Polynomials and the Steenrod Algebra

Grant Walker and Reg Wood

last updated 25 March 2010

Preface

This book investigates the Steenrod algebra $\mathcal{A}_2$ over the field of two elements $\mathbb{F}_2$ in a purely algebraic context by its action on the polynomial algebra $P(n)$ in n variables over $\mathbb{F}_2$ and treats the 'hit' problem of finding the irreducibles of this action as a recurrent theme throughout. The reader is expected to have a basic knowledge of algebra. Of particular importance is the action of the semigroup $M(n)$ of $n \times n$ matrices over $\mathbb{F}_2$ on $P(n)$ which is intimately related to the action of $\mathcal{A}_2$. Properly formulated the irreducibles or 'cohits', as we shall call them, in various degrees form interesting modular representations of the general linear group $GL(n) \subset M(n)$. This is a difficult area and the authors have tried to make the book as self contained as possible. Frequently, definitions and constructions are introduced and first explored for a small number of variables and extended to the general case in later chapters. Although this can involve a certain amount of repetition it has the advantage of leading to interesting results at an early stage by elementary methods. The action of matrices on polynomials can be treated in the rather sophisticated context of modular representation theory of algebraic groups and semigroups, but it is often difficult to find in the literature specific applications to the general linear groups of low order.

Topology does not feature in this book except by way of reference in some of the remarks sections at the ends of the chapters and in the appendices. This may seem surprising because the Steenrod algebra is at the heart of algebraic topology and has been one of the main tools in the development and problem-solving capability of cohomology theory ever since Steenrod's introduction of the Steenrod squaring operations in 1947. The book of Steenrod and Epstein [142], based on lectures by Steenrod, has been a mainstay in learning about the Steenrod algebra from a topological point of view. In fact the hit problem itself first arose in a topological context to do with classical cobordism theory and was first posed by Frank Peterson in 1986 at a meeting of the American Mathematical Society.

The Steenrod algebra is the algebra of universal stable operations for cohomology theory over the field of two elements. In a sense this book deals with just one example of a topological space, namely the product of n copies of infinite real projective space whose cohomology is $P(n)$. However, this example is of paramount importance because a knowledge of the action of $\mathcal{A}_2$ on $P(n)$ for all n determines $\mathcal{A}_2$ itself. But there is another reason for studying the Steenrod algebra in a purely algebraic context. The hit problem can be formulated for any graded module over a graded algebra. There are $\mathcal{A}_2$-modules, derived from $P(n)$, which cannot be the cohomology of any topological space, for example Dickson algebras of invariants of $GL(n)$ for n sufficiently large. This gives wider scope for interesting hit problems by looking at invariants and eigenspaces of certain subgroups of the general linear groups. We shall investigate the hit problem for symmetric groups and certain cyclic groups. We also include an elementary proof of part of the Adams-Gunawardena-Miller theorem at the prime 2 which states

that the self linear transformations of $P(n)$ preserving degree and commuting with the action of $\mathcal{A}_2$ are given by linear substitution, i.e. the action of $M(n)$.

More is known about the hit problem for the prime 2 than for the odd prime analogue. This explains why the book concentrates on the field of two elements. It would require several more volumes to cope adequately with the broader aspects of the hit problem and its applications to topology. We can only offer a brief guide to further reading in the literature and draw attention in a later chapter to some of the outstanding problems in the subject.

Contents

1 Steenrod squares **7**

1.0 Introduction 7
1.1 Steenrod operations Sq^k on $P(n)$ 8
1.2 The action of $M(n)$ on $P^d(n)$ 10
1.3 Further properties of Sq^k 11
1.4 The hit problem 12
1.5 Binomial coefficients mod 2 14
1.6 Spikes 15
1.7 Maps of cohit modules 16
1.8 The hit problem for $P(2)$ 18
1.9 Remarks 20

2 Conjugate Steenrod squares **21**

2.0 Introduction 21
2.1 Formal power series 22
2.2 Conjugate Steenrod operations Xq^k 24
2.3 The functions α, ω and μ 25
2.4 Evaluating $Xq^k(f)$ 27
2.5 The Peterson conjecture 29
2.6 Remarks 31

3 The Steenrod algebra $\mathcal{A}_2$ **33**

3.0 Introduction 33
3.1 The Adem relations 33
3.2 The action of $\mathcal{A}_2$ on $P(n)$ 36
3.3 The admissible basis 38
3.4 The Milnor basis 43
3.5 Remarks 46

4 Products and conjugation in $\mathcal{A}_2$ **47**

4.0 Introduction 47
4.1 The Milnor product formula 47
4.2 The Bullett-Macdonald identity 51

4.3 The conjugation χ 53
4.4 Conjugation and the Milnor basis 55
4.5 Remarks 58

5 Combinatorial structures **59**
5.0 Introduction 59
5.1 Vectors and partitions 60
5.2 Dominance 63
5.3 Vectors of degree d 65
5.4 Partitions of degree d 67
5.5 Minimal spikes 69
5.6 Maximal elements of $\mathbb{V}^d(n)$ and $\mathbb{W}^d(n)$ 71
5.7 Dominance in the Steenrod algebra 74
5.8 The excess function 77
5.9 Remarks 79

6 Filtrations on $Q^d(n)$ **81**
6.0 Introduction 81
6.1 Steenrod and matrix actions on blocks 82
6.2 Reducibility and the spaces $Q^W(n)$ 85
6.3 Concatenation of blocks 87
6.4 Splicing 91
6.5 The Kameko and duplication maps 92
6.6 Head ω-vectors 96
6.7 Remarks 100

A Differential operators **101**

B Catalan numbers and Xq^k **103**

C Blocks and ω-vectors **105**

D Topological applications **107**

Chapter 1

Steenrod squares

1.0 Introduction

In this chapter we introduce our main subject, the algebra of polynomials in n variables over the field of two elements $\mathbb{F}_2$ under the left action of linear operations called **Steenrod squares** and the right action of $n \times n$ matrices. We denote this polynomial algebra by $P(n) = \mathbb{F}_2[x_1, \ldots, x_n]$ for $n \geq 1$. For $n = 1$, 2, 3, it is convenient to use x, y, z for the variables x_1, x_2, x_3. The algebra $P(n)$ is graded by the vector spaces $P^d(n)$ of homogeneous polynomials of degree $d \geq 0$.

In Section 1.1 the Steenrod squares $Sq^k : P^d(n) \to P^{d+k}(n)$ are defined for $k \geq 0$ and their basic properties are established. In particular, Sq^0 is the identity map. In Section 1.2 we explain how $P^d(n)$ is a right module over the semigroup algebra $\mathbb{F}_2 M(n)$, where $M(n) = M(n, \mathbb{F}_2)$ is the multiplicative semigroup of $n \times n$ matrices over $\mathbb{F}_2$. This right action commutes with the left action of the Steenrod squares. By restricting to non-singular matrices, $P^d(n)$ affords a modular representation of the general linear group $GL(n) = GL(n, \mathbb{F}_2)$, and the Steenrod squares are maps of $\mathbb{F}_2 GL(n)$-modules. Further properties of the Steenrod squares are developed in Section 1.3.

In Section 1.4 we introduce the **hit problem**. We call a polynomial **hit** if it is a linear combination of elements in the images of positive Steenrod squares. The hit elements form a $\mathbb{F}_2 M(n)$-submodule $H(n)$ of $P(n)$. The corresponding quotient $Q(n) = P(n)/H(n)$ is the **cohit module**, and the hit problem is to determine $Q(n)$. The cohit modules are of interest in both representation theory and topology. We give the solution of the hit problem for $P(1) = \mathbb{F}_2[x]$ in Section 1.5. In Section 1.6 we define **spike** monomials. These are the monomials which can not appear in the image of a positive Steenrod square. In Sections 1.7 and 1.8 we give the solution of the hit problem for $P(2) = \mathbb{F}_2[x, y]$, using the **Kameko map** and the **duplication map** to connect the cohit modules. The remarks at the end of the chapter provide background information.

1.1 Steenrod operations Sq^k on $P(n)$

Definition 1.1.1. For $n \geq 1$, the algebra $P(n)$ is the polynomial algebra

$$P(n) = \mathbb{F}_2[x_1, \ldots, x_n]$$

in n variables $x_1, \ldots, x_n$ over the field $\mathbb{F}_2 = \{0, 1\}$. It is graded by degree d, so that $P(n) = \sum_{d \geq 0} P^d(n)$, where $P^d(n)$ is the vector space of homogeneous polynomials of degree d. The monomials $x_1^{d_1} \cdots x_n^{d_n}$ such that $d_1 + \cdots + d_n = d$ and $d_i \geq 0$ for $1 \leq i \leq n$ form a basis for $P^d(n)$. If all exponents $d_i = 0$, this is the identity element 1 of $P(n)$, and we identify $P^0(n)$ with $\mathbb{F}_2$. For $n = 0$, we also identify $P(0)$ with $\mathbb{F}_2$.

Since $P(n)$ is freely generated as a commutative algebra by the x_i, an algebra map $\phi : P(n) \to P(n)$ is defined uniquely by assigning a value $\phi(x_i)$ to each variable x_i, $1 \leq i \leq n$. We shall always assume that $\phi(1) = 1$.

Definition 1.1.2. The **total Steenrod square** $Sq : P(n) \to P(n)$ is the algebra map defined by

$$Sq(1) = 1, \quad Sq(x_i) = x_i + x_i^2, \quad \text{for } 1 \leq i \leq n.$$

The **Steenrod squares** $Sq^k : P^d(n) \to P^{d+k}(n)$, for $k \geq 0$ and $d \geq 0$, are the linear maps defined by restricting Sq to $P^d(n)$ and projecting on to $P^{d+k}(n)$. Thus $Sq = \sum_{k \geq 0} Sq^k$ is the formal sum of its graded parts.

Proposition 1.1.3. *For all $x \in P^1(n)$, $Sq(x) = x + x^2$. Thus $Sq^0(x) = x$, $Sq^1(x) = x^2$ and $Sq^k(x) = 0$ for all $k > 1$.*

Proof. Let $x = \sum_{i=1}^n a_i x_i$, where $a_i \in \mathbb{F}_2$. Then $Sq(x) = \sum_{i=1}^n a_i Sq(x_i) = \sum_{i=1}^n a_i(x_i + x_i^2) = \sum_{i=1}^n a_i x_i + \sum_{i=1}^n (a_i x_i)^2 = x + x^2$. The second statement follows by equating graded parts. □

The following is the most important rule for calculating with Steenrod squares.

Proposition 1.1.4. (Cartan formula) *For polynomials $f, g \in P(n)$ and $k \geq 0$,*

$$Sq^k(fg) = \sum_{i+j=k} Sq^i(f) Sq^j(g).$$

Proof. This follows from the multiplicative property $Sq(fg) = Sq(f)Sq(g)$ by equating terms of degree k. □

Proposition 1.1.5. *Sq^0 is the identity map of $P(n)$.*

Proof. Setting $k = 0$ in Proposition 1.1.4, Sq^0 is an algebra map of $P(n)$. Since $Sq^0(1) = 1$ and $Sq^0(x_i) = x_i$ for $1 \leq i \leq n$, $Sq^0(f) = f$ for all $f \in P(n)$. □

By a **Steenrod operation** we mean a linear transformation $\theta : P(n) \to P(n)$ which can be obtained from the Sq^k by the processes of addition and composition. In principle, the above rules allow the evaluation of any Steenrod operation on a polynomial by linearity and recursive use of the Cartan formula.

Example 1.1.6. In $P(2) = \mathbb{F}_2[x, y]$ we have

$$Sq^1(xy) = Sq^1(x)Sq^0(y) + Sq^0(x)Sq^1(y) = x^2y + xy^2.$$

The next two results show how to evaluate a Steenrod square on a monomial.

Proposition 1.1.7. *For all $x \in P^1(n)$,*

$$Sq^k(x^d) = \binom{d}{k} x^{d+k},$$

where the binomial coefficient is reduced mod 2.

Proof. By the multiplicative property of Sq, $Sq(x^d) = (Sq(x))^d = (x + x^2)^d = x^d(1 + x)^d = \sum_{k=0}^{d} \binom{d}{k} x^{d+k}$. The result follows by equating terms of degree $d + k$. □

Proposition 1.1.8. *Let $f = x_1^{d_1} \cdots x_n^{d_n}$ be a monomial in $P(n)$. Then*

$$Sq^k(f) = \sum_{k_1 + \cdots + k_n = k} Sq^{k_1}(x_1^{d_1}) \cdots Sq^{k_n}(x_n^{d_n}).$$

Proof. This follows by induction on n using the Cartan formula 1.1.4. □

Although the above formulae theoretically solve the problem of evaluating Steenrod operations on polynomials, they are not very efficient for treating the hit problem (see Section 1.0), and much work in this book is concerned with the development of more practical ways of evaluating the operations on specific types of polynomials and finding criteria for elements to be hit. The next result explains why Sq^k is called a squaring operation.

Proposition 1.1.9. *For $f \in P^d(n)$, $Sq^k(f) = 0$ for $k > d$ and $Sq^d(f) = f^2$.*

Proof. Since Sq^k is linear and $(f + g)^2 = f^2 + g^2$, we may assume that f is a monomial of degree d. We use induction on d. The base case $d = 0$ holds since $Sq(1) = 1$. For $d > 0$, let $f = xg$, where x is one of the variables x_i and g is a monomial of degree $d - 1$. By the Cartan formula 1.1.4, $Sq^k(f) = xSq^k(g) + x^2Sq^{k-1}(g)$ for $k > 0$. If $k > d$, then by the inductive hypothesis $Sq^k(g) = 0$ and $Sq^{k-1}(g) = 0$, so $Sq^k(f) = 0$. If $k = d$, then by the inductive hypothesis $Sq^k(g) = 0$ and $Sq^{k-1}(g) = g^2$, so $Sq^k(f) = x^2g^2 = f^2$. This establishes the inductive step. □

Proposition 1.1.10. *For all monomials $f \in P(n)$ and all $k \geq 0$, every monomial in $Sq^k(f)$ involves exactly the same variables as f does.*

Proof. The variable x_i is involved in f if and only if $d_i > 0$. If $d_i > 0$, then $k + d_i > 0$ for all k. If $d_i = 0$, then $Sq(x_i^{d_i}) = Sq(1) = 1$. □

1.2 The action of $M(n)$ on $P^d(n)$

Definition 1.2.1. For $n \geq 1$, $GL(n) = GL(n, \mathbb{F}_2)$ is the general linear group of non-singular $n \times n$ matrices, and $M(n) = M(n, \mathbb{F}_2)$ the semigroup of all $n \times n$ matrices, over $\mathbb{F}_2$. For $A = (a_{i,j}) \in M(n)$, A acts on a variable $x_i \in P(n)$ by

$$x_i \cdot A = \sum_{j=1}^{n} a_{i,j} x_j, \quad 1 \leq i \leq n.$$

This action is extended to all polynomials $f \in P(n)$ by substitution, so that $(f \cdot A)(x_1, \ldots, x_n) = f(x_1 \cdot A, \ldots, x_n \cdot A)$.

Let $f, g \in P(n)$ and $A, B \in M(n)$, and let $I_n \in M(n)$ be the identity matrix. Since $1 \cdot A = 1$, $(f+g) \cdot A = f \cdot A + g \cdot A$ and $(fg) \cdot A = (f \cdot A)(g \cdot A)$, $f \mapsto f \cdot A$ is an algebra map of $P(n)$. Since $f \cdot I_n = f$ and $f \cdot (AB) = (f \cdot A) \cdot B$, $P(n)$ is a right $\mathbb{F}_2 M(n)$-module.

Example 1.2.2. The group $GL(2)$ is generated by the two matrices

$$S = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \text{ and } T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},$$

which act on the variables x, y by $x \cdot S = y$, $y \cdot S = x$ and $x \cdot T = x + y$, $y \cdot T = y$. Evaluating in degree 2, we obtain $xy \cdot S = xy$, $x^2 \cdot S = y^2$, $y^2 \cdot S = x^2$ and $xy \cdot T = xy + y^2$, $x^2 \cdot T = x^2 + y^2$, $y^2 \cdot T = y^2$. Evaluating in degree 3, we obtain $x^3 \cdot S = y^3$, $y^3 \cdot S = x^3$, $xy^2 \cdot S = x^2 y$, $x^2 y \cdot S = xy^2$ and $x^3 \cdot T = x^3 + y^3 + xy^2 + x^2 y$, $y^3 \cdot T = y^3$, $xy^2 \cdot T = xy^2 + y^3$, $x^2 y \cdot T = x^2 y + y^3$.

Since the order of the group $GL(n)$ is divisible by 2 for all $n > 1$, $P^d(n)$ affords a modular representation of $GL(n)$. The module $P^1(n)$ gives the natural representation of dimension n, and is a simple module, but in general the structure of $P^d(n)$ is complicated. For example, since $(f+g)^2 = f^2 + g^2$ in $P(n)$, the squaring operation $f \mapsto f^2$ embeds $P^d(n)$ as a submodule in $P^{2d}(n)$ for all d.

Next we prove that the right action of $\mathbb{F}_2 M(n)$ on $P(n)$ commutes with the left action of the Steenrod operations Sq^k. This is central to our whole subject.

Proposition 1.2.3. *For $f \in P(n)$ and $A = (a_{i,j}) \in M(n)$, $Sq(f) \cdot A = Sq(f \cdot A)$.*

Proof. Since Sq and $f \mapsto f \cdot A$ are algebra maps of $P(n)$, we need only check this when f is one of the variables x_i. Then $Sq(x_i) \cdot A = (x_i + x_i^2) \cdot A = x_i \cdot A + (x_i \cdot A)^2$, while $Sq(x_i \cdot A) = Sq(\sum_{j=1}^n a_{i,j} x_j) = \sum_{j=1}^n a_{i,j} Sq(x_j) = \sum_{j=1}^n a_{i,j}(x_j + x_j^2)$. But $\sum_{j=1}^n a_{i,j} x_j^2 = (\sum_{j=1}^n a_{i,j} x_j)^2 = (x_i \cdot A)^2$, so $Sq(x_i) \cdot A = Sq(x_i \cdot A)$. □

In particular, $Sq^k : P^d(n) \to P^{k+d}(n)$ is a map of $\mathbb{F}_2 G$-modules for any subgroup G of $GL(n)$, such as the group of matrices which permute the variables x_i. It also commutes with specializations of the variables given by singular matrices. For example, $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ sets $x = 0$ and $\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$ sets $y = x$ when $n = 2$.

1.3 Further properties of Sq^k

The action of Steenrod operations on polynomials can be reproduced on squares of polynomials by doubling exponents of the operators.

Proposition 1.3.1. *For $f \in P(n)$ and $k \geq 0$, $Sq^{2k}(f^2) = (Sq^k(f))^2$ and $Sq^{2k+1}(f^2) = 0$.*

Proof. By the multiplicative property of Sq,

$$\sum_{d\geq 0} Sq^d(f^2) = \sum_{i\geq 0} Sq^i(f) \sum_{j\geq 0} Sq^j(f).$$

Since $Sq^i(f)$ and $Sq^j(f)$ commute, the terms of degree d vanish in pairs, except for the terms $Sq^k(f)Sq^k(f)$ when $d = 2k$. □

Proposition 1.3.2. *For $f \in P(n)$ and $s, k \geq 0$,*

$$Sq^k(f^{2^s}) = \begin{cases} (Sq^j(f))^{2^s}, & \text{if } k = 2^s j, \\ 0, & \text{otherwise.} \end{cases}$$

In particular, $Sq^k(x^{2^s})$ is x^{2^s} if $k = 0$, $x^{2^{s+1}}$ if $k = 2^s$, and is 0 otherwise.

Proof. This follows from Proposition 1.3.1 by induction on s. □

Proposition 1.3.3. *For $f, g \in P(n)$ and $s, k \geq 0$,*

$$Sq^k(gf^{2^s}) = \sum_{i+2^s j=k} Sq^i(g)(Sq^j(f))^{2^s}.$$

In particular, $Sq^k(gf^{2^s}) = Sq^k(g)f^{2^s}$ if $k < 2^s$.

Proof. This is a consequence of the Cartan formula and Proposition 1.3.2. The second statement is immediate from the first and Proposition 1.1.9. □

Proposition 1.3.4. *For $f \in P(n)$ and $k \geq 0$, $Sq^1Sq^{2k}(f) = Sq^{2k+1}(f)$ and $Sq^1Sq^{2k+1}(f) = 0$.*

Proof. By linearity, we may assume that f is a monomial of degree d. We use induction on d. The base case $d = 0$ is true since $Sq(1) = 1$. For $d > 0$ let $f = xg$ for some variable $x = x_i$. Then by the Cartan formula 1.1.4, $Sq^j(f) = xSq^j(g) + x^2Sq^{j-1}(g)$, and

$$Sq^1Sq^j(f) = xSq^1Sq^j(g) + x^2Sq^j(g) + x^2Sq^1Sq^{j-1}(g).$$

Setting $j = 2k$ and $j = 2k+1$ and using the inductive hypothesis on g to obtain $Sq^1Sq^{2k}(g) = Sq^{2k+1}(g)$, $Sq^1Sq^{2k-1}(g) = 0$ and $Sq^1Sq^{2k+1}(g) = 0$, we have

$$\begin{aligned} Sq^1Sq^{2k}(f) &= xSq^{2k+1}(g) + x^2Sq^{2k}(g) &&= Sq^{2k+1}(f), \\ Sq^1Sq^{2k+1}(f) &= x^2Sq^{2k+1}(g) + x^2Sq^{2k+1}(g) &&= 0. \end{aligned}$$

This establishes the inductive step. □

The last result of this section shows that the action of a Steenrod square commutes with partial differentiation. In Appendix A we explain how the Steenrod squares may be interpreted in terms of differential operators.

Proposition 1.3.5. *For $f \in P^d(n)$ and a variable $x = x_i$ we have*

$$Sq^k(\frac{\partial f}{\partial x}) = \frac{\partial}{\partial x}(Sq^k(f)).$$

Proof. The result is clear for $f = 1$ and $f = x$, and for $k = 0$. As in Proposition 1.3.4, we may assume by induction on degree that f is a monomial of the form $f = xg$ and that the result is proved for g. By the Cartan formula 1.1.4, $Sq^k(\frac{\partial}{\partial x}(xg)) = Sq^k(g + x\frac{\partial g}{\partial x}) = Sq^k(g) + xSq^k(\frac{\partial g}{\partial x}) + x^2Sq^{k-1}(\frac{\partial g}{\partial x})$ for $k > 0$. On the other hand, $\frac{\partial}{\partial x}(Sq^k(xg)) = \frac{\partial}{\partial x}(xSq^k(g) + x^2Sq^{k-1}(g)) = Sq^k(g) + x\frac{\partial}{\partial x}(Sq^k(g)) + x^2\frac{\partial}{\partial x}(Sq^{k-1}(g))$. By the inductive hypothesis on g it follows that $Sq^k(\frac{\partial f}{\partial x}) = \frac{\partial}{\partial x}(Sq^k(f))$, completing the inductive step. □

1.4 The hit problem

The Steenrod squaring operations allow us to express many elements of $P^d(n)$ in terms of polynomials of lower degree. For example, $x^5 = Sq^2(x^3)$, $x^4y = Sq^2(x^2y)$ and $x^3y^2 = Sq^1(x^3y) + Sq^2(x^2y)$. By exchanging the variables x and y and taking linear combinations, it is clear that every element of $P^5(2)$ can be written in this form.

Definition 1.4.1. A polynomial $f \in P^d(n)$ is **hit** if it satisfies a **hit equation**

$$f = \sum_{k=1}^{d} Sq^k(f_k), \text{ where } f_k \in P^{d-k}(n).$$

The set of hit polynomials (or **hits** for short) in $P^d(n)$ is a vector subspace $H^d(n)$. For $f, g \in P^d(n)$ we write $f \sim g$ if $f + g$ is hit.

Of course, since we are working mod 2, $f + g = f - g$. Thus $f \sim g$ is an equivalence relation on $P^d(n)$, the equivalence classes being cosets of $H^d(n)$. By Proposition 1.2.3, the Steenrod squares commute with the action of $M(n)$, and so $H^d(n)$ is an $\mathbb{F}_2M(n)$-submodule of $P^d(n)$.

Definition 1.4.2. The **hit problem** is to determine the **cohit module**

$$Q^d(n) = P^d(n)/H^d(n)$$

for each $n \geq 1$ and $d \geq 0$.

The hit problem can be put at different levels. The most basic question is to ask whether or not all homogeneous polynomials of degree d in $P(n)$ are hit. At the next level, we can ask for the dimension of $Q^d(n)$ as a vector space over $\mathbb{F}_2$. Further, we can ask for a basis of $Q^d(n)$. For example, as a quotient space of $P^d(n)$, a basis can be chosen from the hit equivalence classes of monomials. (The reader is warned that we shall normally refer to elements of $Q^d(n)$ as 'polynomials' when strictly speaking they are equivalence classes of polynomials.) Finally, we can seek information about $Q^d(n)$ as a $\mathbb{F}_2GL(n)$-module, or as a $\mathbb{F}_2M(n)$-module.

Example 1.4.3. In $P(1) = \mathbb{F}_2[x]$, x^3 is not hit because $Sq^2(x) = 0$ and $Sq^1(x^2) = 0$. Hence $Q^3(1)$ is 1-dimensional, generated by x^3. In $P(2) = \mathbb{F}_2[x, y]$ we have $Sq^1(x) = x^2$ and $Sq^1(y) = y^2$. Hence x^2 and y^2 are hit, but xy is not hit and so $Q^2(2)$ is 1-dimensional, generated by xy. The discussion at the start of this section shows that $Q^5(2) = 0$.

From Proposition 1.1.9 we have the following result.

Proposition 1.4.4. *For any $f \in P^d(n)$, where $d > 0$, f^2 is hit.* □

Proposition 1.4.5. *Let $g \in P^d(n)$ be hit and let $x \in P^1(n)$. Then xg^2 is hit.*

Proof. Let $g = \sum_{i>0} Sq^i(g_i)$ be a hit equation for g. By Proposition 1.3.1, $Sq^i(g_i)^2 = Sq^{2i}(g_i^2)$, and by the Cartan formula 1.1.4, $xSq^{2i}(g_i^2) = Sq^{2i}(xg_i^2)$. Hence $xg^2 = \sum_{i>0} x(Sq^i(g_i)^2) = \sum_{i>0} Sq^{2i}(xg_i^2)$. □

As far as the vector space structure of $Q(n)$ is concerned, the hit problem can be decomposed into smaller problems in the following way. For each subset $Y \subseteq \{1, \ldots, n\}$, let $P(Y)$ be the subspace of $P(n)$ spanned by monomials which are divisible by x_i if and only if $i \in Y$. (If $Y = \emptyset$, then $P(Y) = P(0) = \mathbb{F}_2$.) By Proposition 1.1.10, $P(Y)$ is preserved by the action of the Steenrod squares, and so we have a corresponding vector space $H(Y)$ of hit elements and quotient space $Q(Y) = P(Y)/H(Y)$. Then as vector spaces over $\mathbb{F}_2$,

$$P(n) = \oplus_Y P(Y), \quad H(n) = \oplus_Y H(Y), \quad Q(n) = \oplus_Y Q(Y).$$

If Y, Y' have the same cardinality, then $Q(Y) \cong Q(Y')$ by permuting variables and using Proposition 1.2.3. Since there are $\binom{n}{k}$ subsets Y with cardinality k, this gives the following result.

Proposition 1.4.6. *For $1 \leq k \leq n$, let $Y_k = \{1, 2, \ldots, k\}$ and let $P[k] = P(Y_k)$, $Q[k] = Q(Y_k)$. Then $\dim Q^d(n) = \sum_{k=1}^n \binom{n}{k} \dim Q^d[k]$ for $d > 0$.* □

Example 1.4.7. We compute $\dim Q^3(3)$. Since x^3 is not hit in $P(1)$ (Example 1.4.3), $\dim Q^3[1] = 1$. By Example 1.1.6, $xy^2 \sim x^2y$, but clearly these monomials are not hit, so $\dim Q^3[2] = 1$. Finally $\dim Q^3[3] = 1$ since the only monomial in $P^3[3]$ is xyz. Hence $\dim Q^3(3) = 3 + 3 + 1 = 7$.

1.5 Binomial coefficients mod 2

In general it is not easy to determine if a monomial is hit. However, the 1-variable case can be solved immediately using Proposition 1.1.7. For this, we need to evaluate binomial coefficients $\binom{a}{b}$ mod 2. This is done by comparing the binary expansions of a and b, which express a and b uniquely as the sum of distinct 2-powers.

Definition 1.5.1. For $a \geq 0$ let $a = 2^{j_1} + 2^{j_2} + \cdots + 2^{j_r}$, where $j_1 > j_2 > \cdots > j_r \geq 0$. Then

$$\mathrm{bin}(a) = \{2^{j_1}, 2^{j_2}, \ldots, 2^{j_r}\}$$

is the set of terms in the binary expansion of a.

