-2k}$ are determined modulo 2 by*

$$\binom{b-1-k}{a-2k} \equiv \binom{(b-1)-1-k}{(a+1)-2k} + \binom{b-1-k}{(a+1)-2k} + \binom{b-1-(k+1)}{(a+1)-2(k+1)}$$

unless $k = b-1$ and either $a = 2b-2$ or $a = 2b-3$ in which case

$$\binom{b-1-(b-1)}{(2b-2)-2(b-1)} = 1 \quad \text{and} \quad \binom{b-1-(b-1)}{(2b-3)-2(b-1)} = 0.$$

Proof. We first show that these are properties of the binomial coefficients. We may rewrite the right-hand side as

$$\binom{b-1-k}{a-2k+1} + \binom{(b-1-k)-1}{a-2k+1} + \binom{(b-1-k)-1}{a-2k-1}.$$

Modulo 2, this is

$$\begin{aligned} & \binom{b-1-k}{a-2k+1} + \binom{(b-1-k)-1}{a-2k+1} + \binom{(b-1-k)-1}{a-2k} + \\ & + \binom{(b-1-k)-1}{a-2k} + \binom{(b-1-k)-1}{a-2k-1}. \end{aligned}$$

That, modulo 2, this is

$$\binom{b-1-k}{a-2k+1} + \binom{b-1-k}{a-2k+1} + \binom{b-1-k}{a-2k} \equiv \binom{b-1-k}{a-2k}$$

is implied by Pascal's rule which holds for all integer variables, positive or not, with the single exception $\binom{-1}{-1} + \binom{-1}{0} \neq \binom{0}{0}$. This exceptional case is encountered here when $k = b-1$ and either $a = 2b-2$ or $a = 2b-3$.

It remains to show that an \mathbf{F} -valued function of a, b, k which satisfies these properties must be $\binom{b-1-k}{a-2k}$ modulo 2 on the given range. The equivalence may

be used to establish this by induction, a recursion in the quantity $2b - a$ (the induction is grounded by giving the explicit values when $2b - a \leq 3$). \square

Definition. *Kristensen's operator*, denoted here by K , is the linear operator of degree -1 on $\mathbf{F}[\mathbf{X}]$ defined on the basis of monomials by

$$K(\mathbf{X}_{\mathbf{m}}) = \sum_j \mathbf{X}_{\mathbf{m}^{(j)}},$$

where $\mathbf{m}^{(j)}$ is the same as \mathbf{m} except that its j th entry is $m_j - 1$ [Kr63,65, Gr75, Wo98].

Using the universal mapping property of the tensor algebra, we may introduce a coalgebra structure on $\mathbf{F}[\mathbf{X}]$ via the cocommutative coproduct defined by sending \mathbf{X}_n , $n \geq 1$, to $\sum_{0 \leq i \leq n} \mathbf{X}_i \otimes \mathbf{X}_{n-i}$. Then the commutative algebra $\mathbf{F}[\mathbf{X}]^*$ acts on $\mathbf{F}[\mathbf{X}]$ by stripping and the action of \mathbf{X}_1^* is that of K . The stripping action behaves well with respect to products. In our special case this behaviour takes the following form.

Lemma. For $F, G \in \mathbf{F}[\mathbf{X}]$, $K(FG) = K(F)G + FK(G)$.

Proof. This is immediate from the definition if F, G are monomials, and the general case reduces to this one by linearity. \square

For integers a, b , we define polynomials $R_{a,b}$ recursively as follows. Unless $0 < a < 2b$, let $R_{a,b} = 0$. If $0 < a < 2b$, then, if $2b - a = 1$, let $R_{2b-1,b} = \mathbf{X}_{2b-1}\mathbf{X}_b$, and, if $2b - a > 1$, let $R_{a,b} = R_{a+1,b-1} + K(R_{a+1,b})$.