Example 1.5.2. As $\mathrm{bin}(11) = \{1, 2, 8\}$, $\mathrm{bin}(12) = \{4, 8\}$, $\mathrm{bin}(27) = \{1, 2, 8, 16\}$, the next result provides a quick way to check that $\binom{27}{11}$ is odd and $\binom{27}{12}$ is even.

Proposition 1.5.3. *For $a \geq b \geq 0$, the binomial coefficient $\binom{a}{b}$ is odd if and only if* $\mathrm{bin}(b) \subseteq \mathrm{bin}(a)$.

Proof. Working in $\mathbb{F}_2[x]$, the term x^b appears in the expansion of

$$(1+x)^a = \prod_{i=1}^{r}(1+x)^{2^{j_i}} = \prod_{i=1}^{r}(1+x^{2^{j_i}}).$$

if and only if all summands in the binary expansion of b are summands in the binary expansion of a. □

We can now combine Propositions 1.1.7 and 1.5.3 to solve the case $n = 1$ of the hit problem. For this we need only determine the degrees d such that x^d is hit.

Theorem 1.5.4. *In $P(1) = \mathbb{F}_2[x]$, x^d is hit if and only if d is not of the form $2^j - 1$, where $j \geq 0$. Hence $Q^d(1) \cong \mathbb{F}_2$ if $d = 2^j - 1$, and $Q^d(1) = 0$ otherwise.*

Proof. By definition, x^d is hit if and only if $Sq^b(x^a) = x^d$, where $a + b = d$ and $a, b > 0$. By Proposition 1.1.7, $Sq^b(x^a) = \binom{a}{b}x^{a+b}$. Let $d = 2^j - 1$ and consider any decomposition $d = a + b$ where $a, b > 0$. Then $\mathrm{bin}(a)$ is a subset of $\mathrm{bin}(d) = \{1, 2, \ldots, 2^{j-1}\}$ and $\mathrm{bin}(b)$ is the complementary subset. By Proposition 1.5.3, $\binom{a}{b} = 0$ if $b > 0$. Hence x^d is not hit. On the other hand, if d is not of the form $2^j - 1$, then, for some i, 2^{i+1} is a term in the binary expansion of d, but 2^i is not. Let $b = 2^i$ and $a = d - 2^i$. Then $2^i \in \mathrm{bin}(a)$, and hence $\binom{a}{b} = 1$ by Proposition 1.5.3. Hence x^d is hit. □

1.6 Spikes

Theorem 1.5.4 has implications which reach beyond the case $n = 1$. In this section we apply it in two ways. The first application uses the fact that the hit polynomials form a $\mathbb{F}_2 M(n)$-module.

Proposition 1.6.1. *Let $A \in \mathbb{F}_2 M(n)$ and $f \in P(n)$. If $f \cdot A$ is not hit, then f is not hit.*

Proof. If $f = \sum_k Sq^k(f_k)$, then by Proposition 1.2.3 $Sq^k(f_k) \cdot A = Sq^k(f_k \cdot A)$, so $f \cdot A = \sum_k Sq^k(f_k \cdot A)$. Thus, if f is hit, then $f \cdot A$ is also hit. □

Proposition 1.6.2. *For $j \geq 0$, no monomial of degree $2^j - 1$ in $P(n)$ is hit.*

Proof. We specialize all the variables to a single variable x to obtain x^{2^j-1}. By Theorem 1.5.4, this is not hit, so the result follows from Proposition 1.6.1. □

The second application of Theorem 1.5.4 gives a special class of monomials which are of fundamental importance in the hit problem. Not only are they not hit, but they can not appear as terms in a polynomial in the image of Sq^k for $k > 0$ when it is expressed irredundantly as a sum of monomials.

Definition 1.6.3. A **spike** in $P(n)$ is a monomial of the form $x_1^{2^{j_1}-1} \cdots x_n^{2^{j_n}-1}$, where $j_1, \ldots, j_n \geq 0$.

Proposition 1.6.4. *For $k > 0$ and $f \in P(n)$, a spike cannot be a term in $Sq^k(f)$. Every monomial basis of $Q(n)$ contains all the spikes in $P(n)$.*

Proof. Let $x_1^{2^{j_1}-1} \cdots x_n^{2^{j_n}-1}$ be a spike. By Proposition 1.5.4, $x_i^{2^{j_i}-1}$ cannot be the ith factor $Sq^{k_i}(x^{d_i})$ of any term in the expansion of $Sq^k(x_1^{d_1} \cdots x_n^{d_n})$ by the Cartan formula 1.1.8 unless $k_i = 0$. Hence $k_i = 0$ for all i, and so $k = 0$. Hence a spike cannot appear as a term in a hit equation. Consequently the spikes in degree d are linearly independent modulo $H^d(n)$, and so any monomial basis of $Q^d(n)$ must contain them all. □

The next example shows that, even when $n = 2$, the spikes are not sufficient to give a basis for $Q(n)$.

Example 1.6.5. In $P(2) = \mathbb{F}_2[x, y]$ we have $Sq^1(x^2) = 0 = Sq^1(y^2)$, and Example 1.1.6 exhibits the only essential hit equation in $P^3(2)$. It follows that x^2y and xy^2 are not hit, and one of them, but not both, must be included in a monomial basis for $Q^3(2)$. Hence the dimensions of $Q^d(2)$ for $d = 1, 2, 3$ are $2, 1, 3$ respectively, with generating sets $\{x, y\}$, $\{xy\}$ and $\{x^3, x^2y, y^3\}$.

1.7 Maps of cohit modules

In this section we find criteria for an element in $P^d(2)$ to be hit, and define two important $\mathbb{F}_2GL(2)$-module maps $\kappa : Q^d(2) \to Q^{2d+2}(2)$ and $\delta : Q^{2d+1}(2) \to Q^{4d+3}(2)$, which will be generalized later to the n-variable case. A polynomial $f \in P^d(2)$ can be written uniquely as $f = xyg^2 + h^2$ if d is even and as $f = xg^2 + yh^2$ if d is odd. Thus from $xyg_1^2 + h_1^2 = xyg_2^2 + h_2^2$ or $xg_1^2 + yh_1^2 = xg_2^2 + yh_2^2$ we may deduce $g_1 = g_2$ and $h_1 = h_2$, by 'comparing coefficients' of xy and 1, or of x and y. We begin with the even degree case.

Proposition 1.7.1. *For $d \geq 1$, $f = xyg^2 + h^2$ is hit if and only if g is hit.*

Proof. If f is hit, then since h^2 is hit we may assume that $f = xyg^2$. Using Proposition 1.3.4 to collect the terms in a hit equation which involve odd Steenrod squares, we can write $f = Sq^1(f') + \sum_{i>0} Sq^{2i}(f'_i)$ where f' has odd degree and f'_i has even degree for all i. Then $f' = xg'^2 + yh'^2$ and $f'_i = xyg'^2_i + h'^2_i$. Hence by Proposition 1.3.1

$$f = xyg^2 = x^2g'^2 + y^2h'^2 + \sum_{i>0} \left(xy(Sq^i g'_i)^2 + x^2y^2(Sq^{i-1}g'_i)^2 + (Sq^i h'_i)^2 \right).$$

Comparing coefficients of xy gives $g = \sum_{i>0} Sq^i(g'_i)$. Hence g is hit. Reversing the argument shows that if g is hit then so is f. □

Definition 1.7.2. For $d \geq 0$, the **Kameko map** $\kappa : P^d(2) \to P^{2d+2}(2)$ is the linear map defined by $\kappa(g) = xyg^2$.

Thus the image of κ is the subspace of polynomials f of the form $f = xyg^2$, and the cokernel of κ is the subspace of polynomials f of the form $f = h^2$, which are hit. Hence Proposition 1.7.1 gives the following result.

Proposition 1.7.3. *For $d \geq 0$, the Kameko map induces a linear isomorphism $Q^d(2) \to Q^{2d+2}(2)$, which we also denote by κ. The inverse map is induced by the surjective linear map $\kappa' : P^{2d+2}(2) \to P^d(2)$ defined by $\kappa'(xyg^2 + h^2) = g$.* □

We now turn to the odd degree case.

Proposition 1.7.4. *For d odd, $Q^d(2) = 0$ unless $d = 2^j - 1$ for some $j \geq 1$.*

Proof. First let $d = 4k + 1$ where $k > 0$, and let $f \in P^d(2)$ be a monomial. By permuting x and y if necessary, we may assume that $f = xg^2$ for some monomial $g \in P^{2k}(2)$. If $g = g'^2$, then g is hit by Proposition 1.4.4 and so f is hit by Proposition 1.4.5. If $g = xyh'^2$, then $Sq^1(x^3yh'^4) = x^4yh'^4 + f$. Since $x^4yh'^4$ is hit by Proposition 1.4.5, it follows that f is hit.

Next let $d = 8k+3$ where $k > 0$, and let $f \in P^d(2)$ be a monomial. As before we may assume that $f = xg^2$ for some monomial $g \in P^{4k+1}(2)$. Then g is hit by the preceding case. By Proposition 1.4.5, it follows that f is hit. Continuing in this way, the result is true for integers d of the form $2^r k + (2^r - 1)$ for $r \geq 2$ by induction on r. □

The next result gives a necessary and sufficient condition for an odd degree polynomial in $P(2)$ to be hit, corresponding to Proposition 1.7.1 for even degrees. Note that the converse of Proposition 1.4.5 is false, for example $g = y^3+xy^2$ is not hit since it contains a spike, but $xg^2 = xy^6+x^3y^4 = Sq^1(xy^5+x^3y^3)+Sq^2(x^2y^3)$ is hit.

Proposition 1.7.5. *For $d \geq 2$, $f = xg^2 + yh^2 \in P^{2d-1}(2)$ is hit if and only if $g \sim yg', h \sim xg'$ for some $g' \in P^{d-2}(2)$.*

Proof. Let $f = Sq^1(f') + \sum_{i>0} Sq^{2i}(f'_i)$. Then f' has even degree and f'_i has odd degree, so $f' = xyg'^2 + h'^2$ and $f'_i = xg'^2_i + yh'^2_i$. Hence

$$f = xg^2 + yh^2 = (x^2y + xy^2)g'^2 + \sum_{i>0} \left(x(Sq^i g'_i)^2 + y(Sq^i h'_i)^2\right)$$

Comparing coefficients of x, y gives $g = yg' + \sum_{i>0} Sq^i(g'_i)$, $h = xg' + \sum_{i>0} Sq^i(h'_i)$. This proves the result in one direction. The converse follows easily by reversing the argument. □

Proposition 1.7.6. *For $d > 0$, the linear map $\delta : P^{2d-1}(2) \to P^{4d-1}(2)$ defined by $\delta(xg^2) = x^3g^4$ and $\delta(yh^2) = y^3h^4$ sends hits to hits.*

Proof. Suppose $f = xg^2 + yh^2$ is hit. Then by Proposition 1.7.5, for some $g' \in P^{d-2}(2)$ we have $g = yg' + \sum_{i>0} Sq^i(g'_i)$, $h = xg' + \sum_{i>0} Sq^i(h'_i)$. Hence $xg^2 = xy^2g'^2 + \sum_{i>0} Sq^{2i}(xg'^2_i) \sim y(xyg'^2)$ and $yh^2 = yx^2g'^2 + \sum_{i>0} Sq^{2i}(yh'^2_i) \sim x(xyg'^2)$. By Proposition 1.7.1 we see that $\delta(f)$ is hit. □

Definition 1.7.7. For $d \geq 0$, the **duplication map** $\delta : Q^{2d-1}(2) \to Q^{4d-1}(2)$ is the linear map induced by δ.

Proposition 1.7.8. *For $d \geq 0$, the Kameko map $\kappa : Q^d(2) \to Q^{2d+2}(2)$ and the duplication map $\delta : Q^{2d+1}(2) \to Q^{4d+3}(2)$ are maps of $\mathbb{F}_2GL(2)$-modules. Further, δ is a map of $\mathbb{F}_2M(2)$-modules, as also is κ when d is positive and even.*

Proof. Let the matrices S, T be as in Example 1.2.2, and let

$$M = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Then S and T generate $GL(2)$ and S, T and M generate the semigroup $M(2)$. Clearly κ and δ commute with the action of S.

We show that they also commute with the action of T. First consider κ. For $f \in P^d(2)$ we have $\kappa(f) = xyf^2$. Then $\kappa(f) \cdot T = (x+y)y(f \cdot T)^2 = xy(f \cdot T)^2 + y^2(f \cdot T)^2$. Since the second term is hit, $\kappa(f) \cdot T \sim xy(f \cdot T)^2 = \kappa(f \cdot T)$, as required. Now consider δ, and let $f = xg^2 + yh^2$. Then $f \cdot T = (x+y)(g \cdot T)^2 + y(h \cdot T)^2$. Hence $\delta(f \cdot T) = (x^3+y^3)(g \cdot T)^4 + y^3(h \cdot T)^4$, while

$\delta(f)\cdot T = (x+y)^3(g\cdot T)^4+y^3(h\cdot T)^4 = (x^3+y^3+x^2y+xy^2)(g\cdot T)^4+y^3(h\cdot T)^4$. Using the hit equation $Sq^1(xy(g\cdot T)^4) = (x^2y+xy^2)(g\cdot T)^4$, we see that $\delta(f)\cdot T \sim \delta(f\cdot T)$, as required. Hence κ and δ are $\mathbb{F}_2GL(2)$-module maps.

Next we show that δ commutes with the action of M. If $f = xg^2 + yh^2$ then $f \cdot M = x(g \cdot M)^2$. Hence $\delta(f \cdot M) = x^3(g \cdot M)^4 = x^3g^4 \cdot M$ and $\delta(f) \cdot M = (x^3g^4 + y^3h^4) \cdot M = x^3g^4 \cdot M$ as required. Now consider κ. Since $\kappa(f) = xyf^2$ is divisible by y for all f, $\kappa(f) \cdot M = 0$. If $d > 0$ is even, then $f = xyg^2 + h^2$ and $f \cdot M = (h \cdot M)^2$ is hit. □

Example 1.7.9. At the level of polynomials, neither $\kappa : P^d(2) \to P^{2d+2}(2)$ nor $\delta : P^{2d-1}(2) \to P^{4d-1}(2)$ commutes with T. Also $\kappa : Q^d(2) \to Q^{2d+2}(2)$ does not commute with M when $d = 2^j - 1$, $j \geq 0$, since we have seen that $\kappa(f) \cdot M = 0$ for all f, whereas $\kappa(x^d \cdot M) = \kappa(x^d) = x^{2d+1}y$ is a spike.

1.8 The hit problem for $P(2)$

We show in this section how the iteration of the Kameko and duplication maps can be used to link $Q^d(2)$ to $Q^0(2)$, $Q^1(2)$ or $Q^3(2)$ as $GL(2)$-modules. We begin by working out the $GL(2)$-module structures of these three cases.

We know that $Q^0(2) = P^0(2)$ is the trivial 1-dimensional representation and that $Q^1(2) = P^1(2)$ is the natural representation of $GL(2)$. For $d = 2, 3$ we reduce the calculations of Example 1.2.2 modulo hits to obtain $(xy) \cdot T = xy$ and $(xy) \cdot S = xy$ in $Q^2(2)$, so that $Q^2(2)$ is the trivial representation. In $Q^3(2)$,

$$x^3 \cdot T = x^3 + y^3, \quad y^3 \cdot T = y^3, \quad x^2y \cdot T = x^2y + y^3,$$
$$x^3 \cdot S = y^3, \quad y^3 \cdot S = x^3, \quad x^2y \cdot S = x^2y.$$

We see that $g = x^3+y^3+x^2y$, viewed as an element of $Q^3(2)$, is a $GL(2)$-invariant, and that x^3, y^3 generate a complementary 2-dimensional submodule, which is the image of $\delta : Q^1(2) \to Q^3(2)$. Thus $\delta : Q^1(2) \to Q^3(2)$ is not surjective, and so the next result fails for $j = 1$.

Proposition 1.8.1. *For $d = 2^j - 1$ and $j > 1$, the duplication map $\delta : Q^d(2) \to Q^{2d+1}(2)$ is an isomorphism.*

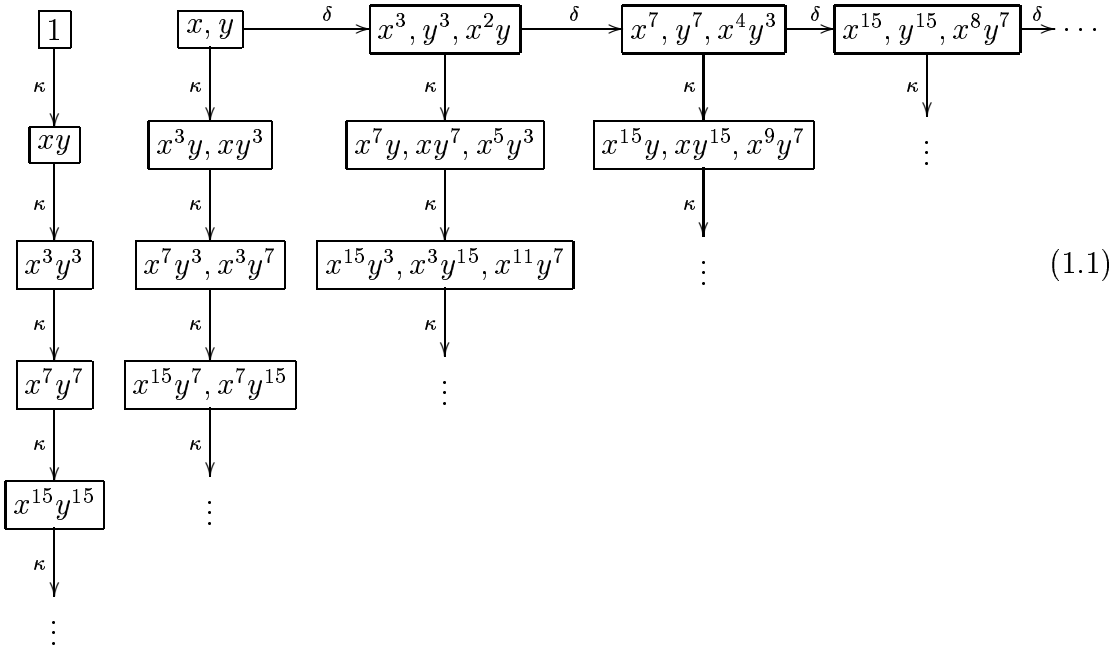
Proof. We first show that $\delta : Q^d(2) \to Q^{2d+1}(2)$ is surjective. After exchanging x and y if necessary, a monomial in $Q^{2d+1}(2)$ has the form $f = gh^8$ for some monomial h, where $g = x^7$, x^3y^4, x^6y or x^5y^2. In the first two cases, f is clearly in the image of δ. Since $Sq^2(x^3y^2h^8) = (x^5y^2 + x^3y^4)h^8$ and $Sq^1(x^5yh^8) = (x^6y + x^5y^2)h^8$, $x^3y^4h^8 \sim x^5y^2h^8 \sim x^6yh^8$. This shows that δ is surjective.

For all $j > 1$, $Q^d(2)$ contains two spikes x^{2^j-1} and y^{2^j-1} and monomials divisible by xy. By Propositions 1.6.2 and 1.4.6 it follows that $\dim Q^d(2) \geq 3$. Since $\dim Q^3(2) = 3$ by Example 1.6.5, we see that δ is an isomorphism. □

Iteration of the Kameko map κ shows that, for d' even, $Q^{d'}(2)$ is isomorphic to $Q^2(2)$ or $Q^d(2)$ for d odd. It then follows that either $Q^d(2) = 0$ for d odd by Proposition 1.7.4, or $Q^d(2)$ is isomorphic to $Q^1(2)$ or to $Q^3(2)$ by iteration of the duplication map δ. We may identify $GL(2)$ with the symmetric group $\Sigma(3)$ of permutations of a set of three elements, and then interpret $Q^3(2)$ as the regular representation of $\Sigma(3)$, permuting x^2y, $x^3 + x^2y$ and $y^3 + x^2y$. Thus we can summarize the solution of the hit problem for $P^d(2)$ in terms of $GL(2)$-modules as follows.

Theorem 1.8.2. *If $Q^d(2) \neq 0$, then it is isomorphic to the trivial module $Q^0(2)$, the natural module $Q^1(2)$ or the permutation module $Q^3(2)$ as a $\mathbb{F}_2GL(2)$-module. More precisely, $Q^d(2) \neq 0$ if and only if $d = 2^{j_1} + 2^{j_2} - 2$ for some $j_1 \geq j_2 \geq 0$, and $Q^d(2)$ is the trivial module if $j_1 = j_2$, the natural module if $j_1 = j_2 + 1$, and the permutation module if $j_1 > j_2 + 1$.* □

The following diagram illustrates the results of this section by giving a monomial basis for $Q^d(2)$ for $d \leq 30$, together with the maps κ and δ. In each case, the last listed monomial generates $Q^d(2)$ as a $\mathbb{F}_2GL(2)$-module.



(1.1)

All maps shown, except the first δ, are isomorphisms of $\mathbb{F}_2GL(2)$-modules, and all others, except the first row of κ, are also isomorphisms of $\mathbb{F}_2M(2)$-modules. The inverse Kameko map $\kappa' : Q^{2d+2}(2) \to Q^d(2)$, for $d = 2^j - 1$ and $j > 1$, is not an isomorphism of $\mathbb{F}_2M(2)$-modules, because the matrix M acts as zero in the first module but not in the second. However, $\kappa \circ \delta \circ \kappa' : Q^{2d+2}(2) \to Q^{4d+4}(2)$ is an isomorphism, because M now acts as zero in both modules. Hence, for all d, $Q^d(2)$ is isomorphic as a $\mathbb{F}_2M(2)$-module to $Q^{d'}(2)$ where $d' = 0, 1, 2, 3, 4$ or 8. These six cases are distinguished by dimension and by the action of M.

1.9 Remarks

The basic facts about Steenrod operations recorded in this chapter can be found in Steenrod-Epstein [142], frequently replicated and modified in various books and conference proceedings [53, 88, 136, 168, 169]. The total Steenrod operation Sq goes back at least to [10], and is described in [136] as a perturbation of the Frobenius map. The terminology of spikes and hit elements is due to Singer [134]. The Kameko map appears in Kameko's thesis [75] and the duplication map in [158]. The hit problem was first raised by Frank Peterson [118]. A natural interpretation of Steenrod squaring operations is explained in Appendix A in terms of differential operators. The hit problem can be posed in its dual form which is to find the intersection of the kernels of the dual operations Sq_k acting on a divided power algebra over $\mathbb{F}_2$. Following [33, 121], we consider the dual problem in a later chapter.

Chapter 2

Conjugate Steenrod squares

2.0 Introduction

In this chapter we introduce a second family of linear operations on $P(n) = \mathbb{F}_2[x_1, \ldots, x_n]$, the **conjugate Steenrod squares** $Xq^k : P^d(n) \to P^{d+k}(n)$. Although it is not immediately apparent from the definition, Xq^k can be expressed as a sum of compositions of the Steenrod squares Sq^k, and conversely. It follows from Chapter 1 that the left action of Xq^k on $P(n)$ also commutes with the right action of matrices in $M(n)$, and also that the operations Xq^k can be used in the hit problem. We apply them to prove Peterson's conjecture that every polynomial $f \in P^d(n)$ is hit, so that $Q^d(n) = 0$, if there are no spikes in $P^d(n)$.

As for Sq^k, we introduce the operations Xq^k by means of the **total conjugate Steenrod square** $Xq = \sum_{k\geq 0} Xq^k$. For this, we need to extend our algebraic framework. For $x \in P^1(n)$, $Xq(x) = x + x^2 + \cdots + x^{2^k} + \cdots$ is a formal power series, not a polynomial, and so in Section 2.1 we review some properties of algebras of formal power series. In Section 2.2 we define $Xq^k : P^d(n) \to P^{d+k}(n)$ for $k \geq 0$ and establish properties analogous to those of Sq^k. In particular, when $k = 2^j - 1$, Xq^k raises an element of $P^1(n)$ to its 2^jth power. Thus decompositions of integers as sums of integers of the form $2^j - 1$, which have already appeared in connection with spike monomials in Chapter 1, also play a part in the study of the operations Xq^k. Section 2.3 introduces the important function $\mu(d)$, the minimum number of integers of the form $2^j - 1$ which sum to d.

In Section 2.4, we show how to express Xq^k recursively as a sum of compositions of Sq^j's for $j \leq k$, using the formulae $\sum_{i+j=k} Sq^i Xq^j = 0$ and $\sum_{i+j=k} Xq^i Sq^j = 0$. In Section 2.5 we apply the operations Xq^k to the hit problem. The key result, Proposition 2.5.2, implies that for polynomials f and g the product of f and $Sq^k(g)$ is hit if and only if the product of $Xq^k(f)$ and g is hit. This is known as the **conjugation trick** or briefly as the **χ-trick**. We conclude by establishing the Peterson conjecture, Theorem 2.5.5.

2.1 Formal power series

Let R be a commutative ring with 1. A **(formal) power series** over R in one variable x has the form $f = \sum_{i \geq 0} f_i x^i$, where $f_i \in R$. If $g = \sum_{i \geq 0} g_i x^i$ is another such series, then $f = g$ if and only if $f_i = g_i$ for all $i \geq 0$. The **ith term** of f is $f_i x^i$. We identify x^0 with the identity element 1 of R, so that $f_0 x^0$ is identified with $f_0 \in R$. This is the **constant term** of f, and it is alternatively written as $f(0)$. Power series are added and subtracted term by term, and are multiplied using the Cauchy product $h = fg$ where $h_k = \sum_{i+j=k} f_i g_j$ for $k \geq 0$. These operations make formal power series into an algebra $R[[x]]$. We write $\mathbb{F}_2[[x]]$ alternatively as $\overline{P}(1)$, and identify the subalgebra of elements with a finite number of nonzero terms with the polynomial algebra $P(1)$. In contrast to $P(1)$, the algebra $\overline{P}(1)$ is not graded, since f is not obtained from its terms by algebra operations.

More generally, a formal power series over R in n variables $x_1, \ldots, x_n$ assigns a coefficient in R to each monomial in the variables. We index monomials by listing their exponents in a fixed order, so that the vector $I = (i_1, \ldots, i_n)$ is used to index the monomial $x_1^{i_1} \cdots x_n^{i_n}$. Thus the product of two monomials corresponds to the sum of their exponent vectors.

Definition 2.1.1. Let R be a commutative ring with identity. For $n \geq 1$, a **(formal) power series** over R in n variables $x_1, \ldots, x_n$ is a formal sum

$$f = \sum_I f_I x_1^{i_1} \cdots x_n^{i_n}, \text{ where } f_I \in R,$$

whose terms are indexed by the set of vectors $I = (i_1, \ldots, i_n)$ of integers ≥ 0. The **(formal) power series algebra** $R[[x_1, \ldots, x_n]]$ is the algebra over R whose elements are the power series over R, with sum $f + g$ defined by $(f+g)_I = f_I + g_I$, and product fg defined by $(fg)_I = \sum_{I=J+K} f_J g_K$. For all I, the element $x_1^{i_1} \cdots x_n^{i_n}$ is a **monomial** and has **degree** $|I| = i_1 + \cdots + i_n$. The unique monomial $x_1^0 \cdots x_n^0 = 1$ of degree 0 is the identity element of the algebra. The corresponding coefficient, denoted by f_0 or $f(0)$, is the **constant term** of f.

As in the case $n = 1$, we identify the subalgebra of $R[[x_1, \ldots, x_n]]$ generated by $x_1, \ldots, x_n$ with the polynomial algebra $R[x_1, \ldots, x_n]$. When R is a field, this is the vector space spanned by the monomials. Since $\deg(fg) = \deg(f) + \deg(g)$ for polynomials f and g, the only invertible elements of $R[x_1, \ldots, x_n]$ are the invertible elements of R, regarded as polynomials of degree 0. On the other hand, every polynomial with invertible constant term has a multiplicative inverse in the power series algebra. The geometric series $(1-x)^{-1} = \sum_{k \geq 0} x^k$ is a familiar example, and implies the following more general result.

Proposition 2.1.2. *A power series f is invertible in $R[[x_1, \ldots, x_n]]$ if and only if its constant term f_0 is invertible in R.*

Proof. If f is invertible and $g = f^{-1}$, then $f_0 g_0 = 1$, so f_0 is invertible. Conversely, if f_0 is invertible then by taking a constant multiple we may assume that $f_0 = 1$. Let $h = 1 - f$, so that $h_0 = 0$. Then for $k > 0$, h^k has no terms of degree $< k$. It follows that the expansion $f^{-1} = (1-h)^{-1} = \sum_{k \geq 0} h^k$ is valid in $R[[x_1, \ldots, x_n]]$, since the sum contains only a finite number of terms of a given degree. □

Example 2.1.3. In $\overline{P}(1)$, let $f = 1 + \sum_{k \geq 0} x^{2^k}$ and $g = \sum_{k \geq 0} x^{2^k - 1}$. Then $f + f^2 = x$ and $f = 1 + xg$. Hence $x/f = 1 + f = xg$ and so $g = f^{-1}$.

Proposition 2.1.4. *Let $f = 1 + \sum_{k \geq 0} x^{2^k}$ in $\overline{P}(1)$. Then the coefficient of x^r in f^{r+1} is zero for all $r \geq 1$.*

Proof. We write $r = 2^k(2s-1)$ where $k \geq 0$ and $s \geq 1$, and argue by induction on k. For the base case $k = 0$, $r + 1 = 2s$ and $f^{r+1} = (f^2)^s = (1 + x^2 + x^4 + \cdots)^s$ contains no odd powers of x.

Thus let $k \geq 1$, and assume that the result holds for $k - 1$. Then $f^{r+1} = (1 + x + x^2 + x^4 + \cdots)(1 + x^{2^k} + x^{2^{k+1}} + \cdots)^{2s-1}$. Since r is even and the second factor is a power series in x^2, the term x in the first factor can not contribute to the coefficient of x^r. Thus the coefficient of x^r in f^{r+1} is also the coefficient of x^r in $f^{r+2} = (1 + x^2 + x^4 + x^8 + \cdots)(1 + x^{2^k} + x^{2^{k+1}} + \cdots)^{2s-1}$. Let $y = x^2$, so that $x^r = y^{r'}$ where $r' = r/2 = 2^{k-1}(2s-1)$. By the induction hypothesis, the coefficient of $y^{r'}$ in $(1 + y + y^2 + y^4 + \cdots)^{r'+1}$ is zero. But $(1 + y + y^2 + y^4 + \cdots)^{r'+1} = (1 + y + y^2 + y^4 + \cdots)(1 + y + y^2 + y^4 + \cdots)^{r'} = (1 + y + y^2 + y^4 + \cdots)(1 + y^{2^{k-1}} + y^{2^k} + \cdots)^{2s-1} = f^{r+2}$. This completes the induction. □

Given polynomials f and g in $R[x]$, we can form the compositions $(f \circ g)(x) = f(g(x))$ and $(g \circ f)(x) = g(f(x))$ by substituting $g(x)$ for x in f or $f(x)$ for x in g respectively. In other words, we think of polynomials as functions from R to R, although we cannot literally treat them in this way when R is finite. For example, if $f = x + x^2$ and $g = 1 + x \in P(1)$, then $f \circ g = (1 + x) + (1 + x)^2 = x + x^2$ and $g \circ f = 1 + x + x^2$. Composition is associative but not commutative, the polynomial x is a two-sided identity, and $\deg(f \circ g) = \deg(f)\deg(g)$. In the above example, $f \circ f = x + x^4$ and $g \circ g = x$. Thus g is invertible with respect to composition, and is the only invertible element of $P(1)$ other then x itself.

The composition product is more important when we consider the power series algebra $\overline{P}(1)$. The composition $f \circ g$ of two power series is defined if and only if $g_0 = 0$, so that substitution of g for x in f gives an expression with a finite number of terms in each degree.

Proposition 2.1.5. *Let $f = x + x^2$ and $g = \sum_{k \geq 0} x^{2^k}$ in $\overline{P}(1)$. Then $f \circ g = x = g \circ f$, so that $g = f^{(-1)}$ is the compositional inverse of f.*

Proof. We have $f \circ g = (x + x^2 + x^4 + \cdots) + (x + x^2 + x^4 + \cdots)^2 = (x + x^2 + x^4 + \cdots) + (x^2 + x^4 + x^8 + \cdots) = x$, and $g \circ f = (x + x^2) + (x + x^2)^2 + (x + x^2)^4 + \cdots = (x + x^2) + (x^2 + x^4) + (x^4 + x^8) + \cdots = x$. □

The compositional inverse corresponds to the usual notion of inverse function. Thus the above example can be treated in a more familiar way by saying that the quadratic equation $u = x + x^2$ in $\mathbb{F}_2[[x]]$ has the solution $x = \sum_{k\geq 0} u^{2^k}$, thus establishing an isomorphism between $\mathbb{F}_2[[x]]$ and $\mathbb{F}_2[[u]]$.

Proposition 2.1.6. *Let R be a commutative ring with 1. Then the set of power series $f = \sum_{i\geq 0} f_i x^i \in R[[x]]$ such that $f_0 = 0$ and f_1 is invertible in R is a group under composition, with identity element x.*

Proof. Let $f(x) = f_1x+f_2x^2+f_3x^3+\cdots$ and $g(x) = g_1x+g_2x^2+g_3x^3+\cdots$ satisfy $f(g(x)) = x$. By substituting $g(x)$ for x in f and equating coefficients, we obtain the sequence of equations $f_1g_1 = 1$, $f_1g_2 + f_2g_1^2 = 0$, $f_1g_3 + 2f_2g_1g_2 + f_3g_1^3, \ldots$. Thus f_1 and g_1 are invertible, and by replacing f and g by f/f_1 and g/g_1 we may assume that $f_1 = g_1 = 1$. Then the jth equation expresses g_j in terms of $g_1, \ldots, g_{j-1}$ and $f_1, \ldots, f_j$. Hence, if f is given, the equations have a unique solution g, which can be calculated by recursion on j. Similarly, if g is given, there is a unique solution for f. Thus every $f \in S$ has both a right inverse g and a left inverse h, and so $g = (h \circ f) \circ g = h \circ (f \circ g) = h$ is a two-sided inverse. □

It follows from Proposition 2.1.2 that $\overline{P}(1)$ is a principal ideal domain with a single descending chain of nonzero ideals (x^r) where $r \geq 0$. Its field of fractions $\overline{P}((1))$ has elements f/g where $f, g \in \overline{P}(1)$ and $g \neq 0$. If $r \geq 0$ is minimal such that $g_r \neq 0$, then $g = x^r h$ where h is invertible. Thus $f/g = x^{-r} f\, h^{-1}$ is a **formal Laurent series** in x.

2.2 Conjugate Steenrod operations Xq^k

For $k \geq 0$, we define an operation Xq^k on $P(n)$ which raises degree by k. We follow the method used in Section 1.1 to define Sq^k.

Definition 2.2.1. The **total conjugate Steenrod square** $Xq : P(n) \to \overline{P}(n)$ is the algebra map defined by

$$Xq(1) = 1, \quad Xq(x_i) = \sum_{j=0}^{\infty} x_i^{2^j}, \text{ for } 1 \leq i \leq n.$$

The **conjugate Steenrod squares** $Xq^k : P^d(n) \to P^{d+k}(n)$, for $k \geq 0$ and $d \geq 0$, are the linear maps defined by restricting Xq to $P^d(n)$ and projecting on to $P^{d+k}(n)$, so that $Xq = \sum_{k=0}^{\infty} Xq^k$ is the formal sum of its graded parts.

Proposition 2.2.2. *For all $x \in P^1(n)$, $Xq(x) = \sum_{j=0}^{\infty} x^{2^j}$. Thus $Xq^k(x) = x^{2^j}$ if $k = 2^j - 1$ and $Xq^k(x) = 0$ otherwise. More generally, for $r \geq 0$, $Xq^k(x^{2^r}) = x^{2^j}$ if $k = 2^j - 2^r$, where $j \geq r$, and $Xq^k(x^{2^r}) = 0$ otherwise.*

Proof. Let $x = \sum_{i=1}^{n} a_i x_i$, where $a_i \in \mathbb{F}_2$. Then $Xq(x) = \sum_{i=1}^{n} a_i Xq(x_i) = \sum_{j=0}^{\infty}(\sum_{i=1}^{n} a_i x_i^{2^j}) = \sum_{j=0}^{\infty}(\sum_{i=1}^{n} a_i x_i)^{2^j} = \sum_{j=0}^{\infty} x^{2^j}$. Since Xq is multiplicative, $Xq(x^{2^r}) = (\sum_{j=0}^{\infty} x^{2^j})^{2^r} = \sum_{j=r}^{\infty} x^{2^j}$. The second statement follows by equating graded parts. □

Proposition 2.2.3. (Cartan formula) *For polynomials $f, g \in P(n)$ and $k \geq 0$,*

$$Xq^k(fg) = \sum_{i+j=k} Xq^i(f)Xq^j(g).$$

Proof. This follows from the multiplicative property $Xq(fg) = Xq(f)Xq(g)$ by equating terms of degree k. □

Proposition 2.2.4. *$Xq^0 = Sq^0$, the identity map of $P(n)$.*

Proof. The proof is the same as for Sq^0 in Proposition 1.1.5. □

As for Sq^k, in principle these rules allow the evaluation of any conjugate squaring operation on a polynomial.

Example 2.2.5. In $P(2) = \mathbb{F}_2[x, y]$ we have for $j \geq 0$

$$Xq^{2^j-1}(xy) = Xq^{2^j-1}(x)y + xXq^{2^j-1}(y) = x^{2^j}y + xy^{2^j}.$$

We shall see later that $Xq^{2^j-1}(f) = Sq^{2^{j-1}}Sq^{2^{j-2}}\cdots Sq^2Sq^1(f)$ for all f.

Proposition 2.2.6. *For $f \in P^d(n)$, $Xq^k(f) = 0$ if k is not the sum of d integers of the form $2^j - 1$.*

Proof. By linearity we may assume that f is a monomial. Let $f = y_1 \cdots y_d$, where each y_i is one of the variables x_j. Then $Xq(f) = \prod_{i=1}^{d}(y_i + y_i^2 + y_i^4 + \cdots + y_i^{2^j} + \cdots) = y_1 \cdots y_d \prod_{i=1}^{d}(1 + y_i + y_i^3 + \cdots + y_i^{2^j-1} + \cdots)$. If k is not the sum of d integers of the form $2^j - 1$, then there are no terms of degree $d + k$ in $Xq(f)$, and so $Xq^k(f) = 0$. □

2.3 The functions α, ω and μ

The binary expansion of an integer $d \geq 0$ as a sum of distinct 2-powers naturally plays a fundamental part in our work. Recall from Section 1.5 that $\mathrm{bin}(d)$ is the set of terms in the binary expansion of d. We can write this in the form

$$d = \sum_{j \geq 1} 2^{j-1}\omega_j(d),$$

where $\omega_j(d) = 1$ if $2^{j-1} \in \mathrm{bin}(d)$ and $\omega_j(d) = 0$ if $2^{j-1} \notin \mathrm{bin}(d)$.

Definition 2.3.1. For $d \geq 0$, $\omega(d) = (\omega_1(d), \omega_2(d), \ldots)$ is the vector whose entries are the digits 0 or 1 in the binary expansion of d, taken in ascending order of 2-powers, and $\alpha(d) = |\mathrm{bin}(d)| = \sum_{j\geq 1} \omega_j(d)$ is the number of 1's.

For example, $\alpha(11) = 3$, $\omega(11) = (1,1,0,1,0,\ldots)$, $\alpha(27) = 4$, $\omega(27) = (1,1,0,1,1,0,\ldots)$, $\alpha(0) = 0$, $\omega(0) = (0,0,\ldots)$. By allowing equal terms, we can write d as the sum of r 2-powers for any r such that $\alpha(d) \leq r \leq d$. We observe that $\mathrm{bin}(d)$ is the unique set of terms in such a sum with $r = \alpha(d)$.

We have seen in Chapter 1 that monomials in $P(n)$ whose exponents are all of the form $2^j - 1$, or **spikes**, play a key role in the hit problem. In particular, it is important to know whether spikes occur in $P^d(n)$, and if so, how many. For example $P^{17}(2)$ contains no spikes, but $P^{17}(3)$ contains 6 spikes, namely $x_1^7x_2^7x_3^3$, $x_1^{15}x_2x_3$ and the monomials obtained from these by permuting variables. By selecting three of the four variables, these yield 24 spikes in $P^{17}(4)$, and there are no more, since 17 is odd and so is not the sum of 4 numbers of the form $2^j - 1$. For $n > 4$, on the other hand, counting the spikes in $P^{17}(n)$ such as $x_1^7x_2^3x_3^3x_4^3x_5$ involves additional exponent vectors.

Definition 2.3.2. For $d > 0$, $\mu(d)$ is the minimum number of terms of the form $2^j - 1$, with repetitions allowed, whose sum is d. We also define $\mu(0) = 0$.

Thus $\mu(d) = 1$ if and only if $d = 2^j - 1$ for some $j \geq 1$, corresponding to the solution of the hit problem for the 1-variable case, Theorem 1.5.4, and $\mu(d) = 2$ if and only if $d = (2^{j_1} - 1) + (2^{j_2} - 1)$ where $j_1 \geq j_2 \geq 1$, corresponding to the solution of the hit problem for the 2-variable case, Theorem 1.8.2. Clearly $\mu(d) \equiv d \bmod 2$ for all d.

Example 2.3.3. Since $19 = 15 + 3 + 1$ and $\mu(19) > 1$, $\mu(19) = 3$. Since $12 = 7 + 3 + 1 + 1$ and $\mu(12) > 2$, $\mu(12) = 4$.

Thus $\mu(d)$ is analogous to $\alpha(d)$, but we consider d as a sum of integers of the form $2^j - 1$ rather than 2^j. The next result relates these two functions.

Proposition 2.3.4. *For all $d \geq 0$, $\mu(d) \leq k$ if and only if $\alpha(d + k) \leq k$. Hence $\mu(d)$ is the minimum k such that $\alpha(d + k) \leq k$.*

Proof. Suppose $\mu(d) \leq k$. We can write $d = \sum_{i=1}^{k}(2^{j_i} - 1)$ where $j_i \geq 0$. Then $d + k = \sum_{i=1}^{k} 2^{j_i}$, showing that $\alpha(d + k) \leq k$. Conversely if $\alpha(d + k) \leq k$ then $d + k = \sum_{i=1}^{m} 2^{j_i}$ for some $m \leq k$. If $m < k$ then not all j_i can be zero. Hence, by splitting a 2-power into two equal parts, we can express $d + k$ as the sum of $m + 1$ 2-powers. By iteration, we may assume $m = k$. Then $d = \sum_{i=1}^{k}(2^{j_i} - 1)$, and so $\mu(d) \leq k$. □

Proposition 2.3.5. *For all $d \geq 0$, $\mu(d) \leq k$ if and only if $\mu(2d + k) \leq k$.*

Proof. By Proposition 2.3.4, $\mu(d) \leq k$ if and only if $\alpha(d + k) \leq k$, and similarly $\mu(2d + k) \leq k$ if and only if $\alpha(2d + 2k) \leq k$. But $\alpha(2d + 2k) = \alpha(d + k)$. □

We conclude this section by calculating $\mu(d)$ in some special cases.

Proposition 2.3.6. *$\mu(2^n - k) = k$ for $1 \leq k \leq n$.*

Proof. The decomposition $2^n - k = (2^{n-k+1} - 1) + \sum_{j=1}^{k-1}(2^{n-j} - 1)$ shows that $\mu(2^n - k) \leq k$. If $\mu(2^n - k) \leq k-1$, then Proposition 2.3.4 gives the contradiction $n = \alpha(2^n - 1) \leq k - 1$. Hence $\mu(2^n - k) = k$. □

Proposition 2.3.7. *Let $d = 2^n - n - 1$ for $n \geq 1$. Then $\mu(d) = n - 1$, and $d = \sum_{j=1}^{n-1}(2^j - 1)$ is the unique expression for d as the sum of $n - 1$ integers of the form $2^j - 1$.*

Proof. $\mu(d) \geq n - 1$ by Proposition 2.3.4, since assuming $\mu(d) \leq n - 2$ gives the contradiction $n - 1 = \alpha(2^n - 3) \leq n - 2$. Thus the stated decomposition implies that $\mu(2^n - n - 1) = n - 1$. The uniqueness with $n - 1$ summands follows from that of the decomposition $2^n - 2 = \sum_{j=1}^{n-1} 2^j$. □

2.4 Evaluating $Xq^k(f)$

We begin by reducing the evaluation of a conjugate Steenrod square Xq^k on a monomial to the 1-variable case.

Proposition 2.4.1. *Let $f = x_1^{d_1} \cdots x_n^{d_n}$ be a monomial in $P(n)$. Then*

$$Xq^k(f) = \sum_{k_1 + \cdots + k_n = k} Xq^{k_1}(x_1^{d_1}) \cdots Xq^{k_n}(x_n^{d_n}).$$

Proof. As for Sq^k, this follows from the Cartan formula 2.2.3 using induction on n. □

The next result is the analogue of Proposition 1.1.7 for Xq^k. Recall that

$$\binom{a}{b} = \frac{a(a-1)\cdots(a-b+1)}{b!} \tag{2.1}$$

defines binomial coefficients for negative integers a. For example, $\binom{-1}{b} = (-1)^b$ for all $b \geq 0$. Thus for $a > 0$

$$\binom{-a}{b} = \frac{-a(-a-1)\cdots(-a-b+1)}{b!} = (-1)^b \binom{a+b-1}{b}. \tag{2.2}$$

Proposition 2.4.2. *For all $x \in P^1(n)$,*

$$Xq^k(x^d) = \binom{d+2k}{k} x^{d+k},$$

where the binomial coefficient is reduced mod 2.

Proof. By (2.2), $\binom{d+2k}{k} = \binom{-d-k-1}{k} \bmod 2$. We use the second form to prove the result. Let $u = x + x^2$. We work in $\mathbb{F}_2[[x]]$, keeping in mind the isomorphism with $\mathbb{F}_2[[u]]$ explained at the end of Proposition 2.1.5. Then $Xq(u) = x$ and $Sq(x) = u$. Hence $Xq(u^d) = x^d$ and $Xq^k(u^d) = c_{d,k}u^{d+k}$, where $c_{d,k}$ is the coefficient of u^{k+d} in x^d. Thus $c_{d,k}$ is the coefficient of u^{-1} in $u^{-d-k-1}x^d$. Now $u^{-d-k-1}x^d = (x+x^2)^{-d-k-1}x^d = x^{-k-1}(1+x)^{-d-k-1} = \sum_{j\geq 0}\binom{-d-k-1}{j}x^{j-k-1}$.

Now for $r \geq 0$, x^r is a formal power series in u, and so the coefficient of u^{-1} in x^r is zero. Thus only negative powers of x can contribute to the above sum, so that $c_{d,k}$ is the coefficient of u^{-1} in

$$\sum_{j=0}^{k}\binom{-d-k-1}{j}x^{j-k-1}. \tag{2.3}$$

The term with $j = k$ is $\binom{-d-k-1}{k}x^{-1}$, and we have $x^{-1} = (u+u^2+u^4+\cdots)^{-1} = u^{-1}(1+u+u^3+u^7+\cdots)^{-1} = u^{-1}(1+u+u^2+u^4+\cdots)$ by Example 2.1.3. Thus the coefficient of this term in (2.3) is $\binom{-d-k-1}{k}$.

To complete the proof, we show that the coefficient of u^{-1} in x^{-r} is even when $r > 1$. This follows from Proposition 2.1.4, since $x^{-r} = (u+u^2+u^4+\cdots)^{-r} = u^{-r}(1+u+u^3+u^7+\cdots)^{-r} = u^{-r}(1+u+u^2+u^4+\cdots)^r$ by Example 2.1.3, and hence the required coefficient is the coefficent of u^{r-1} in $(1+u+u^2+u^4+\ldots)^r$. □

In Appendix B we give an alternative proof of Proposition 2.4.2 which connects this result with the Catalan numbers.

The following formulae relate the two families of operations Sq^k and Xq^k.

Proposition 2.4.3. *Let $k \geq 1$ and $f \in P(n)$. Then*

$$\textbf{(i)}\ \sum_{i+j=k} Sq^i \circ Xq^j(f) = 0, \quad \textbf{(ii)}\ \sum_{i+j=k} Xq^i \circ Sq^j(f) = 0.$$

Proof. Let $\theta^k = \sum_{i+j=k} Sq^i \circ Xq^j$ and $\phi^k = \sum_{i+j=k} Xq^i \circ Sq^j$ for $k \geq 0$. By a straightforward calculation using Propositions 1.1.4 and 2.2.3, θ^k and ϕ^k are also evaluated on products of polynomials $f, g \in P(n)$ by the Cartan formula

$$\theta^k(fg) = \sum_{i+j=k}\theta^i(f)\theta^j(g), \quad \phi^k(fg) = \sum_{i+j=k}\phi^i(f)\phi^j(g).$$

Thus it suffices to show that $\theta^k(x) = 0$ and $\phi^k(x) = 0$ when $x = x_i$ is one of the variables. Since $Xq \circ Sq(x) = Xq(x+x^2) = Xq(x) + Xq(x)^2 = \sum_{j=0}^{\infty} x^{2^j} + \sum_{j=0}^{\infty} x^{2^{j+1}} = x$, $\phi^k(x) = 0$ when $k > 0$. By Propositions 1.1.3 and 2.2.2, $Sq^i(Xq^j(x)) = 0$ unless there is an integer $r \geq 0$ such that $j = 2^r - 1$ and $i = 0$ or $i = 2^r$. Thus $\theta^k(x) = 0$ if k is not of the form $2^s - 1$. But if $k = 2^s - 1$ with $s \geq 1$, then $\theta^k(x) = Sq^0Xq^{2^s-1}(x) + Sq^{2^{s-1}}Xq^{2^{s-1}-1}(x) = x^{2^s} + x^{2^s} = 0$. Hence $\theta^k(x) = 0$ for all $k \geq 1$. □

These formulae can be used, in principle at least, to calculate $Xq^k(f)$ recursively in terms of $Sq^j(f)$, $1 \leq j \leq k$. In particular, since $Sq^1Sq^1(f) = 0$ and $Sq^1Sq^2(f) = Sq^3(f)$ for all $f \in P(n)$ by Proposition 1.3.4, we obtain using **(i)**

$$\begin{aligned} Xq^1(f) &= Sq^1(f), \\ Xq^2(f) &= Sq^2(f) + Sq^1Xq^1(f) = Sq^2(f), \\ Xq^3(f) &= Sq^3(f) + Sq^2Xq^1(f) + Sq^1Xq^2(f) = Sq^2Sq^1(f). \end{aligned}$$

Continuing in this way, we obtain $Xq^4(f) = Sq^4(f) + Sq^2Sq^2(f)$ and $Xq^5(f) = Sq^4Sq^1(f) + Sq^2Sq^2Sq^1(f)$. (Using **(ii)** leads to the same formulae.) This procedure is not efficient, because the operations Sq^k satisfy further relations not given by Proposition 1.3.4, to be proved in Chapter 3. For example, $Sq^2Sq^2Sq^1(f) = 0$ for all f, so that $Xq^5(f) = Sq^4Sq^1(f)$. However, Proposition 2.4.3 gives the following result.

Proposition 2.4.4. *For $k \geq 0$, $Xq^k : P(n) \to P(n)$ is a Steenrod operation of degree k. Hence $Xq^k(f)$ is hit for all $f \in P^d(n)$ and $k > 0$.* □

By way of illustration, we solve the 1-variable hit problem again (see Theorem 1.5.4) using the operations Xq^k.

Example 2.4.5. For $d = 2^j - 1$, let $d = a + b$ where $a, b > 0$. Let 2^i be the minimum element of $\mathrm{bin}(b)$, so that $b \equiv 2^i \bmod 2^{i+1}$. Then $a \equiv 2^i - 1 \bmod 2^{i+1}$, and so $2^i \notin \mathrm{bin}(a + 2b)$. Hence $\binom{a+2b}{b} = 0 \bmod 2$, and so x^d is not hit.

Now assume that $d \neq 2^j - 1$, and choose a and b, as in the proof of 1.5.4, so that $2^i \notin \mathrm{bin}(d)$ and $2^{i+1} \in \mathrm{bin}(d)$. Let $b = 2^i$ and $a = d - 2^i$. Then $a + 2b = d + 2^i$ and so $2^i \in \mathrm{bin}(a + 2b)$. Hence $\binom{a+2b}{b} = 1 \bmod 2$, and so x^d is hit.

2.5 The Peterson conjecture

Proposition 2.5.1. *For elements $f, g \in P^d(n)$*

$$fXq^k(g) + Sq^k(f)g = \sum_{i=1}^{k} Sq^i(fXq^{k-i}(g)).$$

Proof. We write down the first term on the left of the equation and expand the terms on the right by the Cartan formula in the rows of the following array.

$$\begin{array}{lllll} fXq^k(g) & & & & \\ fSq^1Xq^{k-1}(g) & + Sq^1(f)Xq^{k-1}(g) & & & \\ fSq^2Xq^{k-2}(g) & + Sq^1(f)Sq^1Xq^{k-2}(g) & + Sq^2(f)Xq^{k-2}(g) & & \\ \vdots & \vdots & \vdots & \ddots & \\ fSq^k(g) & + Sq^1(f)Sq^{k-1}(g) & + Sq^2(f)Sq^{k-2}(g) + & \cdots + & Sq^k(f)g \end{array}$$

By Proposition 2.4.3**(i)**, the column sums of the above array are zero, except for the last term in the bottom line, which gives the first term on the left of the hit equation. □

Since all the terms on the right are of the form $Sq^i(h_i)$ where $i > 0$, the formula of Proposition 2.5.1 is a hit equation. Thus we have proved the following result, which provides a direct means of applying the conjugate squaring operations Xq^k to the hit problem without having first to evaluate them in terms of Steenrod squares.

Proposition 2.5.2. (The conjugation *or* χ**-trick)** *For $f, g \in P(n)$ we have $fSq^k(g) \sim Xq^k(f)g$. In particular, $fSq^k(g)$ is hit if and only if $Xq^k(f)g$ is hit.* □

We can combine this with Proposition 2.2.6 to obtain a useful condition for a homogeneous polynomial to be hit.

Proposition 2.5.3. *If $f = gh^2 \in P^d(n)$ where $\mu(\deg(h)) > \deg(g)$, then f is hit.*

Proof. Let $a = \deg(g)$, $b = \deg(h)$. By Proposition 1.1.9 we have $h^2 = Sq^b(h)$. By Proposition 2.2.6 $Xq^b(g) = 0$ since $\mu(b) > a$. Then the χ-trick 2.5.2 shows that $f = gh^2 = gSq^b(h) \sim Xq^b(g)h = 0$. □

Example 2.5.4. Let $n = 3$ and let $f = x^7y^2x^2$. Then $f = gh^2$ where $g = x$ has degree 1 and $h = x^3yz$ has degree 5. Since $\mu(5) = 3 > 1$, f is hit. The same argument applies to any monomial of the form $f = x^ay^2z^2$ of odd degree not of the form $2^j - 1$. By Proposition 1.6.2, a monomial of degree $2^j - 1$ cannot be hit.

Peterson conjectured that a minimal set of generators for $P(n)$ does not contain elements in $P^d(n)$ if $\alpha(d + n) > n$. By Proposition 2.3.4, this condition is equivalent to $\mu(d) > n$.

Theorem 2.5.5. (Peterson conjecture) *$Q^d(n) \neq 0$ if and only if $\mu(d) \leq n$.*

Proof. If $\mu(d) \leq n$ then we can write $d = (2^{j_1} - 1) + \cdots + (2^{j_n} - 1)$. Then $P^d(n)$ contains the spike $x_1^{2^{j_1}-1} \cdots x_n^{2^{j_n}-1}$, which is not hit (Proposition 1.6.4) and so $Q^d(n) \neq 0$.

Conversely, suppose $\mu(d) > n$. A monomial $f \in P^d(n)$ has the form $f = gh^2$ where g is the product of a distinct variables, $a \leq n$. Then $d = a + 2b$ where $b = \deg(h)$. Since $\mu(d) > n \geq a$, by Proposition 2.3.5 we have $\mu(b) > a$. By Proposition 2.5.3 we see that f is hit. Hence $Q^d(n) = 0$. □

Proposition 2.5.5 shows that $Q^d(n)$ can only be non-zero in degrees d where spikes exist. Example 2.5.4 shows that Proposition 2.5.3 is actually stronger than Proposition 2.5.5, as it can sometimes be applied to prove that a monomial of degree d is hit in cases where $\mu(d) \leq n$.

2.6 Remarks

The authoritative book by Stanley [138] is a good reference for formal power series: see in particular Volume 2, Section 5.4, where Proposition 2.1.6 appears, and Section 6.1 for formal Laurent series.

The conjugate Steenrod operations Xq^k and the relations of Proposition 2.4.3 were first studied by Thom [149] and Wu [171] in the context of duality for characteristic classes of vector bundles. Following the work of Milnor and Moore on Hopf algebras [94], they were placed in a general algebraic context.