For example,

$$R_{2b-2,b} = R_{2b-1,b-1} + K(R_{2b-1,b}) = 0 + K(\mathbf{X}_{2b-1}\mathbf{X}_b) = \mathbf{X}_{2b-2}\mathbf{X}_b + \mathbf{X}_{2b-1}\mathbf{X}_{b-1}$$

and

$$R_{2b-3,b} = R_{2b-2,b-1} + K(R_{2b-2,b}) = \mathbf{X}_{2b-3}\mathbf{X}_b + \mathbf{X}_{2b-1}\mathbf{X}_{b-2}.$$

Proposition. Assume that $0 < a < 2b$ for integers a, b . Then

$$R_{a,b} = \mathbf{X}_a\mathbf{X}_b + \sum_{0 \leq k} \binom{b-1-k}{a-2k} \mathbf{X}_{a+b-k}\mathbf{X}_k.$$

Proof. The proof is by induction on the quantity $2b - a$. If $2b - a \leq 3$, then the result follows from the lemma and the cases above. Assume then that $2b - a > 3$ and that the conclusion holds in all cases with lower values of $2b - a$. As $R_{a,b} = R_{a+1,b-1} + K(R_{a+1,b})$, by induction

$$\begin{aligned}
R_{a,b} &= \mathbf{X}_{a+1}\mathbf{X}_{b-1} + \sum \binom{(b-1)-1-k}{(a+1)-2k} \mathbf{X}_{a+b-k}\mathbf{X}_k \\
&+ K(\mathbf{X}_{a+1}\mathbf{X}_b + \sum \binom{b-1-k}{(a+1)-2k} \mathbf{X}_{a+b-k+1}\mathbf{X}_k) = \\
&= \mathbf{X}_{a+1}\mathbf{X}_{b-1} + \sum \binom{(b-1)-1-k}{(a+1)-2k} \mathbf{X}_{a+b-k}\mathbf{X}_k \\
&\quad + \mathbf{X}_a\mathbf{X}_b + \sum \binom{b-1-k}{(a+1)-2k} \mathbf{X}_{a+b-k}\mathbf{X}_k \\
&\quad + \mathbf{X}_{a+1}\mathbf{X}_{b-1} + \sum \binom{b-1-k}{(a+1)-2k} \mathbf{X}_{a+b-k+1}\mathbf{X}_{k-1}.
\end{aligned}$$

Thus,

$$\begin{aligned}
R_{a,b} &= \mathbf{X}_a\mathbf{X}_b + \\
&+ \sum \left(\binom{(b-1)-1-k}{(a+1)-2k} + \binom{b-1-k}{(a+1)-2k} + \binom{b-1-(k+1)}{(a+1)-2(k+1)} \right) \mathbf{X}_{a+b-k}\mathbf{X}_k.
\end{aligned}$$

By the lemma the coefficient is $\binom{b-1-k}{a-2k}$ modulo 2 as required. \square

The ideal generated by the *Adem-Wu relators* $R_{a,b}$ is denoted here by I_{AW} . The proposition implies that, modulo I_{AW} , each inadmissible is equivalent to a sum of admissibles, a fact proved in the next section. It also makes possible the identification of I_{AW} as follows.

Corollary. *In $\mathbf{F}[\mathbf{X}]$, the $\mathbf{F}[K]$ -submodule $\sum_{1 \leq b} \mathbf{F}[K]\mathbf{X}_{2b-1}\mathbf{X}_b$ is the linear subspace spanned by the Adem-Wu relators. The ideal I_{AW} is the ideal of $\mathbf{F}[\mathbf{X}]$ generated by this module and as such admits K .*

Proof. The first point follows from the facts that K applied to an Adem-Wu relator is a sum of Adem-Wu relators and that $\mathbf{X}_{2b-1}\mathbf{X}_b$ is an Adem-Wu relator. The last is a consequence of the formula for the action of K on a product. \square

The next result is a straightforward corollary of Cartan's formula. It serves to show that $\text{Ker}\rho$ admits K .

Lemma. *Let $F \in \mathbf{F}[\mathbf{X}]$, $f \in \mathbf{F}[\mathbf{x}]$ and $i \geq 1$. Then*

$$F(fx_i) \equiv F(f)x_i + K(F)(f)x_i^2$$

modulo the principal ideal generated by x_i^4 .

Proof. By linearity we may assume that F is a monomial $\mathbf{X}_{\mathbf{m}}$, $\mathbf{m} \geq \mathbf{0}$, and so must show that

$$\mathbf{X}_{\mathbf{m}}(fx_i) \equiv \mathbf{X}_{\mathbf{m}}(f)x_i + K(\mathbf{X}_{\mathbf{m}})(f)x_i^2 = \mathbf{X}_{\mathbf{m}}(f)x_i + \sum_j \mathbf{X}_{\mathbf{m}(j)}(f)x_i^2.$$

By Cartan's formula,

$$\mathbf{X}_{\mathbf{m}}(fx_i) = \sum_{\mathbf{k}+\boldsymbol{\ell}=\mathbf{m}} \mathbf{X}_{\mathbf{k}}(f)\mathbf{X}_{\boldsymbol{\ell}}(x_i),$$

which is

$$\mathbf{X}_{\mathbf{m}}(f)\mathbf{X}_{\mathbf{0}}(x_i) + \sum_j \mathbf{X}_{\mathbf{m}^{(j)}}(f)\mathbf{X}_{\mathbf{1}}(x_i) = \mathbf{X}_{\mathbf{m}}(f)x_i + \sum_j \mathbf{X}_{\mathbf{m}^{(j)}}(f)x_i^2$$

plus the terms of the form $\mathbf{X}_{\mathbf{k}}(f)\mathbf{X}_{\boldsymbol{\ell}}(x_i)$ in which $|\boldsymbol{\ell}| \geq 2$. For such a term $\mathbf{X}_{\boldsymbol{\ell}}(x_i) = 0$ unless $\boldsymbol{\ell}_{\ell}(\boldsymbol{\ell}) = 1$ and the previous non-zero entry of $\boldsymbol{\ell}$ is 2 whence it follows that x_i^4 divides $\mathbf{X}_{\boldsymbol{\ell}}(x_i)$. \square

If F is a monomial of length 1, i.e., $F = \mathbf{X}_m$, $m > 0$, then the equivalence is an equality. This result is not difficult to derive directly, and another proof of Cartan's formula can be based on it.

Corollary. *The ideal $\text{Ker}\rho$ admits K and so $I_{\text{AW}} \subseteq \text{Ker}\rho$.*

Proof. Take $F \in \text{Ker}\rho$, $f \in \mathbf{F}[\mathbf{x}]$ and $i \notin \text{supp } f$. Then the lemma implies that

$$K(F)(f)x_i^2 + gx_i^4 = 0$$

for some $g \in \mathbf{F}[\mathbf{x}]$ whence, as $\text{supp } K(F)(f) \subseteq \text{supp } f$, $K(F)(f) = 0$ as required. The last point follows from the theorem of the previous section. \square

4. Serre's theorem. The Steenrod algebra

We begin by demonstrating the spanning aspect of Serre's theorem.

Proposition. *Each element of $\mathbf{F}[\mathbf{X}]$ is equivalent to a sum of admissible monomials modulo the ideal I_{AW} . More precisely, for a vector \mathbf{m} there is a set A of admissible vectors such that $\mathbf{X}_{\mathbf{m}} \equiv \sum_{\mathbf{a} \in A} \mathbf{X}_{\mathbf{a}}$ modulo I_{AW} and, for all $\mathbf{a} \in A$, $\mathbf{a} \geq \mathbf{m}$ in the left lexicographic ordering.*

Proof. We show that the second statement is true for all \mathbf{m} of degree d by induction using the left lexicographic ordering. The maximal element here is the admissible vector $d\mathbf{e}_1 = (d, 0, \dots)$ for which the monomial is the indeterminate \mathbf{X}_d . Assume that the result is true for all \mathbf{n} in degree d for which $\mathbf{m} < \mathbf{n}$.