The μ-function has been investigated by several authors [39, 134]. The proof of the Peterson conjecture was first published in [163] with the version presented here appearing in [164] alongside Peterson's application to his original problem [119]. Campbell and Selick had verified the conjecture up to $n = 5$. The χ-trick was known to Peterson and other topologists since at least the 1970s in the context of bordism theory and appears for example in [112].

Chapter 3

The Steenrod algebra $\mathcal{A}_2$

3.0 Introduction

In this chapter we define the **mod 2 Steenrod algebra** $\mathcal{A}_2$, with generators denoted symbolically by Sq^k, $k \geq 0$, and prove that the Steenrod squaring operations Sq^k defined in Chapter 1 make the polynomial algebra $P(n)$ into a left $\mathcal{A}_2$-module. The commuting actions of $\mathcal{A}_2$ on the left and of $\mathbb{F}_2 M(n)$ on the right of $P(n)$ form our central object of study. The essence of the hit problem is to find a minimal set of generators of $P(n)$ as an $\mathcal{A}_2$-module.

In Section 3.1 we define $\mathcal{A}_2$ formally by generators Sq^k subject to the **Adem relations**. We introduce **admissible monomials** $Sq^A = Sq^{a_1} \cdots Sq^{a_s}$ and show that they span $\mathcal{A}_2$ as a vector space over $\mathbb{F}_2$. In Section 3.2 we establish the $\mathcal{A}_2$-module structure of $P(n)$ by showing that the operations Sq^k on $P(n)$ defined in Chapter 1 satisfy the Adem relations. We show that Steenrod operations of degree d are faithfully represented by their action on the product of the variables $x_1 x_2 \cdots x_n$ when $n \geq d$. In Section 3.3 we prove that the admissible monomials form a vector space basis for $\mathcal{A}_2$, and in Section 3.4 we introduce a second important basis, the **Milnor basis**.

3.1 The Adem relations

Definition 3.1.1. The **Steenrod algebra** $\mathcal{A}_2$ is the associative algebra over $\mathbb{F}_2$ generated by symbols Sq^k, $k \geq 0$, subject to the relations $Sq^0 = 1$, the identity element of $\mathcal{A}_2$, and the **Adem relations**

$$Sq^a Sq^b = \sum_{j=0}^{[a/2]} \binom{b-j-1}{a-2j} Sq^{a+b-j} Sq^j, \quad \text{for } \ 0 < a < 2b, \tag{3.1}$$

where the binomial coefficient is reduced mod 2.

The algebra $\mathcal{A}_2$ is not commutative. For example, $Sq^1Sq^2 = Sq^3 \neq Sq^2Sq^1$. Nonetheless it is convenient to adapt polynomial language to refer to elements of $\mathcal{A}_2$.

Definition 3.1.2. Let $A = (a_1, a_2, \ldots, a_s)$, where $s \geq 1$ and $a_1, \ldots, a_s \geq 0$. The product $Sq^A = Sq^{a_1}Sq^{a_2}\cdots Sq^{a_s}$ is a **monomial** in $\mathcal{A}_2$ with **exponent vector** A.

Every element $\theta \in \mathcal{A}_2$ is a sum of monomials, but this expression is not unique. The example of $Sq^1Sq^2 = Sq^3$ shows that monomials with different exponent vectors can be equal as elements of $\mathcal{A}_2$. However, all terms in the Adem relation (3.1) have the same exponent sum $a + b$, and this is also unchanged by inserting or omitting factors $Sq^0 = 1$. This gives the following result.

Proposition 3.1.3. *The algebra $\mathcal{A}_2$ is graded by giving Sq^k degree k for $k \geq 0$. If $Sq^A \neq 0$ then Sq^A has degree $|A| = a_1 + \cdots + a_s$, where $A = (a_1, a_2, \ldots, a_s)$.* $\square$

Definition 3.1.4. For $k \geq 0$, we denote by $\mathcal{A}_2^k$ the vector space over $\mathbb{F}_2$ spanned by the set of monomials Sq^A of degree k. We write $\deg(\theta) = k$ if $\theta \in \mathcal{A}_2^k$. Thus $\mathcal{A}_2 = \sum_{k\geq 0} \mathcal{A}_2^k$. We also denote by $\mathcal{A}_2^+ = \sum_{k>0} \mathcal{A}_2^k$ the two-sided ideal of $\mathcal{A}_2$ generated by Sq^k, $k > 0$.

Example 3.1.5. Some examples of the Adem relations are as follows.

(1) $Sq^1Sq^1 = 0$, and in general $Sq^{2k-1}Sq^k = 0$ for $k \geq 1$.

(2) $Sq^1Sq^2 = Sq^3$ and $Sq^1Sq^3 = 0$. In general, $Sq^1Sq^{2k} = Sq^{2k+1}$, $Sq^1Sq^{2k+1} = 0$ for $k \geq 0$.

(3) $Sq^2Sq^2 = Sq^3Sq^1$ and $Sq^2Sq^3 = Sq^5 + Sq^4Sq^1$. In general, $Sq^2Sq^k = Sq^{k+1}Sq^1$ if $k \equiv 1$ or $2 \bmod 4$, $Sq^2Sq^k = Sq^{k+2} + Sq^{k+1}Sq^1$ if $k \equiv 3$ or $0 \bmod 4$.

(4) $Sq^3Sq^2 = 0$, $Sq^3Sq^3 = Sq^5Sq^1$, $Sq^3Sq^4 = Sq^7$, and $Sq^3Sq^5 = Sq^7Sq^1$. In general, $Sq^3Sq^k = Sq^{k+2}Sq^1$ if k is odd, $Sq^3Sq^k = 0$ if $k \equiv 2 \bmod 4$ and $Sq^3Sq^k = Sq^{k+3}$ if $k \equiv 0 \bmod 4$.

These relations show that the generating set Sq^k, $k \geq 1$, is far from minimal. A minimal generating set is given in Proposition 3.2.4. The Adem relations also contain much redundancy, for example relations **(4)** above follow from relations **(2)** and **(3)** using $Sq^1Sq^2 = Sq^3$.

As these examples show, the Adem relations can be used to express a composition of two squaring operations as a linear combination of compositions of the form Sq^aSq^b, where $a \geq 2b$. In fact, every term on the right hand side of (3.1) has the form Sq^cSq^d, where $c > 2d$, since the conditions $2j \leq a < 2b$ give $a + b - j > 2j$. We shall see shortly that the Adem relations can be used to reduce a longer product Sq^A to a standard form.

Definition 3.1.6. Let $A = (a_1, \ldots, a_s)$ where $s \geq 1$ and $a_1, \ldots, a_s \geq 0$. The monomial Sq^A (or the vector A) is **admissible** if $A = (a)$, $a \geq 0$, or if $s > 1$, $a_s > 0$ and $a_i \geq 2a_{i+1}$ for $1 \leq i \leq s-1$.

We shall prove in the next section that the admissible monomials form a basis for $\mathcal{A}_2$ as a vector space over $\mathbb{F}_2$. The table below shows this basis in degrees ≤ 9.

degree d	admissible basis of $\mathcal{A}_2^d$	$\dim \mathcal{A}_2^d$
0	$Sq^0 = 1$	1
1	Sq^1	1
2	Sq^2	1
3	Sq^3, Sq^2Sq^1	2
4	Sq^4, Sq^3Sq^1	2
5	Sq^5, Sq^4Sq^1	2
6	Sq^6, Sq^5Sq^1, Sq^4Sq^2	3
7	$Sq^7, Sq^6Sq^1, Sq^5Sq^2, Sq^4Sq^2Sq^1$	4
8	$Sq^8, Sq^7Sq^1, Sq^6Sq^2, Sq^5Sq^2Sq^1$	4
9	$Sq^9, Sq^8Sq^1, Sq^7Sq^2, Sq^6Sq^3, Sq^6Sq^2Sq^1$	5

We shall use two linear orderings on vectors, the left lexicographic order, or **left order** $<_l$, and the *reversed* right lexicographic order, or **right order** $<_r$. Since $Sq^0 = 1$, we can add trailing zeros to the vector A without altering Sq^A, and so we treat A as an infinite sequence with only a finite number of nonzero entries. The following definition collects our conventions for vectors.

Definition 3.1.7. A **vector** is a sequence $V = (v_1, v_2, \ldots)$ of integers ≥ 0 with only a finite number of nonzero entries. The **length** $\mathrm{len}(V)$ is the maximum j such that $v_j > 0$. If $v_j = 0$ for all j then $\mathrm{len}(V) = 0$. The **modulus** of V is the sum $|V| = \sum_{i \geq 1} v_i$, and its **degree** $\deg(V) = \sum_{i \geq 1} 2^{i-1} v_i$. Given two vectors $A = (a_1, a_2, \ldots)$ and $B = (b_1, b_2, \ldots)$, $A <_l B$ if and only if, for some k, $a_j = b_j$ for $1 \leq j < k$ and $a_k < b_k$, and $A <_r B$ if and only if, for some k, $a_j = b_j$ for $j > k$ and $a_k > b_k$.

We can use these orderings to order the admissible monomials of degree d by ordering their exponent vectors. In the table above, the two orderings coincide and the admissible basis is listed in decreasing order. The left and right orders are different when $d \geq 12$, since $(9,2,1) >_l (8,4)$ but $(9,2,1) <_r (8,4)$.

Proposition 3.1.8. *Every element of $\mathcal{A}_2$ is a sum of admissible monomials.*

Proof. By linearity is is sufficient to prove this for a monomial Sq^A. Since Sq^a is admissible for all $a \geq 0$, we may assume that $A = (a_1, a_2, \ldots, a_s)$, where $s \geq 2$ and $a_i > 0$ for $1 \leq i \leq s$, and there is some k such that $a_k < 2a_{k+1}$. Using the Adem relation with $a = a_k$ and $b = a_{k+1}$, we may write Sq^A as a sum of monomials Sq^B where B is obtained from A by replacing the pair (a, b) by the

pair $(a+b-j, j)$ for some j such that $0 \leq j \leq [a/2] < b$. Then $B >_{l,r} A$. Hence a non-admissible monomial Sq^A can be written as a sum of monomials Sq^B which are greater than Sq^A in both the left and the right orders. Iterating this procedure, Sq^A can be expressed as a sum of admissible monomials. □

The following example shows how the Adem relations can be used to convert an element of $\mathcal{A}_2$ to admissible form.

Example 3.1.9. $Sq^4Sq^2Sq^3 = Sq^4(Sq^5 + Sq^4Sq^1) = (Sq^9 + Sq^8Sq^1 + Sq^7Sq^2) + (Sq^7Sq^1 + Sq^6Sq^2)Sq^1 = Sq^9 + Sq^8Sq^1 + Sq^7Sq^2 + Sq^6Sq^2Sq^1$.

3.2 The action of $\mathcal{A}_2$ on $P(n)$

In this section we shall prove that the linear operations $Sq^k : P(n) \to P(n)$ defined in Chapter 1 define an action of the Steenrod algebra on $P(n)$. In the first place, the correspondingly named elements Sq^k generate $\mathcal{A}_2$ as an algebra. Thus given $\theta \in \mathcal{A}_2$ there is a corresponding linear operation $\theta : P(n) \to P(n)$ defined by addition and composition of the squaring operations Sq^k.

For this action to be well defined, it must be compatible with all relations between the generators Sq^k of $\mathcal{A}_2$. For example $Sq^0 = 1$ is a relation in $\mathcal{A}_2$, and Proposition 1.1.5 states that Sq^0 is the identity map of $P(n)$. Hence the action is compatible with the relation $Sq^0 = 1$. In the same way, Proposition 1.3.4 states that the action is compatible with the Adem relations $Sq^1Sq^{2k} = Sq^{2k+1}$, $Sq^1Sq^{2k+1} = 0$ for $k \geq 0$. Since the Adem relations are a set of defining relations for $\mathcal{A}_2$, our task is to extend this argument to all the Adem relations (3.1).

We argue by induction on the grading in $\mathcal{A}_2$, and we begin by extending this to all integers by defining $Sq^k = 0$ for $k < 0$, so that $\mathcal{A}_2^d = 0$ for $d < 0$. We shall prove that the operations Sq^k on $P(n)$ are compatible with the relation (3.1) for all integers a and b, and not only for $0 < a < 2b$. There is nothing to prove if $a + b < 0$, giving the base of our induction. We shall see in Section 3.3 that the 'extra' relations (3.1) obtained by omitting the condition $0 < a < 2b$ are not only properties of the linear operations $Sq^k : P(n) \to P(n)$, but are in fact relations in $\mathcal{A}_2$. Hence they are implied by the Adem relations for $0 < a < 2b$.

If $a < 0$ then (3.1) is again trivial, since the left hand side is zero and there are no terms on the right. But if $a \geq 0$ and $a \geq 2b$, the right hand side includes terms with $j \geq b$. In this case, we use equation (2.2) to interpret the mod 2 binomial coefficient. For example, when $a = 2$, $b = -1$ (3.1) states that $Sq^2Sq^{-1} = \binom{-2}{2}Sq^1Sq^0 + \binom{-3}{0}Sq^0Sq^1$, which reduces to the trivial relation $Sq^1 + Sq^1 = 0$.

Proposition 3.2.1. *For all integers a and b, let $R^{a,b} \in \mathcal{A}_2$ be the element*

$$R^{a,b} = Sq^aSq^b + \sum_{j=0}^{[a/2]} \binom{b-j-1}{a-2j} Sq^{a+b-j}Sq^j,$$

Then $R^{a,b}(f) = 0$ for all $f \in P(n)$.

Proof. As explained above, we may assume that the result is true for all a', b' with $a' + b' < a + b$. By linearity, we may assume that f is a monomial of degree $d \geq 0$. Thus we fix a and b and argue by induction on d. If $d = 0$, then $f = 1$ and so $Sq^k(f) = f$ if $k = 0$, $Sq^k(f) = 0$ otherwise. Hence $R^{a,b}(f) = 0$.

For $d > 0$, let $f = xg$ for some variable $x = x_i$. By the Cartan formula 1.1.8

$$\begin{aligned} Sq^a Sq^b(xg) &= Sq^a(xSq^b g + x^2 Sq^{b-1} g) \\ &= xSq^a Sq^b g + x^2(Sq^{a-1} Sq^b + Sq^a Sq^{b-1})g + x^4(Sq^{a-2} Sq^{b-1} g). \end{aligned}$$

Applying this expansion to each term in $R^{a,b}(g)$, we have

$$R^{a,b}(xg) = xR^{a,b}(g) + x^2 S^{a,b}(g) + x^4 T^{a,b}(g), \tag{3.2}$$

where

$$S^{a,b} = Sq^{a-1} Sq^b + Sq^a Sq^{b-1} + \sum_{j=0}^{[a/2]} \binom{b-j-1}{a-2j} (Sq^{a+b-j-1} Sq^j + Sq^{a+b-j} Sq^{j-1})$$

and

$$T^{a,b} = Sq^{a-2} Sq^{b-1} + \sum_{j=0}^{[a/2]} \binom{b-j-1}{a-2j} Sq^{a+b-j-2} Sq^{j-1}.$$

Thus $T^{a,b} = R^{a-2,b-1}$, by shifting the summation index. We claim that $S^{a,b} = R^{a-1,b} + R^{a,b-1}$. By shifting the summation index in the second term we have

$$S^{a,b} = Sq^{a-1} Sq^b + Sq^a Sq^{b-1} + \sum_{j=0}^{[a/2]} \left\{ \binom{b-j-1}{a-2j} + \binom{b-j-2}{a-2j-2} \right\} Sq^{a+b-j-1} Sq^j,$$

while $R^{a-1,b} + R^{a,b-1}$ is a similar sum where the coefficient of $Sq^{a+b-j-1} Sq^j$ is

$$\binom{b-j-1}{a-1-2j} + \binom{b-j-2}{a-2j}.$$

and writing $c = b-j-1$, $d = a-2j$ we have $\binom{c}{d} + \binom{c-1}{d-2} = \binom{c-1}{d} + \binom{c-1}{d-1} + \binom{c-1}{d-2} = \binom{c-1}{d} + \binom{c}{d-1}$, proving the claim.

Since the operations $T^{a,b} = R^{a-2,b-1}$ and $S^{a,b} = R^{a-1,b} + R^{a,b-1}$ correspond to elements of lower degree in $\mathcal{A}_2$, both of them are the zero operation on $P(n)$. In particular, $T^{a,b}(g) = S^{a,b}(g) = 0$. By (3.2), $R^{a,b}(xg) = xR^{a,b}(g) = 0$ by the induction hypothesis on d. □

Since the Adem relations are a set of defining relations for $\mathcal{A}_2$, we have proved the following theorem.

Theorem 3.2.2. *The operations Sq^k on $P(n)$ define a left $\mathcal{A}_2$-module structure on $P(n)$.* □

The hit problem of Section 1.4 asks for a minimal generating set for $P(n)$ as an $\mathcal{A}_2$-module. For example, Theorem 1.5.4 states that $P(1)$ has the minimal generating set $\{1, x, x^3, x^7, \ldots\}$. In the notation of Definition 3.1.4, a polynomial $f \in P(n)$ is hit if and only if $f \in \mathcal{A}_2^+ P(n)$. Thus the hit elements $H^d(n) = P^d(n) \cap \mathcal{A}_2^+ P(n)$.

In principle, we can evaluate the operation of any element of $\mathcal{A}_2$ on $P(n)$ using the results of Chapter 1.

Proposition 3.2.3. *Let Sq^A be a monomial in $\mathcal{A}_2$. Then for $x \in P^1(n)$*

$$Sq^A(x^{2^s}) = \begin{cases} x^{2^s}, & \text{if } A = (0), \\ x^{2^r}, & \text{if } r > s \text{ and } A = (2^{r-1}, 2^{r-2}, \ldots, 2^s), \\ 0, & \text{otherwise.} \end{cases}$$

Proof. From Proposition 1.3.3, $Sq^k(x^{2^s}) = (Sq^j(x))^{2^s}$ if $k = 2^s j$ and is zero otherwise. Applying Proposition 1.1.3, we obtain $Sq^k(x^{2^s}) = x^{2^s}$ if $k = 0$, $Sq^k(x^{2^s}) = x^{2^{s+1}}$ if $k = 2^s$, and $Sq^k(x^{2^s}) = 0$ otherwise. The result follows by induction on r. □

Proposition 3.2.4. *The elements Sq^{2^j}, $j \geq 0$, form a minimal generating set for $\mathcal{A}_2$ as an algebra over $\mathbb{F}_2$.*

Proof. Since $Sq^0 = 1$, $\mathcal{A}_2$ is generated by the elements Sq^k for $k \geq 1$. If k is not a 2-power, then $k = 2^r(2s+1)$ where $r \geq 0$ and $s \geq 1$. Let $a = 2^r$ and $b = 2^{r+1}s$ in the Adem relation (3.1). Since $2^r \in \mathrm{bin}(b-1)$, the coefficient $\binom{b-1}{a} = 1 \bmod 2$ by Proposition 1.5.3, and so the term $Sq^{a+b} = Sq^k$ appears in the Adem relation. Hence Sq^k is in the subalgebra generated by $Sq^1, Sq^2, \ldots, Sq^{k-1}$. It follows by iterating the argument that $\mathcal{A}_2$ is generated by the elements Sq^{2^j}, $j \geq 0$.

Next we show that this generating set is minimal. Using the grading on $\mathcal{A}_2$, it is sufficient to show that Sq^{2^j} is not in the subalgebra generated by Sq^i for $i < 2^j$. This follows from Proposition 3.2.3, since $Sq^i(x^{2^j}) = 0$ for $0 < i < 2^j$, while $Sq^{2^j}(x^{2^j}) = x^{2^{j+1}}$. □

3.3 The admissible basis

In this section we show that admissible monomials are linearly independent, so that by Proposition 3.1.8 they form a vector space basis for $\mathcal{A}_2$. We do this by showing that if $n \geq d$, the polynomials obtained by applying admissible monomials of degree d to the product $x_1 \cdots x_n$ are linearly independent in $P^{n+d}(n)$.

A polynomial in $P(n)$ is **symmetric** if it is invariant under all permutations of the variables $x_1, \ldots, x_n$. Standard examples are the elementary symmetric

function e_k (the sum of all k-fold products of distinct variables), the complete symmetric function h_k (the sum of all monomials of degree k), and the power sum function $p_k = x_1^k + \cdots + x_n^k$. In particular, the product $x_1 \cdots x_n$ is symmetric.

Proposition 3.3.1. *If $f \in P(n)$ is symmetric, then $\theta(f)$ is symmetric for all $\theta \in \mathcal{A}_2$.*

Proof. This follows from Proposition 1.2.3, since permutations of the variables are given by the action on $P(n)$ of the permutation matrices in $GL(n)$. □

Given a monomial $f \in P(n)$, the sum of all the distinct monomials that can be formed from f by permuting the variables $x_1, \ldots, x_n$ is a 'monomial symmetric function'. The functions e_k and p_k are special cases of these. Every symmetric polynomial can be uniquely written as a sum of monomial symmetric functions, and in particular a symmetric polynomial where all exponents are 2-powers can be uniquely written as a sum of monomial symmetric functions of the same form.

Recall from Definition 2.3.1 that for $d \geq 0$, $\omega(d) = (\omega_1(d), \omega_2(d), \ldots)$ records the sequence of digits 0 or 1 in the (reversed) binary expansion of d.

Definition 3.3.2. Given a monomial $f = x_1^{d_1} \cdots x_n^{d_n}$ in $P(n)$, its **ω-vector** $\omega(f) = \sum_{i=1}^n \omega(d_i)$.

We normally omit trailing zeros in writing ω-vectors if they do not add clarity. For example, if $f = x_1^7 x_2^4 x_3 \in P^{12}(3)$, $\omega(f) = \omega(7) + \omega(4) + \omega(1) = (1,1,1) + (0,0,1) + (1,0,0) = (2,1,2)$. It has length 3, modulus 5 and degree 12.

Definition 3.3.3. For $V = (v_1, \ldots, v_s)$ and $n \geq |V|$, the **Cartan symmetric function** $c(V) \in P(n)$ is the monomial symmetric function of degree $\deg(V)$ whose exponents are 2-powers and whose ω-vector is V. The **leading monomial** of $c(V)$ is the monomial whose exponents are in decreasing order.

For example, for $n \geq 5$, $c(2,1,2)$ is the sum of all distinct monomials obtained from the leading monomial $x_1^4 x_2^4 x_3^2 x_4 x_5$ by permuting the variables. In particular, for $n \geq 1$, $c(n) = x_1 \cdots x_n$ is the product of the variables in $P(n)$. The next result prepares for our proof of linear independence of admissible monomials $Sq^A \in \mathcal{A}_2^d$ by showing that $Sq^A(c(n)) = \sum_i c(V_i)$ is a sum of Cartan symmetric functions.

Proposition 3.3.4. *If $c(V)$ is a Cartan symmetric function and $\theta \in \mathcal{A}_2$, then $\theta(c(V))$ is a sum of Cartan symmetric functions $c(W)$ such that $|W| = |V|$ and $\deg(W) = \deg(\theta) + \deg(V)$.*

Proof. By Proposition 3.2.3 and the Cartan formula, all exponents in $\theta(c(V))$ are 2-powers. By Proposition 3.3.1, $\theta(c(V))$ is symmetric, and so it can be written uniquely as a sum of Cartan symmetric functions. If $c(W)$ appears in the sum, then $c(W)$ and $c(V)$ involve the same variables by Proposition 1.1.10, and so $|W| = |V|$. □

We next introduce the 'block' notation for monomials in $P(n)$. This is very useful for practical calculations involving the action of Steenrod operations. In Chapter 6 we develop techniques for work on the hit problem using blocks, but here we introduce it to illustrate Cartan symmetric functions. We represent a monomial by writing the ω-vectors of its exponents as the rows of an array. In writing a block, we generally omit trailing zeros in the rows, but if all entries in a row are zero then we show at least one as a marker.

Example 3.3.5. We show some monomials in $P(2)$ and their associated 2-blocks. The symbol '$-$' between two equal digits indicates that all intermediate positions are filled by that digit.

$x^{11}y^7$	x^6	$x^{2^k-1}y^{2^k}$
1 1 0 1	0 1 1	$1-1$
1 1 1	0	$0-0\ 1$

Definition 3.3.6. An **n-block** B is an array with n rows which are vectors with entries 0 or 1. The entry $b_{i,j}$ in the ith row and jth column is defined for all $j \geq 1$. The **α-vector** $\alpha(B) = (\alpha_1(B), \ldots, \alpha_n(B))$ is the row sum vector, i.e. $\alpha_i(B) = \sum_{j\geq 1} b_{i,j}$, and the **$\omega$-vector** $\omega(B) = \omega_1(B), \omega_2(B), \ldots,)$ is the column sum vector, i.e. $\omega_i(B) = \sum_{i=1}^{n} b_{i,j}$. The **degree** $\deg(B) = \deg(\omega(B))$.

We associate to a monomial $f = x_1^{d_1} \cdots x_n^{d_n}$ in $P(n)$ the n-block F whose i-th row is $\omega(x_i^{d_i})$ for $1 \leq i \leq n$. In particular, the n-block with all zero entries corresponds to the constant monomial 1. Thus $\omega(F) = \omega(f)$ and $\deg(F) = \deg(f)$. We also define $\alpha(f) = \alpha(F)$.

Definition 3.3.7. The **leading block** of a Cartan symmetric function $c(V)$ is the n-block corresponding to its leading monomial.

Thus if $V = (v_1, v_2, \ldots,)$, the leading block of $c(V)$ is an n-block with one digit 1 in each row, where $n = |V|$ and the last v_1 rows have a 1 in column 1, the preceding v_2 rows have a 1 in column 2, and so on.

Definition 3.3.8. For $\theta \in \mathcal{A}_2$, the **leading term** of $\theta(c(n)) = \sum_i c(V_i)$ is the function $c(V_i)$ such that V_i is minimal in the right order $<_r$, and the **leading block** of $\theta(c(n))$ is the leading block of its leading term.

Example 3.3.9. The diagram

$$\begin{array}{cccccccccccccccccccc}
1 & & 0 & 1 & & 1 & & 0 & 1 & & 0 & 0 & 1 & & 1 & & 0 & 1 & & 0 & 0 & 1 \\
1 & & 0 & 1 & & 1 & & 1 & & & 0 & 1 & & & 1 & & 0 & 1 & & 0 & 0 & 1 \\
1 & \overset{Sq^6}{\rightarrow} & 0 & 1 & & 1 & \overset{Sq^1}{\rightarrow} & 1 & & \overset{Sq^5}{\rightarrow} & 0 & 1 & & & 1 & \overset{Sq^2}{\rightarrow} & 1 & & \overset{Sq^4}{\rightarrow} & 1 & & \\
1 & & 0 & 1 & , & 1 & & 1 & & & 0 & 1 & & , & 1 & & 1 & & & 1 & & \\
1 & & 0 & 1 & & 1 & & 1 & & & 1 & & & & 1 & & 1 & & & 1 & & \\
1 & & 0 & 1 & & 1 & & 1 & & & 1 & & & & 1 & & 1 & & & 1 & &
\end{array}$$

shows the leading blocks of $Sq^A c(6)$ for Sq^6, Sq^5Sq^1 and Sq^4Sq^2. The corresponding leading terms are $c(0,6)$, $c(2,3,1)$ and $c(4,0,2)$. More generally, for all $n \geq 6$ the leading terms of $Sq^A c(n)$ are $c(n-6,6)$, $c(n-4,3,1)$ and $c(n-2,0,2)$ respectively, since an element of $\mathcal{A}_2^d$ can act non-trivially on at most d rows of a block. As they have different leading terms, the three polynomials $Sq^A c(n)$ are linearly independent in $P^{n+6}(n)$, and hence the admissible monomials Sq^6, Sq^5Sq^1 and Sq^4Sq^2 are linearly independent in $\mathcal{A}_2^6$.

This is justified as follows. By the Cartan formula and Proposition 1.3.2, each operation Sq^k maps a block with a single digit 1 in each row to a set of new blocks by moving some of the 1's one column to the right. Each new block corresponds to an expression for k as a sum of 2-powers. Note that the monomials represented by the blocks shown above can arise in only one way from the action of the operations Sq^k on the full polynomials. (The point can be seen by working out examples such as $Sq^1Sq^1c(2) = 0$ or $Sq^3Sq^2c(5) = 0$ using blocks.) Note also that, in the cases of the composite operations Sq^5Sq^1 and Sq^4Sq^2, all digits 1 in the leading block which are moved by the first operation are moved again by the second operation. This is possible because Sq^5Sq^1 and Sq^4Sq^2 are admissible, and clearly a block formed in this way is $<_r$-minimal.

We next associate to an admissible monomial Sq^A the vector R given by deleting the first entry of $\omega(B)$, where B is the leading block of $Sq^A c(n)$. The vector R depends only on A and not on n. In the three cases of Example 3.3.9 these vectors are (6), $(3,1)$ and $(0,2)$.

Definition 3.3.10. The **Milnor vector** of an admissible vector $A = (a_1, \ldots, a_s)$, or of Sq^A, is the vector $R = (r_1, \ldots, r_s)$, where $r_j = a_j - 2a_{j+1}$ for $1 \leq j < s$ and $r_s = a_s$.

Proposition 3.3.11. *The map $A \mapsto R$ sending an admissible vector A to its Milnor vector R is a bijection from the set of all admissible vectors to the set of all vectors, and it preserves the right order $<_r$. If $Sq^A \in \mathcal{A}_2^d$, then $|R| = 2a_1 - d \leq d$ and $d = \sum_{j=1}^s (2^j - 1)r_j$.*

Proof. Given $R = (r_1, r_2, \ldots)$ with $r_j \geq 0$ for all j, the linear equations (3.3.10) can be solved recursively for $j = s, s-1, \ldots, 1$ to give $a_j = \sum_{i=j}^s 2^{i-j} r_i$. Clearly $A = (a_1, a_2, \ldots)$ is admissible. Let $A = (a_1, \ldots, a_s)$ and $B = (b_1, \ldots, b_s)$ be admissible, with corresponding Milnor vectors $R = (r_1, \ldots, r_s)$ and $S = (s_1, \ldots, s_s)$. If $a_j = b_j$ for $j > k$ and $a_k > b_k$, then $r_j = s_j$ for $j > k$, and $r_k > s_k$. Hence $R <_r S$ if $A <_r B$. The sum of the equations (3.3.10) gives $|R| = a_1 - (a_2 + \cdots + a_s) = 2a_1 - d$. With the jth equation weighted by 2^{j-1}, the sum gives $\deg(R) = a_1$. Then $d = 2\deg(R) - |R| = \sum_{j=1}^s (2^j - 1)r_i$. Since $a_1 \leq d$, $|R| \leq d$. □

The correspondence between admissible and Milnor vectors does not preserve the left order. For example, for $A = (6,3)$ and $A' = (6,2,1)$, $A >_l A'$ but $R = (0,3)$ and $R' = (2,0,1)$, so $R <_l R'$.

Proposition 3.3.12. *Let $A = (a_1, \ldots, a_s)$ be an admissible vector with Milnor vector $R = (r_1, \ldots, r_s)$, and let $n \geq |A|$. Then $Sq^A c(n)$ has leading term $c(R^+)$, where $R^+ = (n - |R|, r_1, r_2, \ldots)$.*

Proof. We evaluate Sq^{a_i} successively for $i = s, s-1, \ldots, 1$ using the Cartan formula. For the first step, $Sq^{a_s} c(n) = c(n - r_s, r_s)$, with leading block B_1 obtained by applying Sq^1 in each of the first $r_s = a_s$ rows. Since $a_{s-1} = 2a_s + r_{s-1}$, a block B_2 in $Sq^{a_{s-1}} Sq^{a_s} c(n)$ is obtained by applying Sq^2 in the first r_s rows of B_1 and Sq^1 in the next r_{s-1} rows. Since B_2 can arise only in this way, the corresponding monomial appears in $Sq^{a_{s-1}} Sq^{a_s} c(n)$, and so $c(n - r_s - r_{s-1}, r_{s-1}, r_s)$ is a term in $Sq^{a_{s-1}} Sq^{a_s} c(n)$. It is the leading term, as all digits 1 in the second column of B_1 have been moved to the third column of B_2.

The contribution to the degree of B_2 from the first $r_s + r_{s-1}$ variables is $4r_s + 2r_{s-1} = 2a_{s-1}$. Since $a_{s-2} = 2a_{s-1} + r_{s-2}$, a block B_3 in $Sq^{a_{s-2}} Sq^{a_{s-1}} Sq^{a_s} c(n)$ is obtained by applying Sq^4 in the first r_s rows of B_2, Sq^2 in the next r_{s-1} rows of B_2, and Sq^1 in the next r_{s-2} rows. Since the block B_3 can arise only in this way, B_3 represents a monomial which appears in $Sq^{a_{s-2}} Sq^{a_{s-1}} Sq^{a_s} c(n)$. Since B_2 is the leading block of $c(n - r_s - r_{s-1} - r_{s-2}, r_{s-2}, r_{s-1}, r_s)$, this Cartan symmetric function is a term in $Sq^{a_{s-2}} Sq^{a_{s-1}} Sq^{a_s} c(n)$. Again, it is the leading term, since all digits 1 in the second and third columns of B_2 have been moved to the right. By iteration, it follows that the leading term in $Sq^A c(n)$ is $c(n - |R|, r_1, \ldots, r_s)$. □

Theorem 3.3.13. *The admissible monomials form a vector space basis for $\mathcal{A}_2$.*

Proof. Using Propositions 3.1.3 and 3.1.8, it is sufficient to prove that the admissible monomials of degree d are linearly independent for all $d \geq 0$. Let $n \geq d$ and consider the linear map $\mathcal{A}_2^d \to P^{n+d}(n)$ defined by $\theta \mapsto \theta(c(n))$. Since an admissible vector A is uniquely determined by its Milnor vector R, it follows from Proposition 3.3.12 that the symmetric polynomials obtained by applying admissible monomials Sq^A of degree d to $c(n)$ have different leading terms. Hence the polynomials $Sq^A c(n)$ are linearly independent, and so the elements Sq^A are also linearly independent. □

It follows that we can use the action of $\mathcal{A}_2$ on $P(n)$ to study the structure of $\mathcal{A}_2$ itself, since elements of $\mathcal{A}_2^d$ are faithfully represented by the action of the corresponding operations on $P(n)$ when $n \geq d$. In fact, we have seen that it suffices to consider the 'universal' case provided by the action on the product of variables $c(n) = x_1 \cdots x_n$. Thus we have proved the following result.

Proposition 3.3.14. *For $0 \leq d \leq n$, the linear map $\mathcal{A}_2^d \to P^{n+d}(n)$ defined by $\theta \mapsto \theta(c(n))$ is injective.* □

As a first example of this method, Proposition 3.2.1 has the following consequence. Recall that $Sq^k = 0$ if $k < 0$.

Proposition 3.3.15. *The Adem relation (3.1) holds in $\mathcal{A}_2$ for all integers a and b.* □

3.4 The Milnor basis

In this section, we introduce the most frequently used basis of the Steenrod algebra $\mathcal{A}_2$, the **Milnor basis**. The Milnor basis consists of elements $Sq(R)$ indexed by vectors $R = (r_1, r_2, \ldots)$ of integers ≥ 0, of which only finitely many are > 0. If A is the admissible vector associated to R by (3.3.10), $Sq(R)$ has the same degree as Sq^A. By Proposition 3.3.11, this means that $Sq(R) \in \mathcal{A}_2^d$ where $d = \sum_{j=1}^{s}(2^j-1)r_j$. Since $|R| = \sum_{j\geq 1} r_j$, $|R| \leq d$, with equality only for $R = (d)$. Recall from Definition 3.3.3 that $c(V) \in P(n)$ is the Cartan symmetric function corresponding to a vector V such that $|V| = n$.

Theorem 3.4.1. *Let $R = (r_1, r_2, \ldots)$ be a vector, let $d = \sum_{j\geq 1}(2^j - 1)r_j$, and let $n \geq d$. Then there is an unique element $Sq(R) \in \mathcal{A}_2^d$ such that*

$$Sq(R)c(n) = c(n - |R|, r_1, r_2, \ldots).$$

The elements $Sq(R)$ form a vector space basis for $\mathcal{A}_2$. This basis, the **Milnor basis***, is triangularly related to the admissible basis for the right order $<_r$.*

Proof. As above, let A be the admissible vector associated to R. Then d is the degree of $Sq^A \in \mathcal{A}_2$. By Proposition 3.3.12, $Sq^A c(n)$ is a sum of Cartan symmetric functions

$$Sq^A c(n) = c(R^+) + \sum_i c(V_i), \tag{3.3}$$

where $R^+ = (n - |R|, r_1, r_2, \ldots)$, $c(V_i)$ has degree $n + d$, and $V_i >_r R^+$ for all terms in the sum. Thus we obtain a triangular system of linear equations in $\mathcal{A}_2^d$ by listing equations (3.3) in increasing $<_r$ order for A. Solving these equations recursively, we obtain an element

$$Sq(R) = Sq^A + \sum_j Sq^{B_j} \tag{3.4}$$

in $\mathcal{A}_2^d$ such that $Sq(R)c(n) = c(R^+)$ and each B_j in the sum is admissible with $B_j >_r A$. By Proposition 3.3.14, equation (3.4) determines $Sq(R)$ uniquely.

Again by Proposition 3.3.14, since $n \geq d$ every element of $\mathcal{A}_2^d$ is determined uniquely by its action on $c(n)$. Hence equation (3.3) gives a corresponding equation

$$Sq^A = Sq(R) + \sum_i Sq(S_i) \tag{3.5}$$

in $\mathcal{A}_2^d$, where $S_i^+ = V_i$ for all i. Since $V_i >_r R^+$, $S_i > R$. By Theorem 3.3.13, the admissible monomials form a basis for $\mathcal{A}_2$. Hence the elements $Sq(R)$ also form a basis. By (3.5), the two bases are triangularly related in the order $<_r$. □

The table below shows the Milnor basis in degrees ≤ 9.

degree d	Milnor basis of $\mathcal{A}_2^d$	$\dim \mathcal{A}_2^d$
0	$Sq(0) = 1$	1
1	$Sq(1)$	1
2	$Sq(2)$	1
3	$Sq(3), Sq(0,1)$	2
4	$Sq(4), Sq(1,1)$	2
5	$Sq(5), Sq(2,1)$	2
6	$Sq(6), Sq(3,1), Sq(0,2)$	3
7	$Sq(7), Sq(4,1), Sq(1,2), Sq(0,0,1)$	4
8	$Sq(8), Sq(5,1), Sq(2,2), Sq(1,0,1)$	4
9	$Sq(9), Sq(6,1), Sq(3,2), Sq(0,3), Sq(2,0,1)$	5

For $d \geq 0$, Sq^d is the $<_r$-maximal admissible monomial, and $A = (d, 0, 0, \ldots)$ has Milnor vector $R = A$. Hence $Sq(d) = Sq^d$, so that the squaring operations Sq^d are in the Milnor basis. In principle, we can express a given admissible monomial $Sq^A \in \mathcal{A}_2^d$ in the Milnor basis by evaluating it on $c(d)$ and using (3.3) and (3.5).

Example 3.4.2. The table below gives the conversion from the admissible basis to the Milnor basis in degree 9. For example, the first line of the table corresponds to the equation $Sq^6Sq^2Sq^1c(n) = c(n-3,2,0,1) + c(n-3,0,3) + c(n-5,3,2)$.

	$Sq(2,0,1)$	$Sq(0,3)$	$Sq(3,2)$	$Sq(6,1)$	$Sq(9)$
$Sq^6Sq^2Sq^1$	1	1	1	0	0
Sq^6Sq^3	0	1	1	1	0
Sq^7Sq^2	0	0	1	0	0
Sq^8Sq^1	0	0	0	1	1
Sq^9	0	0	0	0	1

The combinatorics of keeping track of the multiplicities (mod 2) of the monomials which arise in this type of calculation soon become unmanageable. Milnor's product formula, which we discuss in the next chapter, provides a much more efficient method. The next example is a special case of this formula.

Example 3.4.3. We evaluate Sq^aSq^b on $c(a+b)$. Thus $Sq^bc(a+b) = c(a,b)$, and $Sq^ac(a,b)$ is a sum of terms of the form $c(2k, a+b-3k, k)$, where $0 \leq k \leq [a/2]$. We count the number of ways in which the leading term of $c(2k, a+b-3k, k)$ can arise. This depends on the choice of $b-k$ of $a+b-3k$ rows where the move $x_i \mapsto x_i^2$ is effected by Sq^b rather than Sq^a. Thus

$$Sq^aSq^bc(d) = \sum_{k=0}^{[a/2]} \binom{a+b-3k}{b-k} c(2k, a+b-3k, k). \tag{3.6}$$

It follows that in the Milnor basis

$$Sq(a)Sq(b) = \sum_{k=0}^{[a/2]} \binom{a+b-3k}{b-k} Sq(a+b-3k, k).$$

The Cartan formula 1.1.4 can be generalized to Milnor basis elements. We illustrate the argument using the special case $Sq(r) = Sq^r$. Writing the variables in $P(m+n)$ as $x_1, \ldots, x_m, y_1, \ldots, y_n$, we observe that

$$Sq^r(x_1 \cdots x_m y_1 \cdots y_n) = \sum_{s+t=r} Sq^s(x_1 \cdots x_m) Sq^t(y_1 \cdots y_n),$$

since a choice of r variables to be squared on the left hand side corresponds to a choice of s of the x's and t of the y's to be squared on the right hand side, where $r = s+t$. Since the squaring operations Sq^k commute with specializations of the variables (Proposition 1.2.3), the equation $Sq^r(fg) = \sum_{s+t=r} Sq^s(f)Sq^t(g)$ holds for monomials f and g, and hence for polynomials by linearity.

Proposition 3.4.4. *For all Milnor basis elements $Sq(R)$ and $f, g \in P(n)$,*

$$Sq(R)(fg) = \sum_{R=S+T} Sq(S)f \; Sq(T)g.$$

Proof. Recall that $Sq(R)c(n) = c(R^+)$, the sum of all monomials in $x_1, \ldots, x_n$ in which r_i of the variables are raised to the 2^jth power for $j \geq 1$. We rename the variables as $x_1, \ldots, x_m, y_1, \ldots, y_n$ and consider $Sq(R)c(m+n)$. Each monomial in $Sq(R)c(m+n)$ corresponds to a choice of s_j of the x's and t_j of the y's to be raised to the power 2^j, where $s_j + t_j = r_j$ for $j \geq 1$. Writing $S = (s_1, s_2, \ldots)$ and $T = (t_1, t_2, \ldots)$ we obtain

$$Sq(R)(x_1 \cdots x_m y_1 \cdots y_n) = \sum_{S+T=R} Sq(S)(x_1 \cdots x_m) Sq(T)(y_1 \cdots y_n),$$

and the proof is completed as in the special case $Sq(R) = Sq^r$. □

The Milnor basis provides a more systematic description of the elements Xq^k introduced in Chapter 2.

Proposition 3.4.5. *For $k \geq 0$, Xq^k is the sum of all Milnor basis elements $Sq(R)$ in $\mathcal{A}_2^k$.*

Proof. Since Xq is multiplicative,

$$Xq(c(n)) = Xq(x_1)Xq(x_2)\cdots Xq(x_n) = \prod_{i=1}^{n} (x_i + x_i^2 + x_i^4 + \cdots),$$

and this is the sum of all monomials in $P(n)$ with 2-powers as exponents. Thus the terms of degree k in $Xq(c(n))$ give the sum of all Cartan symmetric functions of degree $n+k$. This is $\sum_R c(R^+)$, where the sum is over all R such that $Sq(R) \in \mathcal{A}_2^k$. Taking $n \geq k$, the result follows from Proposition 3.3.14. □

Proposition 3.4.6. *The dimension of $\mathcal{A}_2^d$ as a vector space over $\mathbb{F}_2$ is the coefficient of t^d in $\prod_{j=1}^{\infty} 1/(1-t^{2^j-1})$.*

Proof. For $d > 0$, $Sq(R)$ has degree $d = \sum_{j\geq 1}(2^j-1)r_j$, where $R = (r_1, r_2, \ldots)$. Thus $\dim \mathcal{A}_2^d$ is the number of solutions $R = (r_1, r_2, \ldots)$ of the equation

$$r_1 + 3r_2 + \cdots + (2^j - 1)r_j + \cdots = d, \tag{3.7}$$

where $r_j \geq 0$ for $j \geq 1$. A solution of (3.7) gives an expression for d as the sum of $|R|$ terms, of which r_j are equal to $2^j - 1$ for $j \geq 1$. Since $1/(1-t^{2^j-1}) = \sum_{i=0}^{\infty} t^{i(2^j-1)}$, this corresponds to a term of degree d in the product of formal power series $\prod_{j=1}^{\infty} 1/(1-t^{2^j-1})$. □

3.5 Remarks

The standard reference for the development of the Steenrod algebra as an algebra of stable cohomology operations is [142], which is based on lectures by Norman Steenrod himself. The Sections 3.1 and 3.2 follow Chapter I of [142] in general, but we avoid explicit use of algebraic topology, and rely instead on the action of the squaring operations Sq^k on the polynomial algebra $P(n)$ to supply information about $\mathcal{A}_2$.

Jean-Pierre Serre [125] derived the Adem relations from the action of the Sq^k on the Cartesian product of n copies of infinite real projective space $\mathbb{R}P^\infty$. The algebra $P(n)$ can be identified with the cohomology algebra $H^*(\mathbb{R}P^\infty \times \cdots \times \mathbb{R}P^\infty; \mathbb{F}_2)$ in a natural way, so that the action has the properties introduced in Chapter 1. Serre showed that the Steenrod squares generate all stable operations in mod 2 cohomology and established the basis of admissible monomials in $\mathcal{A}_2$. For background on symmetric functions relevant to Section 3.2, see for example Chapter 1 of [87]. For the history of algebraic topology up to 1960, see [41].

The foundations for the study of the internal structure of the algebra $\mathcal{A}_2$ were laid independently in 1958 by Adams [1] and Milnor [93]. Both gave new (and closely related) additive bases for $\mathcal{A}_2$, but their approaches differed; while Adams followed the methods of Cartan and Serre, Milnor's method was based on his fundamental work with John Moore on Hopf algebras [94]. In Section 3.4 we present Milnor's basis using the methods of Adams.

Chapter 4

Products and conjugation in $\mathcal{A}_2$

4.0 Introduction

In Section 4.1 we discuss the formula for multiplying elements of the Milnor basis. The existence of this formula is a major reason for the preference given in the literature to the Milnor basis over the admissible basis. In Section 4.2 we give the **Bullett-Macdonald formula**, which expresses the Adem relations in a concise form. This provides a proof that the conjugate Steenrod squares Xq^k of Chapter 2 satisfy a set of relations dual to the Adem relations. In Section 4.3, we use this to define the **conjugation** or anti-isomorphism χ of $\mathcal{A}_2$, which exchanges the operations Sq^k and Xq^k. Finally in Section 4.4 we collect a number of formulae involving the conjugation χ.

4.1 The Milnor product formula

An important feature of the Milnor basis is that there is a combinatorial formula for multiplication of basis elements $Sq(R)$.

Definition 4.1.1. A **Milnor matrix** is an array of integers $x_{i,j} \geq 0$ of the form

$$X = \begin{array}{c|ccc} & x_{0,1} & x_{0,2} & \cdots \\ \hline x_{1,0} & x_{1,1} & x_{1,2} & \cdots \\ x_{2,0} & x_{2,1} & x_{2,2} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{array}$$

with only finitely many nonzero entries. The **row vector** $R(X) = (r_1, r_2, \ldots)$, **column vector** $S(X) = (s_1, s_2, \ldots)$ and **diagonal vector** $T(X) = (t_1, t_2, \ldots)$ of X are defined by

$$r_i = \sum_{j \geq 0} 2^j x_{i,j}, \quad s_j = \sum_{i \geq 0} x_{i,j}, \quad t_k = \sum_{i+j=k} x_{i,j}$$

and the **coefficient** $b(X)$ is the product of mod 2 multinomial coefficients

$$b(X) = \prod_{k\geq 1} \binom{t_k}{x_{k,0}\ x_{k-1,1}, \ldots, x_{0,k}}.$$

Proposition 4.1.2. *The product of Milnor basis elements is given by*

$$Sq(R)Sq(S) = \sum_X b(X)Sq(T(X)),$$

where the sum is over all Milnor matrices X with $R(X) = R$ and $S(X) = S$.

A special case of this formula was proved in Example 3.4.3. In this case, the Milnor matrix corresponding to the kth term in the product is

$$X = \begin{array}{c|ccc} & b-k & 0 & \cdots \\ \hline a-2k & k & 0 & \cdots \\ 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{array}$$

and $b(X)$ reduces to a single binomial coefficient.

To implement the product formula, we first enter R in the first column, S in the first row and zeros elsewhere to get the initial Milnor matrix

$$X = \begin{array}{c|ccc} & s_1 & s_2 & \cdots \\ \hline r_1 & 0 & 0 & \cdots \\ r_2 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{array} .$$

All further Milnor matrices can be generated systematically from the initial matrix as follows. We 'move out' each r_i along its row, by subtracting powers of 2 from r_i and inserting entries $x_{i,j}$ so as to satisfy the row constraint $R(X) = R$. The entry $x_{i,j}$ is also restricted by the column constraint $S(X) = S$, which is then used to adjust the top row of the array X. Finally, the coefficient $b(X)$ is easily determined by the following generalization of Proposition 1.5.3, simply by checking, for each diagonal of X, the powers of 2 in the binary expansions of its entries.

Proposition 4.1.3. *Let $a = a_1 + \cdots + a_s$ where $a_1, \ldots, a_s \geq 0$. Then the multinomial coefficient $\binom{a}{a_1,\ldots,a_s}$ is odd if and only if* $\mathrm{bin}(a)$ *is the disjoint union of the sets* $\mathrm{bin}(a_i)$, $1 \leq i \leq s$.

Proof. We have $(x_1 + \cdots + x_s)^a = \prod_k (x_1 + \cdots + x_s)^{2^k} = \prod_k (x_1^{2^k} + \cdots + x_s^{2^k})$ mod 2, where the product is taken over all k such that $2^k \in \mathrm{bin}(a)$. The monomial $x_1^{a_1} \cdots x_s^{a_s}$ appears in the product if and only no two integers a_i and a_j have a power of 2 in common in their binary expansions. □

Example 4.1.4. To evaluate the product $Sq(4,2)Sq(1,2)$, we have to consider the following Milnor matrices, where all entries not shown are zero.

	1	2
4	0	0
2	0	0

	0	2
2	1	0
2	0	0

	1	1
0	0	1
2	0	0

	0	2
4	0	0
0	1	0

	0	1
0	0	1
0	1	0

As only the third and fourth matrices give $b(X) = 1 \bmod 2$, $Sq(4,2)Sq(1,2) = Sq(1,3,1) + Sq(4,2,1)$. We illustrate the combinatorics involved in the proof of Proposition 4.1.2 using the second matrix above, which we denote by X.

The aim of the proof is to calculate $Sq(R)Sq(S)c(n)$, where $n \geq d$ and d is the degree of the product $Sq(R)Sq(S)$. In the example, $Sq(R)$ has degree 10 and $Sq(S)$ has degree 7, so $d = 17$. Since $Sq(1,2)c(17) = c(14,1,2)$, of degree 24, we aim to calculate $Sq(4,2)c(14,1,2)$ as a sum of Cartan symmetric functions $\sum_T c(T)$ of degree 34.

The starting point of the calculation is the equation $Sq(4,2)c(24) = c(18,4,2)$ in degree 34. The polynomial $c(14,1,2)$ can be obtained from the product $c(24)$ as the sum of specializations of the 24 variables, one for each monomial in $c(14,1,2)$, which identify variables in two sets of four and one set of two, by setting $x_i = x_j$ for suitable i, j. These specializations are given by the action of singular matrix substitutions, which commute with the action of $Sq(4,2)$ by Proposition 1.2.3. Hence $Sq(4,2)c(14,1,2)$ is the sum of the corresponding specializations of 24 variables in $c(18,4,2)$.

Let f be the leading term of $c(18,4,2)$. The Milnor matrices X which arise correspond to the different ways of choosing the variables to be identified (in two sets of four and one set of two) from the 24 variables in f, which have exponents 1 in 18 cases, 2 in 4 cases, and 4 in 2 cases. For example, the set of two variables to be identified may be a subset of the set of 4 variables with exponent 2, and the two sets of four variables to be identified may all have exponent 1 in f. This situation gives rise to the second Milnor matrix X in the list above, by moving the entry 1 in the initial Milnor matrix down to the second row and subtracting 2 from the entry 4 in the first column.

Now consider $T^+(X)$. A monomial g in $c(T^+(X))$ arising from f by specialization has 5 variables with exponent 4, two originally with exponent 4 in f, one by identification of two variables in f with exponent 2, and two by identification of four variables in f with exponent 1. The monomial g has 2 variables with exponent 2, both originally with exponent 2 in f. The remaining 10 variables in g, corresponding to the variables with exponent 1 in f not affected by the identifications, have exponent 1 in g. Thus $T^+(X) = (10,2,5)$.

Now $b(X)$ is the multiplicity of $c(T^+(X))$, or equivalently the multiplicity of the monomial g, in $Sq(4,2)c(14,1,2)$. Consider the Cartan formula 3.4.4 applied to a monomial h in $c(14,1,2)$. If g is a term in $Sq(4,2)h$, then there is a set of 5 variables which have the set of exponents $\{1,1,1,1,2\}$ in h and the set of

exponents $\{2,2,4,4,4\}$ in g, and all other variables have the same exponent in h and in g. This can only be achieved by a term in the Cartan formula for $Sq(4,2)h$ which decomposes the vector $(4,2)$ as the sum $(2,0)+(1,0)+(1,0)+(0,1)+(0,1)$. In other words, some variable y_1 which appears as y_1^2 in h appears as y_1^4 in g, while variables y_2, y_3 which appear as y_2, y_3 in h appear as y_2^2, y_3^2 in g, and variables y_4, y_5 which appear as y_4, y_5 in h appear as y_4^4, y_5^4 in g.

Thus there are $\binom{5}{2,1,2}$ choices for the monomial h, according to how the 5 variables with exponent 4 in g are split as variables with exponents 1, 2 or 4 in h. The two variables with exponent 2 in g can only arise by squaring two variables in h, so the corresponding multinomial coefficient reduces to $\binom{2}{2}$. Finally $b(X)$ is the product of these multinomial coefficients, taken over powers of 2 as exponents in g.

Proof of Proposition 4.1.2. For $R = (r_1, r_2, \ldots)$ and $S = (s_1, s_2, \ldots)$, let $d(R) = \sum_{j\geq 1}(2^j - 1)r_j$ and $d(S) = \sum_{j\geq 1}(2^j - 1)s_j$ be the degrees of $Sq(R)$ and $Sq(S)$. By Proposition 3.3.14, the product $Sq(R)Sq(S) \in \mathcal{A}_2^d$ is uniquely determined by $Sq(R)Sq(S)c(d)$, where $d = d(R) + d(S)$.

Since $d \geq d(S)$, $Sq(S)c(d) = c(S^+)$ by Theorem 3.4.1, where $S^+ = (d - |S|, s_1, s_2, \ldots)$. Thus we wish to prove that the Cartan symmetric functions which appear in $Sq(R)c(S^+)$ are of the form $c(T^+)$ where $T^+ = (d - |T|, t_1, t_2, \ldots)$ and $T = (t_1, t_2, \ldots) = T(X)$ for some Milnor array X satisfying $R(X) = R$ and $S(X) = S$, and that $c(T^+)$ appears with coefficient $b(X)$.

We can compute $Sq(R)c(S^+)$ by specialization of variables from the case $Sq(R)c(n) = c(R^+)$, where $R^+ = (n - |R|, r_1, r_2, \ldots)$ and $n \geq d(R)$. As in Section 3.3, we work with the leading blocks of Cartan symmetric functions. Thus $Sq(R)$ acts on the block for $c(n)$ by moving r_1 digits 1 from column 1 to column 2, r_2 digits 1 from column 1 to column 3, and so on. After specialization of the r_1 variables corresponding to digits moved from column 1 to column 2, we can use $Sq(R)$ to move (for example) k digits 1 from column 2 to column 3 and $r_1 - 2k$ digits 1 from column 1 to column 2, where $k < r_1/2$. More generally, specialization of the same r_1 variables allows us to use $Sq(R)$ to move $x_{1,j}$ digits 1 from column $j+1$ to column $j+2$, where $r_1 = x_{1,0} + 2x_{1,1} + 4x_{1,2} + \cdots$.

In the same way, specialization of the r_2 variables corresponding to digits moved from column 1 to column 3 allows us to move k digits 1 from column 2 to column 4 and $r_2 - 2k$ digits 1 from column 1 to column 3, where $k < r_1/2$. More generally, specialization of the same r_2 variables allows us to use $Sq(R)$ to move $x_{2,j}$ digits 1 from column $j+1$ to column $j+3$, where $r_2 = x_{2,0}+2x_{2,1}+4x_{2,2}+\cdots$. Thus for each $i \geq 1$ we obtain an equation $r_i = \sum_{j\geq 0} 2^j x_{i,j}$ by specializing the action of $Sq(R)$ on the r_i variables corresponding to digits in the block for $c(d(R))$ which are moved from column 1 to column $i+1$. Here $x_{i,j}$ is the number of digits moved from column $j+1$ to column $i+j+1$.