If \mathbf{m} is inadmissible, then there is an i for which $m_i < 2m_{i+1}$. But, modulo I_{AW} , $\mathbf{X}_{\mathbf{m}} = \mathbf{X}_{m_1} \cdots \mathbf{X}_{m_{i-1}} \mathbf{X}_{m_i} \mathbf{X}_{m_{i+1}} \cdots$ is equivalent to

$$\sum_{k < m_{i+1}} \binom{m_{i+1} - 1 - k}{m_i - 2k} \mathbf{X}_{m_1} \cdots \mathbf{X}_{m_{i-1}} \mathbf{X}_{m_i + m_{i+1} - k} \mathbf{X}_k \cdots$$

As $m_i + m_{i+1} - k > m_i$, each of these monomials exceeds \mathbf{m} in left lexicographic order and so, by the induction hypothesis, each is equivalent to a sum of yet greater admissibles and the result follows. \square

The identification of our two ideals follows as an immediate corollary.

Theorem. *The ideal $\text{Ker}\rho$ is contained in I_{AW} . Thus $\text{Ker}\rho = I_{\text{AW}}$.*

Proof. Let $F \in \text{Ker}\rho$. By the previous result $F \equiv \sum_{\mathbf{a} \in A} \mathbf{X}_{\mathbf{a}}$ modulo I_{AW} for a set A of admissibles. By Serre's theorem the admissibles are independent modulo $\text{Ker}\rho$ so that A is empty and $F \in I_{\text{AW}}$. \square

Another corollary is Serre's criterion for an element of $\mathbf{F}[\mathbf{X}]$ to act trivially [Se53; Wo98, Theorem 1.5].

Theorem. [Serre.] *An element F of $\mathbf{F}[\mathbf{X}]$ of degree d belongs to $\text{Ker}\rho$ if and only if $F(x_1x_2 \cdots x_d) = 0$.*

Proof. Suppose that $F(x^{1^d}) = 0$. If $F \notin \text{Ker}\rho$, then there is a non-empty set A of admissibles for which $F \equiv \sum_{\mathbf{a} \in A} \mathbf{X}_{\mathbf{a}}$ modulo $\text{Ker}\rho$. By homogeneity $d = \deg \mathbf{a}$ for each $\mathbf{a} \in A$. Thus, by the proof of the proposition in Section 1, $F(x^{1^d}) \neq 0$, contrary to assumption. \square

The *Steenrod algebra* A may now be defined as the quotient

$$A := \mathbf{F}[\mathbf{X}]/\text{Ker}\rho = \mathbf{F}[\mathbf{X}]/I_{\text{AW}}.$$

The images of the generators \mathbf{X}_i are called the *Steenrod squares* and usually denoted Sq^i ; analogous notations are used, e.g., $\text{Sq}^{\mathbf{m}} = \text{Sq}^{m_1}\text{Sq}^{m_2} \cdots$.

Serre's Theorem. *The elements $\text{Sq}^{\mathbf{a}}$ with \mathbf{a} admissible form a basis for the Steenrod algebra.*

References

- [AS72] Abramowitz, M.; Stegun, I.A. (eds.): Handbook of mathematical functions. New York: Dover (1972)
- [Ca50] Cartan, H.: Une théorie axiomatique des carrées de Steenrod, C. R. Acad. Sci. Paris **230**, 425–427 (1950)
- [Gr75] Gray, B.: Homotopy theory. London: Academic Press 1975
- [Kr63] Kristensen, L. (1963): On secondary cohomology operations. Math. Scand. **12**, 57–82
- [Kr65] Kristensen, L. (1965): On a Cartan formula for secondary cohomology operations. Math. Scand. **16**, 97–115
- [Ma99] Massey, W.S.: A history of cohomology theory. In: James, I.M. (ed.): History of topology. Amsterdam: Elsevier 1999, pp. 579–603
- [Se53] Serre, J.-P.: Cohomologie modulo 2 des complexes d'Eilenberg–MacLane. Comment. Math. Helv. **27**, 198–231 (1953)
- [Si97] Silverman, J: Stripping and conjugation in the Steenrod algebra. J. Pure Appl. Algebra **121**, 95–106. (1997)
- [SE62] Steenrod, N.E., Epstein, D.B.A.: Cohomology operations. (Ann. Math. Stud. **50**) Princeton: Princeton University Press 1962
- [Wo98] Wood, R.M.W.: Problems in the Steenrod algebra. Bull. London Math. Soc. **30**, 449–517 (1998)