The total number of digits 1 ending up in column $k+1$ is therefore $t_k = \sum_{i+j=k} x_{i,j}$, where $x_{k,0} = r_k$ and $x_{0,k} = s_k$. The column constraint $s_j = \sum_j x_{i,j}$

keeps track of the number of digits available in column $j+1$ of the block for $c(d-|S|, s_1, s_2, \ldots)$ for possible moves to later columns by $Sq(R)$. The coefficient $b(X)$ keeps track of the possible choices for the ways in which these moves to later columns take place. The same Cartan symmetric function $c(d-|T|, t_1, t_2, \ldots)$ arises regardless of how, for each k, the t_k digits 1 which ended up in column $k+1$ are made up of $x_{k-j,j}$ digits 1 chosen from columns $j+1$, $0 \leq j \leq k$. Thus we obtain $b(X)$ as a product of multinomial coefficients corresponding to the diagonals in the Milnor matrix X. □

Example 4.1.5. Let R be the vector $(1, 1, \ldots, 1)$ of length k. We show by induction on k that $Sq(R) = Sq^{2^{k-1}} Sq^{2^{k-1}-1} \cdots Sq^3 Sq^1$. For the inductive step, the initial Milnor matrix is

$$X = \begin{array}{c|cccc} & 1 & 1 & \cdots & 1 \\ \hline 2^k-1 & 0 & 0 & \cdots & 0 \end{array} .$$

Inspection of the first diagonal shows that $b(X) = 0 \bmod 2$. Since the entry $y_{1,0}$ is odd for all Milnor matrices Y which arise, $b(Y) = 0 \bmod 2$ unless $y_{0,1} = 0$ and $y_{1,1} = 1$. Then by inspecting the second diagonal we see that $b(Y) = 0 \bmod 2$ unless $y_{0,2} = 0$ and $y_{1,2} = 1$. Continuing in this way, we see that the only Milnor matrix Y such that $b(Y) = 1 \bmod 2$ is

$$Y = \begin{array}{c|cccc} & 0 & 0 & \cdots & 0 \\ \hline 1 & 1 & 1 & \cdots & 1 \end{array} .$$

Since $T(Y) = (1, 1, \ldots, 1)$, this completes the induction.

4.2 The Bullett-Macdonald identity

The Bullett-Macdonald identity expresses the Adem relations in a compact form, in much the same way as the multiplicative property of the total squaring operation $Sq : P(n) \to P(n)$ expresses the Cartan formula 1.1.4 in a compact form. In order to keep track of the grading in $\mathcal{A}_2$, we introduce a formal variable t, and work with formal power series in t with coefficients in $P(n)$.

Definition 4.2.1. For $n \geq 1$, let $R = P(n)$ and let $u \in \mathbb{F}_2[[t]]$ be a formal power series in t. The **generalized total squaring operation** $Sq(u) : R \to R[[t]]$ is the linear operation

$$Sq(u) = \sum_{k=0}^{\infty} u^k Sq^k.$$

In particular, when $u = 1$ is the identity element of $\mathbb{F}_2[[t]]$, $Sq(u) = Sq$. For a polynomial $f \in R$, $Sq(t)f = \sum_{k=0}^{\infty} t^k Sq^k(f)$ is a polynomial in $R[t]$ since $Sq^k(f) = 0$ when $k > \deg(f)$. More generally, the fact that $Sq^k(f)$ is nonzero for only finitely many values of k ensures that the coefficient of t^k in $Sq(u)f$ is a polynomial in R, whether or not the constant term of u is 0.

Proposition 4.2.2. *Let $R = P(n)$. For all $u \in \mathbb{F}_2[[t]]$, $Sq(u) : R \to R[[t]]$ is multiplicative, i.e. $Sq(u)(fg) = (Sq(u)f)(Sq(u)g)$ for all $f, g \in R$.*

Proof. We have $Sq(u)(fg) = \sum_{k=0}^{\infty} u^k Sq^k(fg)$, and

$$(Sq(u)f)(Sq(u)g) = \sum_{i=0}^{\infty} u^i Sq^i(f) \sum_{j=0}^{\infty} u^j Sq^j(g) = \sum_{k=0}^{\infty} u^k \left(\sum_{i+j=k} Sq^i(f) Sq^j(g) \right).$$

The result follows from the Cartan formula 1.1.4. □

The Bullett-Macdonald identity is a relation between the operations $Sq(u)$. We compose these by extending them to operations $Sq(u) : R[[t]] \to R[[t]]$, with $Sq^k(t) = t$ for all $k \geq 0$. Thus $Sq(u)(\sum_{i\geq 0} f_i t^i) = \sum_{i\geq 0} (Sq(u)f_i)t^i$.

Proposition 4.2.3. (Bullett-Macdonald identity) *Let $n \geq 1$ and let $R = P(n)$. Then the operations $Sq(u)$ on $R[[t]]$ satisfy the relation*

$$Sq(t + t^2)Sq(1) = Sq(1 + t)Sq(t^2).$$

Proof. All four operations in the relation are algebra maps of $R[[t]]$ by Proposition 4.2.2. Since they all act as the identity map of $\mathbb{F}_2[[t]]$, it suffices to check that they agree on an element $x \in P^1(n)$. The left hand side gives $Sq(t+t^2)Sq(1)x = Sq(t+t^2)(x+x^2) = (x+x^2)+(t+t^2)Sq^1(x+x^2)+(t+t^2)^2Sq^2(x+x^2) = (x+x^2)+(t+t^2)x^2+(t^2+t^4)x^4$. Since $Sq(t^2)x = x+t^2Sq^1(x) = x+t^2x^2$, the right hand side is $Sq(1+t)(x+t^2x^2) = (x+t^2x^2)+(1+t)Sq^1(x+t^2x^2)+(1+t)^2Sq^2(x+t^2x^2) = (x+t^2x^2)+(1+t)x^2+(1+t^2)t^2x^4$. Hence $Sq(1+t)Sq(t^2)x = Sq(t+t^2)Sq(1)x$. □

The Adem relations follow from the Bullett-Macdonald identity by equating coefficients of powers of t and using the grading in $\mathcal{A}_2$. By Proposition 3.3.14, the resulting identities of operations on $R = P(n)$ for all n give corresponding identities in $\mathcal{A}_2$. For example, by equating the coefficients of t, we obtain $Sq^1Sq = \sum_{j\geq 0} Sq^{2j+1}$, which is Proposition 1.3.4. Unfortunately the general case is not so straightforward, as we have to take linear combinations of the equations derived in this way to obtain the desired result. Recall that by Proposition 3.3.15, the Adem relation (3.1) holds in $\mathcal{A}_2$ without restriction on a and b.

Proposition 4.2.4. *The identity $Sq(t+t^2)Sq(1) = Sq(1+t)Sq(t^2)$ is equivalent to the set of Adem relations (3.1) with $a, b \geq 0$.*

Proof. The method is the same as for Proposition 2.4.2. We change the variable in the coefficient ring $\mathbb{F}_2[[t]]$ by setting $u = t+t^2$. Then $t = \sum_{i=0}^{\infty} u^{2^i}$ by Proposition 2.1.5. The Bullett-Macdonald identity states that for $a \geq 0$, $\sum_{b\geq 0} Sq^aSq^b$ is the coefficient of u^a in $\sum_{c,d\geq 0}(1+t)^c t^{2d} Sq^c Sq^d$, and so, taking terms in degree $a+b$, Sq^aSq^b is the coefficient of u^a in

$$\sum_{j=0}^{a+b} (1+t)^{a+b-j} t^{2j} Sq^{a+b-j} Sq^j,$$

when this sum is expressed in powers of u. Since $u = t + t^2$, $u^{a+1} = (1+t)^{a+1}t^{a+1}$ and so Sq^aSq^b is the coefficient of u^{-1} in

$$\sum_{j=0}^{a+b}(1+t)^{b-j-1}t^{2j-a-1}Sq^{a+b-j}Sq^j.$$

Thus we wish to show that for fixed a, b and j with $a \geq 0$ and $0 \leq j \leq a+b$, the coefficient of u^{-1} in $f(t) = (1+t)^{b-j-1}t^{2j-a-1}$ is the coefficient $\binom{b-j-1}{a-2j}$ which appears in (3.1). If $2j - a - 1 \geq 0$, then $f(t) = \sum_{i\geq 0}\binom{b-j-1}{i}t^{i+2j-a-1}$ is a formal power series in t, and writing $t = u + u^2 + u^4 + u^8 + \cdots$ this is a formal power series in u, so that the coefficient of u^{-1} is zero. Hence we may assume that $j \leq a/2$, and retain only the negative powers of t in the expansion, namely

$$\sum_{i=0}^{a-2j}\binom{b-j-1}{i}t^{i+2j-a-1}. \tag{4.1}$$

The coefficient of t^{-1} here is $\binom{b-j-1}{a-2j}$, and the coefficient of u^{-1} in $t^{-1} = (u+u^2+u^4+u^8+\cdots)^{-1} = u^{-1}(1+u+u^3+u^7+\cdots)$ is 1 mod 2. As in the proof of Proposition 2.4.2, it follows from Proposition 2.1.4 that for $k > 1$ the coefficient of u^{-1} in t^{-k} is 0 mod 2. We conclude that the coefficient of u^{-1} in (4.1) is $\binom{b-j-1}{a-2j}$ mod 2. □

4.3 The conjugation χ

In this section we introduce the anti-automorphism, or conjugation, of $\mathcal{A}_2$. We begin by reworking some of the material in Section 2.4, and Proposition 2.4.3 in particular. Recall that $\overline{P}(n)$ is the formal power series algebra $\mathbb{F}_2[[x_1, \ldots, x_n]]$.

Proposition 4.3.1. *The maps $Sq : P(n) \to P(n)$ and $Xq : P(n) \to \overline{P}(n)$ can be extended to an inverse pair of algebra automorphisms of $\overline{P}(n)$.*

Proof. An element $f \in \overline{P}(n)$ has the form $f = \sum_I f_I x_1^{i_1}\cdots x_n^{i_n}$, with terms indexed by vectors $I = (i_1, \ldots, i_n)$ of integers ≥ 0 and with coefficients f_I in $\mathbb{F}_2$. We define $Sq(f)$ and $Xq(f)$ in the natural way by $Sq(f) = \sum_I f_I Sq(x_1^{i_1}\cdots x_n^{i_n})$ and $Xq(f) = \sum_I f_I Xq(x_1^{i_1}\cdots x_n^{i_n})$. To see that these are well defined elements of $\overline{P}(n)$, we observe that for a given monomial h in $P(n)$ there are only finitely many monomials of lower degree, and so there are only finitely many monomials g for which h appears in $Sq(g)$ or $Xq(g)$. Clearly Sq and Xq are algebra maps of $\overline{P}(n)$.

Since Sq and Xq are multiplicative, the compositions $Sq\circ Xq$ and $Xq\circ Sq$ must be the identity map of $\overline{P}(n)$ if they fix the generators. Let x_i, where $1 \leq i \leq n$. Then $Sq \circ Xq(x) = Sq(\sum_{j=0}^{\infty} x^{2^j}) = \sum_{j=0}^{\infty} Sq(x^{2^j}) = \sum_{j=0}^{\infty}(x^{2^j} + x^{2^{j+1}}) = x$, so $Sq \circ Xq = 1$. Also $Xq \circ Sq(x) = Xq(x + x^2) = Xq(x) + Xq(x)^2 = \sum_{j=0}^{\infty} x^{2^j} + \sum_{j=0}^{\infty} x^{2^{j+1}} = x$, so $Xq \circ Sq = 1$. □

We recover Proposition 2.4.3 by equating the components of $Xq \circ Sq$ and $Sq \circ Xq$ in degree k. By Proposition 3.3.15, these formulae yield the following identities in $\mathcal{A}_2$.

Proposition 4.3.2. *For $k > 0$, $\sum_{i+j=k} Sq^i Xq^j = 0$ and $\sum_{i+j=k} Xq^i Sq^j = 0$.* □

Example 4.3.3. As in Section 2.4, we can use these formulae to calculate Xq^k recursively in terms of Sq^j, $1 \leq j \leq k$. for example, $Xq^1 = Sq^1$, $Xq^2 = Sq^2$, $Xq^3 = Sq^2Sq^1$, $Xq^4 = Sq^4 + Sq^3Sq^1$ and $Sq^5 = Sq^4Sq^1$, where we have used the Adem relations $Sq^2Sq^2 = Sq^3Sq^1$ and $Sq^1Sq^1 = 0$ to express the results in the admissible basis.

Our next aim is to show that the elements $Xq^k \in \mathcal{A}_2$ satisfy a set of conjugate Adem relations obtained by reversing the product. For example, the relations $Sq^1Sq^2 = Sq^3$, $Sq^2Sq^2 = Sq^3Sq^1$ and $Sq^2Sq^3 = Sq^5 + Sq^4Sq^1$ have conjugate forms $Xq^2Xq^1 = Xq^3$, $Xq^2Xq^2 = Xq^1Xq^3$ and $Xq^3Xq^2 = Xq^5 + Xq^1Xq^4$, which are easily verified using Example 4.3.3. We begin by generalizing Xq as we did for Sq in Section 4.2.

Definition 4.3.4. Let $R = \overline{P}(n)$ and let $u \in \mathbb{F}_2[[t]]$. The **generalized total conjugate squaring operation** $Xq(u) : R \to R[[t]]$ is the linear operation

$$Xq(u) = \sum_{k=0}^{\infty} u^k Xq^k.$$

In particular, $Xq(u) = Xq$ when $u = 1$. The operation $Xq(u)$ is multiplicative, by the same argument as for Proposition 4.2.2. As for $Sq(u)$, we extend $Xq(u)$ to an operation $Xq(u) : R[[t]] \to R[[t]]$ by defining $Xq^k(t) = t$ for all $k \geq 0$. A formal calculation using Proposition 4.3.1 shows that $Sq(u)$ and $Xq(u)$ are inverse operations on $R[[t]]$ for all $u \in \mathbb{F}_2[[t]]$.

Proposition 4.3.5. (Conjugate Bullett-Macdonald identity) *Let $n \geq 1$ and let $R = \overline{P}(n)$. Then the operations $Xq(u)$ on $R[[t]]$ satisfy the relation*

$$Xq(t^2)Xq(1+t) = Xq(1)Xq(t+t^2).$$

Proof. Since $Xq(u) = Sq(u)^{(-1)}$ is the inverse of $Sq(u)$ with respect to composition of operations on $R[[t]]$, this follows from Proposition 4.2.3. □

The argument of Proposition 4.2.4 shows that this identity is equivalent to a set of conjugate Adem relations.

Proposition 4.3.6. *For all integers a and b, the* **conjugate Adem relations**

$$Xq^b Xq^a = \sum_{j=0}^{[a/2]} \binom{b-j-1}{a-2j} Xq^j Xq^{a+b-j}$$

hold in $\mathcal{A}_2$, where $Xq^k = 0$ if $k < 0$. □

Definition 4.3.7. The **conjugation** $\chi : \mathcal{A}_2 \to \mathcal{A}_2$ is the linear map defined on monomials Sq^A by

$$\chi(Sq^{a_1}Sq^{a_2}\cdots Sq^{a_s}) = Xq^{a_s}\cdots Xq^{a_2}Xq^{a_1}. \tag{4.2}$$

In particular, $\chi(Sq^k) = Xq^k$ for $k \geq 0$.

The map χ is well defined, since by Proposition 4.3.6 it is consistent with the Adem relations (3.1), and these, together with the relation $Sq^0 = 1$, are a set of defining relations for $\mathcal{A}_2$.

Proposition 4.3.8. *The conjugation $\chi : \mathcal{A}_2 \to \mathcal{A}_2$ has the following properties.*

(i) *For all $\theta_1, \theta_2 \in \mathcal{A}_2$, $\chi(\theta_1\theta_2) = \chi(\theta_2)\chi(\theta_1)$,*

(ii) *$\chi(Xq^k) = Sq^k$ for $k \geq 0$,*

(iii) *χ^2 is the identity map of $\mathcal{A}_2$.*

Proof. By definition, **(i)** holds when $\theta_1 = Sq^A$ and $\theta_2 = Sq^B$ are monomials in $\mathcal{A}_2$. By linearity it holds for all θ_1 and θ_2. To prove **(ii)**, we apply χ to the relations $\sum_{i+j=k} Sq^i Xq^j = 0$ and $\sum_{i+j=k} Xq^i Sq^j = 0$ of Proposition 4.3.2. This gives $\sum_{i+j=k} \chi(Xq^i)Xq^j = 0$ and $\sum_{i+j=k} Xq^i\chi(Xq^j) = 0$. It follows by induction on k that $\chi(Xq^k) = Sq^k$ for all k. Finally **(ii)** implies **(iii)**, since $\chi(Sq^k) = Xq^k$ and χ^2 is an algebra map of $\mathcal{A}_2$. □

The **opposite algebra** A^{op} of an algebra A has the same elements and the same addition and scalar multiple operations as A, but the multiplication is reversed, i.e. $a * b = ba$. Thus χ is an isomorphism from $\mathcal{A}_2$ to $\mathcal{A}_2^{\mathrm{op}}$.

Example 4.3.9. By the results above, $\mathcal{A}_2$ can be defined using generators Xq^k for $k \geq 0$ subject to $Xq^0 = 1$ and the conjugate Adem relations. We tabulate χ in degree 9 using the admissible basis.

	$Sq^6Sq^2Sq^1$	Sq^6Sq^3	Sq^7Sq^2	Sq^8Sq^1	Sq^9
$Xq^6Xq^2Xq^1$	0	0	1	0	0
Xq^6Xq^3	0	1	0	0	0
Xq^7Xq^2	1	0	0	0	0
Xq^8Xq^1	0	0	1	0	1
Xq^9	1	0	0	1	0

4.4 Conjugation and the Milnor basis

In this section we collect some formulae relating the conjugation χ to the Milnor basis. The Milnor basis element $Sq(R)$ has degree $d(R) = \sum_{j\geq 1}(2^j - 1)r_j$. Note that $d(R) = \deg(Sq(R))$ is not the same as the degree $\deg(R) = \sum_{j\geq 1} 2^{j-1}r_j$ of the *vector* $R = (r_1, r_2, \ldots)$. In fact $|R| + d(R) = 2\deg(R)$, where $|R| = \sum_{j\geq 1} r_j$.

Proposition 4.4.1. *Let $a, b \geq 0$. Then*

$$\textbf{(i)}\ Sq^a Xq^b = \sum_R \binom{\deg(R)}{a} Sq(R), \quad \textbf{(ii)}\ Xq^a Sq^b = \sum_R \binom{|R|}{b} Sq(R).$$

where the sums are over all Milnor basis elements $Sq(R)$ of degree $d(R) = a + b$.

Proof. **(i)** Let $Sq^a Xq^b = \sum_R m_R Sq^R$, where $m_R \in \mathbb{F}_2$. By Proposition 3.4.5, Xq^b is the sum of all Milnor basis elements of degree b. Since $Sq^a = Sq(a)$, it follows from Proposition 4.1.2 that m_R is the sum of the coefficients $b(X)$ for Milnor matrices

$$X = \begin{array}{c|ccc} & r_1 - a_1 & r_2 - a_2 & \cdots \\ \hline a_1 & a_2 & a_3 & \cdots \end{array} .$$

where $a = \sum_{j \geq 1} 2^{j-1} a_j$. Every vector $A = (a_1, a_2, \ldots)$ of degree a such that $a_j \leq r_j$ for $j \geq 1$ gives rise to one such matrix X, and $b(X) = \prod_{j \geq 1} \binom{r_j}{a_j}$.

By Proposition 1.5.3, $\binom{r_j}{a_j} = \binom{2^j r_j}{2^j a_j} \bmod 2$, and $\binom{2^j r_j}{b_j} = 0 \bmod 2$ if b_j is not divisible by 2^j. Hence

$$m_R = \sum_A \prod_{j \geq 1} \binom{r_j}{a_j} = \sum_A \prod_{j \geq 1} \binom{2^j r_j}{2^j a_j} = \sum_B \prod_{j \geq 1} \binom{2^j r_j}{b_j}, \tag{4.3}$$

where the sum is over all $B = (b_1, b_2, \ldots)$ such that $\sum_{j \geq 1} b_j = \sum_{j \geq 1} 2^j a_j = 2a$. Since $\deg(R) = \sum_{j \geq 1} 2^{j-1} r_j$, we have the identity

$$\prod_{j \geq 1} (1 + x)^{2^j r_j} = (1 + x)^{2 \deg(R)}.$$

Comparing coefficients of x^{2a}, (4.3) reduces to $m_R = \binom{2 \deg(R)}{2a} = \binom{\deg(R)}{a}$.

(ii) By Proposition 3.4.5, $Xq^a Sq^b = \sum_S Sq(S) Sq(b)$ where the sum is over all Milnor basis elements $Sq(S)$ of degree a. A term $b(X) Sq(R)$ in the product $Sq(S) Sq(b)$ arises from each Milnor matrix

$$Y = \begin{array}{c|c} & b_1 \\ \hline r_1 - b_1 & b_2 \\ r_2 - b_2 & b_3 \\ \vdots & \vdots \end{array} .$$

Every vector $B = (b_1, b_2, \ldots)$ such that $|B| = b = \sum_{j \geq 1} b_j$ and $b_j \leq r_j$ for $j \geq 1$ gives rise to one such Y, with coefficient $b(Y) = \prod_{j \geq 1} \binom{r_j}{b_j}$. Comparing coefficients of x^b in the identity $\prod_{j \geq 1} (1+x)^{r_j} = (1+x)^{|R|}$, we have $\sum_B \prod_{j \geq 1} \binom{r_j}{b_j} = \binom{|R|}{b}$. $\square$

We conclude this section by using Proposition 4.4.1**(i)** to obtain some formulae for $\chi(Sq^k) = Xq^k$. Recall from Definition 2.3.2 that $\mu(d)$ is the minimum number of terms in an expression for $d = d(R)$ as a sum of integers of the form $2^j - 1$, and so $\mu(d) \leq |R| \leq \deg(R)$.

Proposition 4.4.2. **(i)** *$Xq^{2^n-k} = Sq^{2^{n-1}} Xq^{2^{n-1}-k}$ for $1 \leq k \leq n$, and*

(ii) $Xq^{2^n-n-1} = Sq^{2^{n-1}} Xq^{2^{n-1}-n-1} + Sq^{2^{n-1}-1} Sq^{2^{n-2}-1} \cdots Sq^3 Sq^1$.

Proof. **(i)** By Proposition 4.4.1**(i)** and Proposition 3.4.5, **(i)** is equivalent to the statement that $\binom{\deg(R)}{a} = 1 \bmod 2$ when $a = 2^{n-1}$ and $d(R) = 2^n - k$, where $1 \leq k \leq n$. By Proposition 2.3.6, $\mu(2^n - k) = k$, and hence $|R| \geq k$ for all R such that $d(R) = 2^n - k$. Hence $\deg(R) = (d(R) + |R|)/2 \geq 2^{n-1}$. Since also $\deg(R) \leq d(R) < 2^n$, $\binom{\deg(R)}{a} = 1$ by Proposition 1.5.3.

(ii) By Proposition 2.3.7, $2^n - n - 1 = \sum_{j=1}^{n-1}(2^j - 1)$ is the unique partition of $d = 2^n - n - 1$ as the sum of a minimum number of integers of the form $2^j - 1$. Since the number of terms has the same parity as d, all other such partitions have $\geq n + 1$ terms. Thus the minimal partition corresponds to the vector $R = (1, 1, \ldots, 1)$ of length $n - 1$, and $|R| \geq n + 1$ for all other R with $d(R) = d$. Applying Proposition 4.4.1**(i)**, $Sq^{2^{n-1}} Xq^{2^{n-1}-n-1}$ is the sum of all Milnor basis elements of degree $2^n - n - 1$ except $Sq(1, 1, \ldots, 1)$. The result follows from Proposition 3.4.5 and Example 4.1.5. □

By iterating **(i)**, we have $Xq^{2^n-k} = Sq^{2^{n-1}} \cdots Sq^{2^{k-1}} Xq^{2^{k-1}-k}$ for $1 \leq k \leq n$. In particular, $Xq^{2^n-1} = Sq^{2^{n-1}} \cdots Sq^2 Sq^1$.

The next result gives a formula for Xq^d for all d *not* of the form $2^k - 1$. In terms of the vector $\omega(d)$, the formula involves the last digit 0 that is followed by a digit 1, i.e. j is the largest number such that $\omega_j(d) = 0$ and $\omega_{j+1}(d) = 1$.

Proposition 4.4.3. *For $1 \leq k \leq n$ and $2^{k-1} < j \leq 2^k$,*

$$Xq^{2^{n+1}-j} = Sq^{2^n} Xq^{2^n-j} + Sq^{2^n-2^{k-1}} Xq^{2^n+2^{k-1}-j}.$$

Proof. Let $d = 2^{n+1} - j$. We show that $Sq^{2^n} Xq^{2^n-j}$ is the sum of all $Sq(R)$ of degree d with $\deg(R) \geq 2^n$, while $Sq^{2^n-2^{k-1}} Xq^{2^n+2^{k-1}-j}$ is the sum of all $Sq(R)$ of degree d with $\deg(R) < 2^n$. Thus every Milnor basis element $Sq(R)$ of degree d appears in one sum and not in the other, and so the result follows from Proposition 3.4.5.

By Proposition 4.4.1**(i)**, the first sum is $\sum_R \binom{\deg(R)}{2^n} Sq(R)$, and the second is $\sum_R \binom{\deg(R)}{2^n - 2^{k-1}} Sq(R)$, taken over all R with $d(R) = d$. Now $d/2 \leq (|R| + d)/2 = \deg(R)$ and $\deg(R) \leq d$. Since $d = 2^{n+1} - j$ and $2^{k-1} < j \leq 2^k$, we have $2^n - 2^{k-1} \leq d/2$ and $d < 2^{n+1} - 2^{k-1}$. Hence $2^n - 2^{k-1} \leq m < 2^{n+1} - 2^{k-1}$ where $m = \deg(R)$. By Proposition 1.5.3, for numbers m in this range $\binom{m}{2^n}$ is odd if $m \geq 2^n$ and even if $m < 2^n$, while $\binom{m}{2^n - 2^{k-1}}$ is even if $m \geq 2^n$ and odd if $m < 2^n$. □

4.5 Remarks

Milnor [93] gave the formula for the product $Sq(R)Sq(S)$ of his basis elements. He also gave a formula for the conjugate $\chi(Sq(R))$ of a basis element. This is rather complicated, and instead we shall give (in a later chapter) a formula due to Z. Li [85] for $\chi(Sq(R))$.

The conjugation χ on $\mathcal{A}_2$ was introduced in 1951 by Wu Wen-tsun and by R. Thom [150] in their work on characteristic classes of fibre bundles. This map is an essential element, the antipode, of the Hopf algebra structure of $\mathcal{A}_2$. Our approach in Section 4.3 uses the neat formulation by S. R. Bullett and I. G. Macdonald [19] of the Adem relations as a quadratic relation in a ring of formal power series over $\mathcal{A}_2$, which we present in Section 4.2. Our proof of the Bullett-Macdonald identity is based on the argument given in [19] (see also [137]), but avoids the direct use of residue calculus. The conjugation formulae of Section 4.4 were originally proved by Davis [39], Bausum [12] and Silverman [127].

Chapter 5

Combinatorial structures

5.0 Introduction

In this chapter we extend the combinatorial framework we require for work on the hit problem. The central role played by monomials with exponents all of the form $2^j - 1$, or 'spikes', is already clear. In Chapter 1, it was shown that a spike cannot be a term in any hit polynomial, so that the set of spikes in $P^d(n)$ is a subset of any monomial basis for $Q^d(n)$. We also saw that when $n = 1$ every monomial x^d that is not a spike is hit, and that in degrees d where the spikes do not already form a basis for $Q^d(2)$, such a basis is obtained by including one further monomial which lies in the $\mathbb{F}_2GL(2)$-module generated by the spikes. In Chapter 2, we saw that for all n, $Q^d(n) = 0$ when there are no spikes in $P^d(n)$.

In the general case, when $n \geq \mu(d)$, $P^d(n)$ can contain spikes with many different sets of exponents, each corresponding to a decomposition of d as a sum of n integers of the form $2^j - 1$, some of which may be zero. We shall call such a decomposition a **spike partition** of d. Example 1.4.7 gives a simple case where two spike partitions contribute to $Q^3(3)$. We introduce a natural partial ordering on the set $\mathbb{S}^d(n)$ of spike partitions of d with at most n nonzero parts, and show that the poset $\mathbb{S}^d(n)$ is a lattice with unique maximal and minimal elements. In the case of $P^3(3)$, the spike xyz is 'maximal' and the spikes x^3, y^3 and z^3 are 'minimal'.

It is also clear from Chapter 1 that the hit problem involves the structure of $P^d(n)$ as a module over $\mathbb{F}_2GL(n)$ (or over $\mathbb{F}_2M(n)$) in an essential way. Thus a second objective of this chapter is to begin the analysis of this structure at a combinatorial level, based simply on the observation that descending chains of submodules of $P^d(n)$ are obtained by taking the linear span of monomials which are divisible by successively higher 2-powers of the variables. For example the submodule S of $P^3(3)$ generated by monomials divisible by the square of some variable has dimension 9, and its quotient by the hit elements is the 6-dimensional submodule of $Q^3(3)$ generated by the minimal spikes. The **ω-vector** $\omega(f)$ of a

monomial f, introduced in Definition 3.3.2, is the key to describing this structure. In $P^3(3)$, the monomial xyz, with ω-vector $(3, 0, \ldots)$, generates the 1-dimensional quotient of $P^3(3)$ by the submodule S, which is the vector space spanned by the monomials with ω-vector $(1, 1)$.

In general, if f is a monomial in $P^d(n)$ with ω-vector $W = (w_1, \ldots, w_s)$, then $d = \sum_{i=1}^{s} 2^{i-1} w_i$, and so $\omega(f)$ corresponds to a decomposition of d as a sum of 2-powers, each occurring with multiplicity at most n. We shall call such a decomposition an **(n-bounded) binary partition** of d. We introduce a natural partial ordering on the set $\mathbb{B}^d(n)$ of n-bounded binary partitions of d and show that the poset $\mathbb{B}^d(n)$ is a distributive lattice with unique maximal and minimal elements.

In Section 5.1 we establish notation for various sets of vectors with integer entries, and introduce binary and spike partitions. In Section 5.2 we define two partial orderings, **dominance** and **2-dominance**, on the set of vectors. These partial orderings are compatible with the left and right linear orders introduced in Definition 3.1.7. In Section 5.3 we study the poset $\mathbb{V}^d$ of vectors of degree d with respect to 2-dominance, and relate it to $\mathbb{B}^d$, the binary partitions of d. In Section 5.4 we study the poset $\mathbb{W}^d$ of decreasing vectors of degree d, and relate it to $\mathbb{S}^d$, the spike partitions of d. In Section 5.5 we show that $\mathbb{S}^d$ has a unique minimal element $S^{\min}(d)$. In Section 5.6 we find maximal elements of these posets in the n-bounded case.

In the remaining sections, we apply these structures to the Steenrod algebra $\mathcal{A}_2$. In Section 5.7, we show that the dominance order on admissible vectors of degree d gives a lattice $\mathbb{A}^d$ isomorphic to $\mathbb{W}^d$ or $\mathbb{S}^d$. Finally in Section 5.8 we introduce the important **excess** function on $\mathcal{A}_2$, and evaluate it for Milnor basis elements and admissible monomials.

5.1 Vectors and partitions

Definition 5.1.1. Given a positive integer d, a **partition** of d is a way of expressing d as a sum of positive integers, called the **parts** of the partition, disregarding the order of the terms.

For example $d = 4$ has five partitions, namely 4, $3 + 1$, $2 + 2$, $2 + 1 + 1$ and $1 + 1 + 1 + 1$. In this chapter we are concerned with partitions of two special types. All parts of a **binary partition** are 2-powers $1, 2, 4, 8, \ldots$ and all parts of a **spike partition** are integers $1, 3, 7, \ldots$ of the form $2^j - 1$. Thus $d = 4$ has four binary partitions and two spike partitions. By convention, the parts of a partition are written in decreasing order as the entries of a vector. We shall often use multiset notation for partitions, so that for example the binary partition $(4, 2, 2, 1, 1, 1, 1)$ of 12 can be written more compactly as $(42^2 1^4)$ and the spike partition $(3, 3, 3, 1, 1, 1)$ of 12 as $(3^3 1^3)$. When vector notation is used, we extend

the notation as in 3.1.7 by allowing trailing zeros. By convention, we regard the zero vector, or the empty multiset, as a partition of $d = 0$, with no parts.

Recall from Definition 3.1.7 that a *vector* is a sequence $V = (v_1, v_2, \ldots)$ of integers ≥ 0 with only a finite number of nonzero entries, with *modulus* $|V| = \sum_{i\geq 1} v_i$, and *degree* $\deg(V) = \sum_{i\geq 1} 2^{i-1} v_i$.

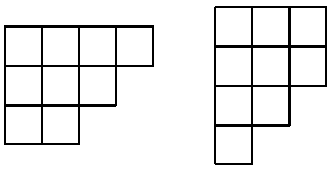
Definition 5.1.2. The vector $V = (v_1, v_2, \ldots)$ is **decreasing** if $v_i \geq v_{i+1}$ for $i \geq 1$, and is **strongly decreasing** if $v_i > v_{i+1}$ for $1 \leq i < \mathrm{len}(V)$.

We write $\mathbb{V}$ for the set of all vectors and $\mathbb{V}^d$ for the set of all vectors of degree d. Note that $\mathbb{V}^d$ is a finite set. We also write $\mathbb{V}(n)$ and $\mathbb{V}^d(n)$ for the corresponding sets of vectors V with entries $v_j \leq n$ for all $j \geq 1$, which we call n-**bounded** vectors. We write $\mathbb{W}$, $\mathbb{W}^d$, $\mathbb{W}(n)$ and $\mathbb{W}^d(n)$ for the corresponding sets of decreasing vectors.

Example 5.1.3. $\mathbb{V}^7 = \{(7), (5,1), (3,2), (3,0,1), (1,3), (1,1,1)\}$, with subsets $\mathbb{W}^7 = \{(7), (5,1), (3,2), (1,1,1)\}$ and $\mathbb{V}^7(3) = \{(3,2), (1,3), (3,0,1), (1,1,1)\}$. So $\mathbb{W}^7(3) = \mathbb{W}^7 \cap \mathbb{V}(3) = \mathbb{W}(3) \cap \mathbb{V}^7 = \{(3,2), (1,1,1)\}$.

A partition of d may be regarded either as a decreasing vector V with $|V| = d$, or as a multiset $(a_1^{m_1} \cdots a_r^{m_r})$ of positive integers whose sum is d. For $i \geq 1$, the ith entry v_i of V is the i**th part** of the partition, and the length of V is the **length** of the partition. Thus parts which are equal to zero do not contribute to the length. The **multiplicity** of a part a_i is its exponent m_i for $1 \leq i \leq r$. We denote the set of partitions of d by $\mathbb{P}^d$. A standard way to represent a partition of d graphically is by its **(Ferrers) diagram**. This is a left justified array of 'boxes' in which for $i \geq 1$ the number of boxes in the ith row is the ith part of the partition. In matrix notation, the diagram of V has a box (i, j) if and only if $v_i \geq j$. The total number of boxes is d.

Example 5.1.4. The Ferrers diagrams of the partitions $(4, 3, 2)$ and $(3, 3, 2, 1)$ in $\mathbb{P}^9$ are shown below.



Given a partition $V = (v_1, v_2, \ldots)$, the **conjugate** (or **transpose**) partition $W = (w_1, w_2, \ldots)$ is defined by transposition of its diagram, i.e. exchange of rows and columns, or reflection in the main diagonal. Thus, for $j \geq 1$, w_j is the number of parts v_i of V such that $v_i \geq j$. For example the partitions $V = (4, 3, 2)$ and $W = (3, 3, 2, 1)$ are conjugates, and we write $W = V^t$ or $V = W^t$.

The ω-vector of a monomial f in P^n was defined in Chapter 3 as the column sum vector of the n-block associated to f. Thus $\mathbb{V}^d(n)$ is the set of ω-vectors of monomials in $P^d(n)$. The exponent vectors $(d_1, \ldots, d_n)$ of monomials in $P^d(n)$,

with their entries taken in decreasing order, are partitions of d. For example, if the exponent vector $(d_1, \ldots, d_n)$ of f is $(3,3,1)$, $(3,2,1,1)$ or $(2,2,1,1,1)$, then $\omega(f) = (3,2)$. In the case of $(3,3,1)$, the corresponding monomials f are spikes. For $n \geq 7$, the spikes of degree 7 in $P(n)$ are the monomials obtained from x_1^7, $x_1^3x_2^3x_3$, $x_1^3x_2x_3x_4x_5$ and $x_1x_2x_3x_4x_5x_6x_7$ by permutations of the variables. Clearly the ω-vector of a spike is decreasing. The ω-vectors of the spikes just listed are the elements of $\mathbb{W}^7$.

The sets $\mathbb{V}^d$ and $\mathbb{W}^d$ correspond in a natural way to two special classes of partitions of d, which we now define.

Definition 5.1.5. A partition $B = (b_1, b_2, \ldots)$ is a **binary partition** if all its nonzero parts b_j are of the form 2^k where $k \geq 0$. A partition $S = (s_1, s_2, \ldots)$ is a **spike partition** if all its nonzero parts s_j are of the form $2^k - 1$, where $k \geq 1$.

We write $\mathbb{B}^d$ for the set of binary partitions of d and $\mathbb{S}^d$ for the set of all spike partitions of d. Thus $\mathbb{B}^d$ and $\mathbb{S}^d$ are subsets of $\mathbb{P}^d$. We also write $\mathbb{B}^d(n)$ and $\mathbb{S}^d(n)$ for the subsets of $\mathbb{B}^d$ and $\mathbb{S}^d$ respectively given by partitions in which all parts occur with multiplicity $\leq n$. Note that the superscript d refers to the modulus of vectors in $\mathbb{B}^d$, $\mathbb{S}^d$ and $\mathbb{P}^d$, but refers to the degree of vectors in $\mathbb{V}^d$ and $\mathbb{W}^d$.

Proposition 5.1.6. *There are bijections between $\mathbb{V}^d$ and $\mathbb{B}^d$, and between $\mathbb{W}^d$ and $\mathbb{S}^d$, such that $\mathbb{V}^d(n)$ and $\mathbb{W}^d(n)$ correspond to $\mathbb{B}^d(n)$ and $\mathbb{S}^d(n)$ respectively.*

Proof. A binary partition $(2^{a_1}, \ldots, 2^{a_r})$ of d corresponds to a monomial $f = x_1^{2^{a_1}} \cdots x_r^{2^{a_r}}$ of degree d where $a_1 \geq \ldots \geq a_r$, and so to an ω-vector $\omega(f) \in \mathbb{V}^d$. Similarly, spike partitions of d correspond to monomials of degree d where all exponents are of the form $2^k - 1$, and $\mathbb{W}^d$ is the set of ω-vectors of these monomials. For the last statement, consider monomials in $P(n)$. □

These correspondences are more transparent in multiset notation. In the case of $\mathbb{B}^d$, the corresponding element of $\mathbb{V}^d$ is given by reading the exponents of the multiset in reverse order, so that (53^21^4) gives $(4,2,1)$. In the case of $\mathbb{S}^d$, we proceed as follows. Let $S = (s_1, s_2, \ldots)$ be a spike partition of d, so that $s_i = 2^{k_i} - 1$ for $i \geq 1$. Then $K = (k_1, k_2, \ldots)$ is a decreasing vector, i.e. a partition. In the notation of Definition 3.3.6, $K = \alpha(f)$ where $f = x_1^{2^{k_1}-1} \cdots x_n^{2^{k_n}-1}$ is a spike with exponents in decreasing order. Then the correspondence of Proposition 5.1.6 associates to S the ω-vector $\omega(f) = K^t$. For example, if $S = (3^31^3)$ then $\alpha(f) = (2^31^3)$ and $\omega(f) = \alpha(f)^t = (6,3)$.

Recall from Definition 3.3.6 that a monomial $f = x_1^{d_1} \cdots x_n^{d_n}$ in $P(n)$ can be represented diagrammatically by an n-**block**, a left justified array with ith row $\omega(x_i^{d_i})$ for $1 \leq i \leq n$. We use matrix notation for blocks, so that, for $i, j \geq 1$, $b_{i,j}$ is the entry of B which is in row i and column j. Thus if B is the block corresponding to f, $b_{i,j} = 1$ if $2^{j-1} \in \text{bin}(d_i)$, and $b_{i,j} = 0$ otherwise. In particular, if f is a spike with exponents in decreasing order, then the corresponding block is obtained by

replacing each box in the diagram of the partition $\alpha(f)$ by a digit 1, following our convention of omitting trailing zeros in the rows. However, if f does not involve all the variables $x_1, \ldots, x_n$, we retain at least one zero as a marker.

Definition 5.1.7. A n-block B is called a **Ferrers** block if it corresponds to a spike monomial $f = x_1^{2^{k_1}-1} \cdots x_n^{2^{k_n}-1}$, where $k_i \geq k_{i+1} \geq 0$ for $1 \leq i < n$.

Example 5.1.8. The Ferrers block below corresponds to $x_1^{15}x_2^3x_3^3$ in $P^{21}(4)$.

$$\begin{array}{cccc} 1 & 1 & 1 & 1 \\ 1 & 1 & & \\ 1 & 1 & & \\ 0 & & & \end{array}$$

If $K = (k_1, k_2, \ldots)$ is a partition of d, so that $|K| = d$, we regard K as $\alpha(f)$, and obtain the Ferrers block corresponding to f by replacing each box in the Ferrers diagram of K by a digit 1. Then $\omega(f) = K^t$, the conjugate partition.

5.2 Dominance

Two linear orders on $\mathbb{V}$, the left order $<_l$ and the right order $<_r$, were defined in Section 3.1. We next define two partial orders on $\mathbb{V}$. Recall that a **partial order** on a set $\mathbb{X}$ is a binary relation $\leq$ which is reflexive ($X \leq X$ for all $X \in \mathbb{X}$), antisymmetric ($X \leq Y$ and $Y \leq X$ only if $X = Y$ for all $X, Y \in \mathbb{X}$) and transitive ($X \leq Y$ and $Y \leq Z$ imply $X \leq Z$ for all $X, Y, Z \in \mathbb{X}$). A set with a partial order is called a **poset**. A subset $\mathbb{X}'$ of $\mathbb{X}$ with the partial order defined by $X \leq Y$ in $\mathbb{X}'$ if and only if $X \leq Y$ in $\mathbb{X}$ is called a **subposet** of $\mathbb{X}$. The usual conventions for the use of the symbols $<$, $>$ and $\geq$ in relation to $\leq$ apply to partial orders.

Definition 5.2.1. For $V, W \in \mathbb{V}$, V is less than or equal to W in **dominance** order, written $V \preceq W$, if for all $k \geq 1$

$$\sum_{j=1}^{k} v_j \leq \sum_{j=1}^{k} w_j.$$

The dominance order is of great importance in the combinatorics of partitions, the theory of symmetric functions and in representation theory. We shall be particularly concerned with the comparison of vectors arising from spike partitions and binary partitions. For this reason we introduce an exponential version of dominance.

Definition 5.2.2. For $V, W \in \mathbb{V}$, V is less than or equal to W in **2-dominance** order, written $V \preceq_2 W$, if for all $k \geq 1$

$$\sum_{j=1}^{k} 2^{j-1} v_j \leq \sum_{j=1}^{k} 2^{j-1} w_j.$$

Proposition 5.2.3. *The left order $<_l$ refines both dominance and 2-dominance on $\mathbb{V}$. The right order $<_r$ refines dominance for vectors of the same modulus, and 2-dominance for vectors of the same degree.*

Proof. It is clear that if $V \prec W$ or $V \prec_2 W$ then $V <_l W$. Let $V \prec W$ where $|V| = |W|$. Then $\sum_{i=1}^{l} v_i = \sum_{i=1}^{l} w_i$, where l is the maximum length of V and W. Let k be the minimum value of j such that $\sum_{i=1}^{j} v_i = \sum_{i=1}^{j} w_i$. Then $v_j = w_j$ for $j > k$ and, since $\sum_{i=1}^{k-1} v_i < \sum_{i=1}^{k-1} w_i$, $v_k > w_k$. Hence $V <_r W$.

Similarly, if $\deg(V) = \deg(W)$ and $V \prec_2 W$, then, for some k, $\sum_{i=1}^{j} 2^{i-1} v_i = \sum_{i=1}^{j} 2^{i-1} w_i$ for $k \leq j \leq l$, and $\sum_{i=1}^{k-1} 2^{i-1} v_i < \sum_{i=1}^{k-1} 2^{i-1} w_i$. As before, it follows that $v_j = w_j$ for $j > k$ and $v_k > w_k$. □

We shall normally use dominance only to compare vectors with the same modulus, and 2-dominance only to compare vectors with the same degree. In both contexts, the relevant form of the dominance order is compatible with both the left and right orders. For partitions of d, regarded as decreasing vectors of modulus d, we prove next that conjugacy reverses the dominance relation. This is not true for the left and right orders, for example when $V = (4, 1, 1)$ and $W = (3, 3)$, we have $V^t = (3, 1, 1, 1)$ and $W^t = (2, 2, 2)$ so that in the left order $V >_l W$ and $V^t >_l W^t$, and in the right order $V <_r W$ and $V^t <_r W^t$.

Proposition 5.2.4. *Let V and W be partitions with $|V| = |W|$. Then $V \succeq W$ if and only if $V^t \preceq W^t$.*

Proof. A pair of integers $k, l \geq 1$ divides the diagram of a partition into 'northwest', 'northeast', 'southwest' and 'southeast' quadrants of boxes (i, j), defined by $i \leq k$ and $j \leq l$, $i \leq k$ and $j > l$, $i > k$ and $j \leq l$, and finally $i > k$ and $j > l$. Let a, b, c, d be the number of boxes in the four quadrants of the diagram of V in the listed order, and let a', b', c', d' be the corresponding numbers for W. Then $a + b + c + d = |V| = |W| = a' + b' + c' + d'$, and $a + b$, $a' + b'$ depend only on k, $a + c$, $a' + c'$ only on l.

As it suffices to prove the implication in one direction, we may assume that $V \succeq W$. Then for all $k \geq 1$ we have $a + b \geq a' + b'$, and we must prove that $a + c \leq a' + c'$ for all $l \geq 1$. Given $l \geq 1$, let k be the largest number such that $s_k \geq l$. Then, with quadrants defined by (k, l), the northwest quadrant of the diagram of W is full and the southeast quadrant empty. Hence $a' = kl$ and $d' = 0$. Hence $a \leq a'$ and so $b \geq b'$. Hence $b + d \geq b' = b' + d'$ and since $|V| = |W|$ it follows that $a + c \leq a' + c'$, and so $V^t \preceq W^t$. □

The following result relates dominance and 2-dominance.

Proposition 5.2.5. *Let $V' = (v_1, v_2, v_2, v_3, v_3, v_3, v_3, \ldots)$ be the vector derived from the vector V in which v_j is repeated 2^{j-1} times for all $j \geq 1$, and let W' be the corresponding vector derived from the vector W. Then $V \preceq_2 W$ if and only if $V' \preceq W'$.*

Proof. By definition $V' \preceq W'$ if and only if $\sum_{j=1}^{k} v'_j \leq \sum_{j=1}^{k} w'_j$ for $1 \leq k \leq n$. Let r be the largest integer such that $2^r - 1 \leq k$ and write $k = 2^r - 1 + s$ where $0 \leq s < 2^r$. Then for all $r \geq 1$ and $0 \leq s < 2^r$

$$\sum_{j=1}^{r} 2^{j-1} v_j + s v_{r+1} \;\leq\; \sum_{j=1}^{r} 2^{j-1} w_j + s w_{r+1} \tag{5.1}$$

Setting $s = 0$ gives $V \preceq_2 W$. Conversely, if $V \preceq_2 W$ then $\sum_{j=1}^{r} 2^{j-1} v_j \leq \sum_{j=1}^{r} 2^{j-1} w_j$ and $\sum_{j=1}^{r+1} 2^{j-1} v_j \leq \sum_{j=1}^{r+1} 2^{j-1} w_j$. By adding $2^r - s$ times the first inequality to s times the second and dividing by 2^r, we obtain (5.1). Hence $V' \preceq W'$. □

5.3 Vectors of degree d

In this section we show that under the bijection of Proposition 5.1.6 between $\mathbb{V}^d$ and $\mathbb{B}^d$, the 2-dominance order on $\mathbb{V}^d$ corresponds to the natural partial order on $\mathbb{B}^d$ by **refinement**. We begin by recalling some standard terms in relation to posets.

Definition 5.3.1. If $\mathbb{X}$ is a poset and $X, Y \in \mathbb{X}$, X **covers** Y if $Y < X$ and there is no element $Z \in \mathbb{X}$ such that $Y < Z < X$. A **grading** on a poset $\mathbb{X}$ is a function gr from $\mathbb{X}$ to the integers ≥ 0 such that $\mathrm{gr}(X) = \mathrm{gr}(Y) + 1$ if X covers Y, and X is then **graded** by gr.

Proposition 5.3.2. **(i)** *If $U <_l V$ in $\mathbb{V}^d$ and k is the smallest number such that $u_k < v_k$, then $v_k - u_k \geq 2$.*

(ii) *For $V, W \in \mathbb{V}^d$, V covers W in the 2-dominance order $\prec_2$ if and only if $w_k = v_k - 2$ and $w_{k+1} = v_{k+1} + 1$ for some $k \geq 1$ and $w_j = v_j$ for $j \neq k, k+1$.*

(iii) *The poset $\mathbb{V}^d$ is graded by the modulus $|V|$ of a vector $V \in \mathbb{V}^d$.*

Proof. For **(i)**, let $U \in \mathbb{V}^d$ be any vector such that $U <_l V$. Then for some $k \geq 1$ we have $u_j = v_j$ for $j < k$ and $u_k < v_k$. Since $\sum_{j \geq 1} 2^{j-1} u_j = d = \sum_{j \geq 1} 2^{j-1} v_j$, by subtraction we have $\sum_{j \geq k} 2^{j-1}(v_j - u_j) = 0$. Cancelling a factor 2^{k-1}, this gives $v_k - u_k = -\sum_{j>k} 2^{j-k}(v_j - u_j) = 0 \bmod 2$.

For **(ii)**, it is clear that if $V, W \in \mathbb{V}^d$ are as described, then $W \prec_2 V$ and $|W| = |V| - 1$. Let $U \in \mathbb{V}^d$ be any vector such that V covers U, let $k \geq 1$ be as in **(i)**, and let W be defined as in **(ii)**. Then $s_j = u_j$ for $j < k$ and, by **(i)**, $s_k \geq u_k$. Since $U \prec_2 V$ and for all $l > k$ we have $\sum_{j=1}^{l} 2^{j-1} w_j = \sum_{j=1}^{l} 2^{j-1} v_j$, $U \prec_2 W$. Since by hypothesis V covers U, it follows that $U = W$. This proves **(ii)**, and **(iii)** follows since $|W| = |V| - 1$ for all choices of k. □

The set $\mathbb{B}^d$ of binary partitions of d has a natural partial ordering by **refinement**.

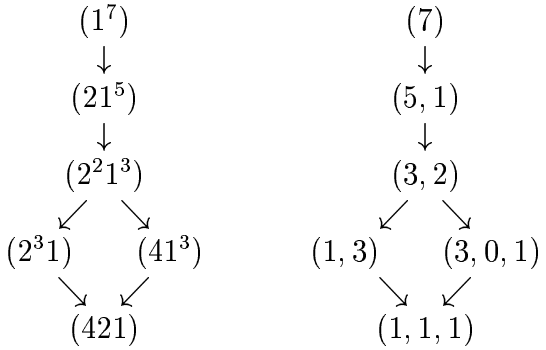
Definition 5.3.3. Let A and B be binary partitions of d. Then A **refines** of B, or $A \geq B$ in the **refinement order** if there are binary partitions of the parts of B which can be combined to give A.

We can refine a binary partition by splitting a part 2^k such that $k \geq 1$ as $2^{k-1} + 2^{k-1}$, so increasing the length of the partition by 1. Given binary partitions A and B of d, $B \leq A$ in the refinement order if A can be obtained from B by a finite number of steps of this type. The binary expansion of d is the unique element of minimal length $\alpha(d) = |\text{bin}(d)|$ in $\mathbb{B}^d$, and (1^d) is the unique element of maximal length. Every element of $\mathbb{B}^d$ can be obtained from the binary expansion of d by a sequence of refinements, and can in turn be reduced to (1^d) by continuing this process. Since the length of the partition is increased by 1 at each step, all such chains in $\mathbb{B}^d$ are of the same length. The length function is a grading on $\mathbb{B}^d$.

For monomials whose exponents are the parts of a binary partition, refinement by splitting corresponds to replacement of a factor x^{2^k} by $y^{2^{k-1}} z^{2^{k-1}}$ for some variables x, y, z with $y \neq z$. Thus in terms of the bijection between binary partitions and vectors given by Proposition 5.1.6, the splitting corresponds to replacement of two consecutive entries (v_{k-1}, v_k) of the ω-vector of the monomial by $(v_{k-1}+2, v_k-1)$. By Proposition 5.3.2, this corresponds to a covering ω-vector for the 2-dominance order on $\mathbb{V}^d$. We summarize these results as follows.

Proposition 5.3.4. *The bijection between $\mathbb{V}^d$ and $\mathbb{B}^d$ given by Proposition 5.1.6 is an isomorphism of graded posets, where $\mathbb{V}^d$ is ordered by 2-dominance and $\mathbb{B}^d$ by refinement, and the grading is the modulus $|V|$ in $\mathbb{V}^d$ and the length $\text{len}(B)$ in $\mathbb{B}^d$. The poset $\mathbb{B}^d$ has a unique maximal element (1^d) and a unique minimal element given by the binary decomposition of d. The poset $\mathbb{V}^d$ has a unique maximal element (d) and a unique minimal element $\omega(d)$.* □

Example 5.3.5. The diagrams below show the refinement order on $\mathbb{B}^7$, the binary partitions of 7, and the 2-dominance order on $\mathbb{V}^7$. Note that refinement of binary partitions implies dominance but is not equivalent to it. For example, $(2^31) \prec (41^3)$, but neither refines the other.



The isomorphic posets $\mathbb{V}^d$ and $\mathbb{B}^d$ are of a particularly nice type known as **distributive lattices**. We recall the definitions.

Definition 5.3.6. A **lattice** $\mathbb{L}$ is a poset in which every pair of elements X, Y has a least upper bound, denoted by $\sup(X, Y)$ or $X \vee Y$, and a greatest lower bound, denoted by $\inf(X, Y)$ or $X \wedge Y$, where these terms have their usual meanings. The lattice $\mathbb{L}$ is **distributive** if for all $X, Y, Z \in \mathbb{L}$

$$X \vee (Y \wedge Z) = (X \vee Y) \wedge (X \vee Z) \text{ and } X \wedge (Y \vee Z) = (X \wedge Y) \vee \ (X \wedge Z),$$

the two conditions being equivalent. A subposet $\mathbb{L}'$ of $\mathbb{L}$ is a **sublattice** if $X \vee Y$ and $X \wedge Y$ are in $\mathbb{L}'$ when X and Y are in $\mathbb{L}'$.

Every distributive lattice is graded, and the grading gr is related to the lattice operations by the formula $\mathrm{gr}(V) + \mathrm{gr}(W) = \mathrm{gr}(V \vee W) + \mathrm{gr}(V \wedge W)$.

The following definition is useful in dealing with 2-dominance.

Definition 5.3.7. For $V \in \mathbb{V}(d)$ the **partial degree vector** $\widehat{V} = (\widehat{v}_1, \widehat{v}_2, \ldots)$, where $\widehat{v}_k = \sum_{j=1}^{k} 2^{j-1} v_j$.

Thus $V \preceq_2 W$ if and only if $\widehat{v}_k \leq \widehat{w}_k$ for all k. $\widehat{V}$ is an increasing sequence such that $\widehat{v}_j \equiv d \bmod 2^j$ for all $j \geq 1$, and $\widehat{v}_l = d$ where $l = \mathrm{len}(V)$. These conditions characterize partial degree vectors of elements of $\mathbb{V}^d$, since V can be recovered from $\widehat{V}$ by the formula $v_j = (\widehat{v}_j - \widehat{v}_{j-1})/2^{j-1}$ for $j \geq 1$, where $v_0 = 0$.

Proposition 5.3.8. *$\mathbb{V}^d$ is a distributive lattice.*

Proof. For $V, W \in \mathbb{V}^d$ the vectors with components $\max(\widehat{v}_j, \widehat{w}_j)$ and $\min(\widehat{v}_j, \widehat{w}_j)$ for $j \geq 1$ are partial degree vectors of elements of $\mathbb{V}(d)$ which we may write as $V \vee W$ and $V \wedge W$. Then $V \vee W$ and $V \wedge W$ have the properties required to be the least upper bound and greatest lower bound of V and W. The distributive laws follow from the corresponding properties $\max(U, \min(V, W)) = \min(\max(U, V), \max(U, W))$, $\min(U, \max(V, W)) = \max(\min(U, V), \min(U, W))$ for a totally ordered set $\{U, V, W\}$. □

5.4 Partitions of degree d

In this section we show that the poset $\mathbb{W}^d$ of decreasing vectors of degree d, ordered by 2-dominance, is a lattice (not distributive, in general). Recall from Proposition 5.1.6 that there is a bijection between $\mathbb{W}^d$ and $\mathbb{S}^d$ which associates to a spike partition S of d the ω-vector of a spike whose exponents are the parts of S.

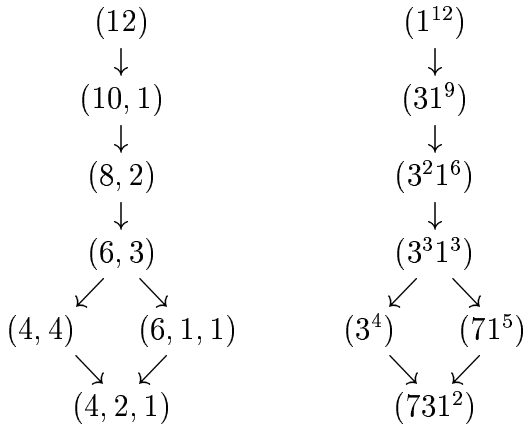
Proposition 5.4.1. *Let $d \geq 0$ and let f and g be spikes of degree d with decreasing exponent vectors S and T. Then $\omega(f) \prec_2 \omega(g)$ if and only if $S \succ T$.*

Proof. Let $V = \omega(f)$ and let $W = \omega(g)$. By Proposition 5.2.5, $V \prec_2 W$ if and only if $V' \prec W'$, where V' and W' are derived from V and W by repeating the jth entry 2^{j-1} times for $j \geq 1$. But V' and W' are the conjugates of the partitions S and T. (For example, when $S = (7,3,1,1)$, $V' = S^t = (4,2,2,1,1,1,1)$ and $V = (4,2,1)$.) By Proposition 5.2.4, $V' \prec W'$ if and only if $T \prec S$. □

Applying Proposition 5.1.6, we have proved the following result.

Proposition 5.4.2. *The poset $\mathbb{W}^d$ of decreasing vectors of degree d with 2-dominance order $\preceq_2$ is isomorphic to the poset $\mathbb{S}^d$ of spike partitions of d with reversed dominance order $\succeq$.* □

Example 5.4.3. We show $(\mathbb{W}^{12}, \preceq_2)$ and $(\mathbb{S}^{12}, \succeq)$ below. The bijection is given by conjugation in $\mathbb{W}^{12}$ followed by the map $k \mapsto 2^k - 1$, or equivalently by $V \mapsto (V')^t$ where V' is constructed as described above.



The poset $\mathbb{W}^d$ is a subposet of the lattice $\mathbb{V}^d$. We shall prove that $\mathbb{W}^d$ is also a lattice, but not in general a sublattice, of $\mathbb{V}^d$, because the same two elements can have different least upper bounds in $\mathbb{W}^d$ and in $\mathbb{V}^d$.

Example 5.4.4. Consider the vectors $V = (14,1,1,1)$ and $W = (4,4,4)$ in $\mathbb{W}^{28}$. Calculating $V \vee W$ in $\mathbb{V}^{28}$ from the partial degree vectors, we obtain $(14,1,3)$, which is not decreasing. However $(14,1,3)$ is covered in $\mathbb{V}^{28}$ by $(14,3,2)$, which we take as $\sup(V,W)$ in $\mathbb{W}^{28}$.

The next result shows that this problem does not arise for the greatest lower bound.

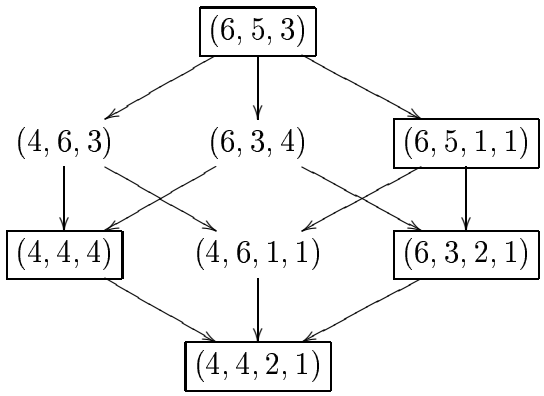
Proposition 5.4.5. *For all V and W in $\mathbb{W}^d$, $V \wedge W$ is in $\mathbb{W}^d$.*

Proof. Recall that if $U = V \wedge W$ then the partial degree vectors $\widehat{U}, \widehat{V}, \widehat{W}$ of U, V, W are related by $\widehat{u}_i = \min(\widehat{v}_i, \widehat{w}_i)$ for $i \geq 1$. For general U the partial degree vector $\widehat{U}$ is related to U by $u_i = (\widehat{u}_i - \widehat{u}_{i-1})/2^{i-1}$, where $u_0 = 0$, and so U is decreasing if and only if $\widehat{u}_i - \widehat{u}_{i-1} \geq (\widehat{u}_{i+1} - \widehat{u}_i)/2$ for $i \geq 1$. The result follows

from the observation that if $a_1 \le a_2 \le a_3$ and $b_1 \le b_2 \le b_3$ are triples of integers such that $a_2 - a_1 \ge (a_3 - a_2)/2$ and $b_2 - b_1 \ge (b_3 - b_2)/2$, then $c_2 - c_1 \ge (c_3 - c_2)/2$ where $c_i = \min(a_i, b_i)$ for $i = 1, 2, 3$. This is easy to check: let $c_1 = a_1$, say, and separate the four cases $(c_2, c_3) = (a_2, a_3)$, (a_2, b_3), (b_2, a_3) and (b_2, b_3). □

It follows from Proposition 5.4.5 that we can define a lattice structure on $\mathbb{W}^d$ by $\inf(V, W) = V \wedge W$ and $\sup V, W = \inf\{U \in \mathbb{W}^d : V \vee W \preceq_2 U\}$. Since $\mathbb{W}^d$ has a unique maximal element (d), the set of such U is non-empty. The next example shows that in general the lattice $\mathbb{W}^d$ is not graded, and hence not distributive.

Example 5.4.6. The diagram below shows the interval $\{(4,4,2,1) \preceq V \preceq (6,5,3)\}$ in the lattice $\mathbb{V}^{28}$. Only the boxed elements are in $\mathbb{W}^{28}$.



5.5 Minimal spikes

Recall from Definition 2.3.2 that $\mu(d)$ is the minimum length of a spike partition of d. Since all its parts are odd, the length of such a partition has the same parity as d.

Proposition 5.5.1. *Let $l = d \bmod 2$ and let $\mu(d) \le l \le d$. Then d has a spike partition of length l.*

Proof. Since a spike partition of length $\mu(d)$ exists, we may assume by recursion on l that d has a spike partition S of length $l - 2$. Since $l \le d$, S has at least one part $s_i > 1$. Then $s_i = 2^{k+1} - 1 = (2^k - 1) + (2^k - 1) + 1$ where $k \ge 1$. By splitting s_i in this way we convert S to a spike partition of length l. □

There may be more than one spike partition of d of length $\mu(d)$, for example $17 = 15 + 1 + 1 = 7 + 7 + 3$. We shall prove that for all d there is a unique spike partition $S^{\min}(d)$ such that either the parts are all distinct, for example $19 = 15 + 3 + 1$, or the parts are distinct except that the smallest part occurs twice, for example $17 = 15 + 1 + 1$. The partition $T = S^{\min}(d)$ has minimum length $\mu(d)$, and we construct it by the 'greedy algorithm'.

Proposition 5.5.2. *Given $d > 0$, let $d_1 = d$ and, for $i \geq 1$, let $d_{i+1} = d_i - t_i$, where t_i is the largest integer of the form $2^j - 1$ such that $t_i \leq d_i$. Then $T = (t_1, t_2, \ldots)$ is a spike partition of d, and*

(i) *if $t_i = t_{i+1}$, then $t_{i+2} = 0$;*

(ii) *T is the unique element of $\mathbb{S}^d$ satisfying* **(i)**;

(iii) *T is the unique element of $\mathbb{S}^d$ such that $\sum_i \alpha(t_i)$ is minimal;*

(iv) $\mathrm{len}(T) = \mu(d)$.

Proof. Clearly $T \in \mathbb{S}^d$ by construction. For **(i)**, let $t_i = 2^k - 1$. Then $2^k - 1 \leq d_i < 2^{k+1} - 1$, and $d_{i+1} = d_i - 2^k + 1$, so $0 \leq d_{i+1} < 2^k$. Now $t_{i+1} = 2^k - 1$ if and only if $2^k - 1 \leq d_{i+1} < 2^{k+1} - 1$, so we must have $d_{i+1} = 2^k - 1$ and $d_{i+2} = 0$. Hence $t_{i+2} = 0$.

For **(ii)**, let $k \geq 1$ and let $1 \leq d \leq 2^{k+1} - 2$. Then the first part of any spike partition of d is is $\leq 2^k - 1$. The spike partitions satifying **(i)** are either strictly decreasing, and so correspond to the $2^k - 1$ non-empty subsets of $\{1, 2, \ldots, k\}$, or they are strictly decreasing except that the last term is repeated, giving a further $2^k - 1$ choices. The construction therefore gives a bijection from the set $\{1, 2, \ldots, 2^{k+1} - 2\}$ to this set of spike partitions.

For **(iii)**, let $S \in \mathbb{S}^d$ be a spike partition $\neq T$. Then by **(i)** we have $s_i = s_{i+1} \geq s_{i+2} > 0$ for some $i \geq 1$. Let $a = s_i$ and let $b = s_{i+2}$, so that $a \geq b > 0$. Since $(2^a - 1) + (2^a - 1) + (2^b - 1) = (2^{a+1} - 1) + (2^{b-1} - 1) + (2^{b-1} - 1)$, we can obtain another spike partition S' of d by replacing the parts $2^a - 1$, $2^a - 1$, $2^b - 1$ by $2^{a+1} - 1$, $2^{b-1} - 1$, $2^{b-1} - 1$, and sorting the parts into decreasing order. Then $\alpha(S') < \alpha(S)$ since $a + 2b - 1 < 2a + b$, proving **(iii)**. Further $\mathrm{len}(S') \leq \mathrm{len}(S)$, and so if $S \neq T$ we can replace S by a spike partition S' with smaller α-value and no greater length. Hence T has minimal length $\mu(d)$, proving **(iv)**. □

Definition 5.5.3. Let $n \geq l = \mu(d)$, so that $S^{\min}(d) = (t_1, \ldots, t_l)$. A spike $f \in P^d(n)$ is called a **minimal spike** if $f = x_1^{t_1} \cdots x_l^{t_l}$, or any monomial obtained from this by permuting exponents. The α-vector $\alpha(f)$ is called the **minimal α-vector** for spikes of degree d, and is denoted by $\alpha^{\min}(d)$, and the ω-vector $\omega(f)$ is called the **minimal ω-vector** for spikes of degree d, and is denoted by $\omega^{\min}(d)$. The Ferrers block corresponding to f is called the **minimal Ferrers block** for spikes in $P^d(n)$, and is denoted by $F^{\min}(n, d)$.

By Propositions 5.5.2 and 5.1.6, $\omega^{\min}(d)$ is the unique minimal element of the poset $\mathbb{W}^d$. The conjugate partition is $\alpha^{\min}(d)$. Proposition 5.5.2**(i)** characterizes $F^{\min}(d)$ as a block with $\mu(d)$ nonzero rows, none of which are equal except possibly the last two. This property is preserved by duplication of the first column of a Ferrers block, and so we obtain the following property of μ. (Compare Proposition 2.3.5.)

Proposition 5.5.4. *If $\mu(d) = k$, then $\mu(2d+k) = k$.* □

Example 5.5.5. Let $d = 21 = 15+3+3$. Then $S^{\min}(d) = (15,3,3)$, $\alpha^{\min}(d) = (4,2,2)$, $\omega^{\min}(d) = (3,3,1,1)$ and for $n = 4$, $F = F^{\min}(n,d)$ is the Ferrers block shown in Example 5.1.8. The corresponding Ferrers block for $d = 45 = 31+7+7$ is obtained by duplicating the first column.

The next result explains how to calculate $\omega^{\min}(d)$ using the function μ.

Proposition 5.5.6. *Let $V = \omega^{\min}(d)$ and let $d_1 = d$. Then v_j is given recursively by $v_j = \mu(d_j)$, $d_{j+1} = (d_j - v_j)/2$ for $j \geq 1$.*

Proof. By Proposition 5.5.2**(iv)**, $v_1 = \mu(d)$, so let $d_2 = (d - \mu(d))/2$. The result follows by induction on d, using minimality of $\omega^{\min}(d)$, since prefixing v_1 to vectors in $\mathbb{W}^{d_2}$ preserves the partial order $\preceq_2$. Alternatively if $T = (t_1, \ldots, t_l)$ satisfies Proposition 5.5.2**(i)** and $t_l > 0$, then the vector T' given by $t'_i = (t_i - 1)/2$ for $1 \leq i \leq l$ is a spike partition of d_2 and satisfies **(i)**. □

5.6 Maximal elements of $\mathbb{V}^d(n)$ and $\mathbb{W}^d(n)$

In this section we consider the subposets $\mathbb{V}^d(n)$ and $\mathbb{W}^d(n)$ of $\mathbb{V}^d$ and $\mathbb{W}^d$ given by vectors whose entries are bounded by n, with partial order $\prec_2$. Of course $\mathbb{V}^d(n) = \mathbb{V}^d$ and $\mathbb{W}^d(n) = \mathbb{W}^d$ if $n \geq d$. Since the minimal element $\omega(x^d)$ of $\mathbb{V}^d$ has all entries 0 or 1, it is also the minimal element of $\mathbb{V}^d(n)$ for all $n \geq 1$. The minimal element $\omega^{\min}(d)$ has first entry $\mu(d)$, and so it is also the minimal element of $\mathbb{W}^d(n)$ for all $n \geq \mu(d)$. If $n < \mu(d)$, then $\mathbb{W}^d(n) = \emptyset$. The statement and proof of Proposition 5.3.2 are valid for $\mathbb{V}^d(n)$ for all n. In particular, $\mathbb{V}^d(n)$ is graded by $|V|$ for $V \in \mathbb{V}^d(n)$. Note that in $\mathbb{V}^d$ an element of $\mathbb{V}^d(n)$ can cover elements which are not in $\mathbb{V}^d(n)$. For example, $(2,2)$ covers $(0,3)$ in $\mathbb{V}^6$. On the other hand, since $\mathbb{W}^d(n)$ is the subset of vectors $V \in \mathbb{W}^d$ with $v_1 \leq d$, if $V \succeq W$ in $\mathbb{W}^d$ and if $V \in \mathbb{W}^d(n)$, then $W \in \mathbb{W}^d(n)$.

Proposition 5.6.1. *For all $n \geq 1$, $\mathbb{V}^d(n)$ is a sublattice of $\mathbb{V}^d$.*

Proof. Let $V, W \in \mathbb{V}^d(n)$ have partial degree vectors $\widehat{V}, \widehat{W}$, i.e. $\widehat{v}_k = \sum_{j=1}^k 2^{j-1} v_j$ and $\widehat{w}_k = \sum_{j=1}^k 2^{j-1} w_j$ for $k \geq 1$. Since $V, W \in \mathbb{V}^d(n)$, $\widehat{v}_k - \widehat{v}_{k-1} \leq 2^{k-1} n$ and $\widehat{w}_k - \widehat{w}_{k-1} \leq 2^{k-1} n$ for $k \geq 1$. To show that $V \vee W$ and $V \wedge W$ are also in $\mathbb{V}^d(n)$, we require $a_k - a_{k-1} \leq 2^{k-1} n$ and $b_k - b_{k-1} \leq 2^{k-1} n$, where $a_k = \max(\widehat{v}_k, \widehat{w}_k)$ and $b_k = \min(\widehat{v}_k, \widehat{w}_k)$.

By exchanging V and W if necessary, we may assume that $\widehat{v}_k \geq \widehat{w}_k$. If also $\widehat{v}_{k-1} \geq \widehat{w}_{k-1}$, then the result holds since $a_k - a_{k-1} = \widehat{v}_k - \widehat{v}_{k-1}$ and $b_k - b_{k-1} = \widehat{w}_k - \widehat{w}_{k-1}$. If $\widehat{v}_{k-1} \leq \widehat{w}_{k-1}$, then $a_k - a_{k-1} = \widehat{v}_k - \widehat{w}_{k-1} \leq \widehat{v}_k - \widehat{v}_{k-1}$ and $b_k - b_{k-1} = \widehat{w}_k - \widehat{v}_{k-1} \leq \widehat{v}_k - \widehat{v}_{k-1}$, so the result holds in this case also. □

By Proposition 5.3.8, the lattice $\mathbb{V}^d(n)$ is distributive. Since $\mathbb{V}^d(n)$ is a finite lattice, it has a unique maximal element $\omega^{\max}(n,d)$. The following algorithm gives this maximal element. Let ξ be the function defined by

$$\xi(n,d) = \begin{cases} d, & \text{if } 0 \le d \le n, \\ n, & \text{if } d > n \text{ and } d \equiv n \bmod 2, \\ n-1, & \text{if } d > n \text{ and } d \not\equiv n \bmod 2. \end{cases}$$

Proposition 5.6.2. *The maximal element $T = \omega^{\max}(n,d)$ of $\mathbb{V}^d(n)$ is the vector T defined recursively by $d_1 = d$ and $t_i = \xi(n,d_i)$, $d_{i+1} = (d_i - t_i)/2$ for $i \ge 1$.*

Proof. It is clear from the construction that T cannot arise from a binary addition move on any vector V with $v_i \le n$ for all i. Hence T is the maximal element of $\mathbb{V}^d(n)$. □

Example 5.6.3. With $n = 3$ and $d = 12$ we obtain $T = (2,3,1)$.

We next consider the poset $\mathbb{W}^d(n)$.

Proposition 5.6.4. *Let $n \ge \mu(d)$. The vector V constructed recursively by $d_1 = d$, $v_0 = n$ and $v_i = \xi(v_{i-1}, d_i)$, $d_{i+1} = (d_i - v_i)/2$ for $i \ge 1$ is the unique maximal element of $\mathbb{W}^d(n)$. It is characterized by $v_{i-1} - v_i = 0$ or 1 for $1 \le i < l$, where $v_0 = n$ and $l = \mathrm{len}(V)$.*

Proof. By definition of the function ξ we see that $v_{i-1} - v_i = 0$ or 1 for $1 \le i < l$. Hence the vector V is decreasing, i.e. $V \in \mathbb{W}^d(n)$. Let $V' \in \mathbb{W}^d(n)$ and let $k = \mathrm{len}(V')$. If $V' \ne V$ then $v'_{i-1} \ge v'_i + 2$ for some i with $1 \le i < k$, where $v'_0 = n$. Since $i < k$, $v'_{i+1} \ne 0$. Let V'' be the vector obtained from V' by replacing the entries (v'_i, v'_{i+1}) by $(v'_i + 2, v'_{i+1} - 1)$. If $v'_{i+1} > v'_{i+2}$, then V'' is decreasing and $V'' \succ V'$. If $v'_{i+1} = v'_{i+2}$, let V''' be the vector obtained from V'' by replacing the entries $(v'_{i+1} - 1, v'_{i+2})$ by $(v'_{i+1} + 1, v'_{i+2} - 1)$. If $v'_{i+2} > v'_{i+3}$, then V''' is decreasing and $V'' \succ V'$. Continuing this process, since $v'_j = 0$ for $j > k$ we can eventually construct $W \in \mathbb{W}^d(n)$ such that $W \succ V'$. Hence V is the unique maximal element of $\mathbb{W}^d(n)$. □

Proposition 5.6.5. *For all $n \ge \mu(d)$, $\mathbb{W}^d(n)$ is a sublattice of $\mathbb{W}^d$.*

Proof. If $U = V \wedge W$ where $V, W \in \mathbb{W}^d(n)$, then $U \in \mathbb{W}^d$ by Proposition 5.4.5, and Since $u_1 = \min(v_1, w_1)$ and $v_1, w_1 \le n$, $u_1 \le n$ and so $U \in \mathbb{W}^d(n)$. By Proposition 5.6.4 $\mathbb{W}^d(n)$ has a unique maximal element, and so the least upper bound $\inf\{U \mid V, W \preceq_2 U\}$ is well-defined. Hence $\mathbb{W}^d(n)$ is a sublattice of $\mathbb{W}^d$. □

Definition 5.6.6. For $n \ge \mu(d)$, we denote the maximal element V of $\mathbb{W}^d(n)$ by $\omega^{\max}(n,d)$ and its conjugate by $\alpha^{\max}(n,d)$. A spike $f \in P^d(n)$ is a **maximal spike** if $\omega(f) = V$. The **maximal Ferrers block** $F^{\max}(n,d)$ for spikes in $P^d(n)$ is the Ferrers block corresponding to the maximal spike f whose exponents are in decreasing order.

Example 5.6.7. Let $d = 21$. Then if $n = 4$ we obtain $\omega^{\max}(n, d) = (3, 3, 3)$, whereas $\omega^{\min}(d) = (3, 3, 1, 1)$, as shown in Example 5.5.5. For $n = 5$, we obtain $\omega^{\max}(n, d) = (5, 4, 2)$.

The following result will be useful later in finding upper bounds for $Q^d(n)$.

Proposition 5.6.8. *Let $n \geq \mu(d)$ and let V be the maximum element of $\mathbb{W}^d(n)$. Let $W >_l V$ where $W \in \mathbb{V}^d(n)$, and let k be defined by $w_k > v_k$ and $w_j = v_j$ for $j < k$. Then $w_k > v_{k-1}$.*

Proof. Since W and V have the same degree, we have $k < \text{len}(V)$. There is nothing to prove if $d \leq n$, since V is then maximal in $\mathbb{V}^d(n)$. Thus we may assume that $v_1 = n$ or $n - 1$ by Proposition 5.6.4, and since $w_1 \equiv d \bmod 2$ we may also assume $k > 1$. Again by Proposition 5.6.4, $v_{k-1} \leq v_k + 1$. By Proposition 5.3.2**(i)**, $w_k > v_k + 1$. Hence $w_k > v_{k-1}$ as required. □

We now consider conditions under which $\omega^{\max}(n, d) = \omega^{\min}(d)$, so that all spikes in $P^d(n)$ have the same ω-vector. A spike in $P^d(n)$ with strictly decreasing α-vector is the unique decreasing spike in $P^d(n)$, since it satisfies the criteria for being both the maximum and minimum in Propositions 5.5.6 and 5.6.4. A necessary condition for this is that $\mu(d) = n$ or $n - 1$. For $n = 1$ or 2 and all $d \geq 0$, there is at most one element in $\mathbb{W}^d(n)$. For $n \geq 5$ and $d > 2$ we see from Propositions 5.5.2 and 5.6.4 that if there is only one spike f in $P^d(n)$ with $\alpha(f)$ decreasing, then $\alpha(f)$ is strictly decreasing. This leads to the following result.

Proposition 5.6.9. *For $n \geq 5$ and $d > 2$, $\mathbb{W}^d(n)$ and $\mathbb{S}^d(n)$ have exactly one element if and only if $\alpha(d + n) = n$.*

Proof. Since $n \geq 5$ and $d > 2$, we may assume from the above that all spikes in $P^d(n)$ have the same ω-vector if and only if $d = \sum_{i=1}^{n}(2^{t_i} - 1)$, where $(t_1, \ldots, t_n)$ is strictly decreasing. Then $\text{bin}(d+n) = \{2^{t_1}, \ldots, 2^{t_n}\}$, and so $\alpha(d+n) = n$. □

Example 5.6.10. The cases $n = 3$ or 4 exhibit some exceptions to the condition $\alpha(d + n) = n$. For $n = 3$ and $d > 2$, $\mathbb{W}^d(3)$ has exactly one element not only when $\alpha(d + 3) = 3$, but also when $d = 2^k - 2$ or $2^k - 3$ and $k \geq 3$. For $n = 4$ and $d > 2$, $\mathbb{W}^d(4)$ has exactly one element not only when $\alpha(d + 4) = 4$, but also when $d = 2^k - 3$ and $k \geq 3$. In these cases the elements of $\mathbb{W}^d(n)$ are the vectors $(2, \ldots, 2)$ or $(3, \ldots, 3, 1)$ of length $k - 1$, and $\alpha(d+n) = 1$ or 2. It is easy to check that $\omega = \omega^{\min}(d) = \omega^{\max}(d)$ using Propositions 5.5.2 and 5.6.4. The corresponding Ferrers blocks are

$$\begin{matrix} 1-1 & 1-1 & 1 \\ 1-1\ , & 1-1 & \\ 0-0 & 1-1 & \end{matrix} \quad \text{and} \quad \begin{matrix} 1-1 & 1 \\ 1-1 & \\ 1-1 & \\ 0-0 & \end{matrix} .$$

5.7 Dominance in the Steenrod algebra

Recall (Definition 3.1.6) that a vector $A = (a_1, \ldots, a_s)$ of length s is **admissible** if $s = 1$ or if $s > 1$ and $a_i \geq 2a_{i+1}$ for $1 \leq i \leq s-1$. The modulus $|A| = \sum_{i=1}^{s} a_i$ is the degree of the corresponding admissible monomial $Sq^A = Sq^{a_1} \cdots Sq^{a_s}$ of the Steenrod algebra $\mathcal{A}_2$. We denote by $\mathbb{A}^d$ the poset of admissible vectors of modulus d, ordered by dominance.

Proposition 5.7.1. *For $d \geq 0$, the poset $(\mathbb{A}^d, \preceq)$ of admissible vectors of modulus d is isomorphic to the poset $(\mathbb{W}^d, \preceq_2)$ of decreasing vectors of degree d.*

Proof. We define inverse bijections between $\mathbb{A}^d$ and $\mathbb{W}^d$, so that $V = (v_1, \ldots, v_s) \in \mathbb{W}^d$ corresponds to $A = (a_1, \ldots, a_s) \in \mathbb{A}^d$ where for $1 \leq i \leq s$

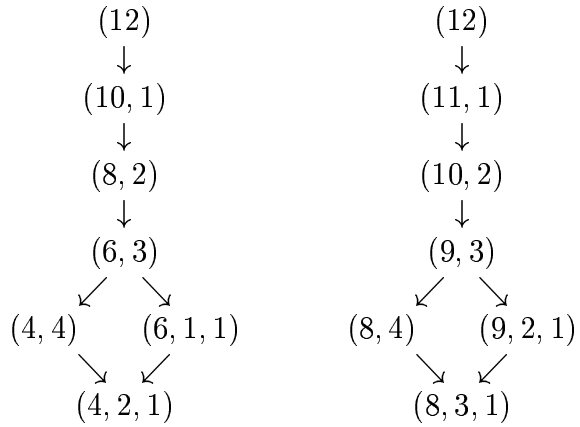
$$a_i = v_i + v_{i+1} + 2v_{i+2} + \cdots + 2^{s-i-1}v_s, \tag{5.2}$$

$$v_i = a_i - (a_{i+1} + \cdots + a_s). \tag{5.3}$$

Then $a_i - 2a_{i+1} = (v_i + v_{i+1} + 2v_{i+2} + \cdots + 2^{s-i-1}v_s) - 2(v_{i+1} + v_{i+2} + 2v_{i+3} + \cdots + 2^{s-i-2}v_s) = v_i - v_{i+1}$, and so $a_i \geq 2a_{i+1}$ if and only if $v_i \geq v_{i+1}$. Hence $A \in \mathbb{A}^d$ if and only if $V \in \mathbb{W}^d$. It is easily checked that the maps $A \mapsto V$ and $V \mapsto A$ are inverses of each other and that $|A| = \deg(V)$.

We must show that $V \succeq_2 V'$ in $\mathbb{W}^d$ if and only if $A \succeq A'$ in $\mathbb{A}^d$, i.e. $\widehat{v_i} \geq \widehat{v'}_i$ for $1 \leq i \leq s$ if and only if $\sum_{j=1}^{i} a_j \geq \sum_{j=1}^{i} a'_j$ for $1 \leq i \leq s$, where $\widehat{v_i} = \sum_{j=1}^{i} 2^{j-1}v_j$. We show that in fact for each i the ith inequality for $V \succeq_2 V'$ is equivalent to the ith inequality for $A \succeq A'$. Thus $\widehat{v_i} \geq \widehat{v'}_i$ if and only if $\sum_{j=1}^{i} 2^{j-1}(a_j - a_{j+1} - \cdots - a_s) \geq \sum_{j=1}^{i} 2^{j-1}(a'_j - a'_{j+1} - \cdots - a'_s)$. Adding the equation $(2^i - 1)(a_1 + \cdots + a_s) = (2^i - 1)(a'_1 + \cdots + a'_s)$ to this inequality and dividing by the common factor 2^i, we obtain $\sum_{j=1}^{i} a_j \geq \sum_{j=1}^{i} a'_j$. □

Example 5.7.2. The posets $\mathbb{W}^{12}$ and $\mathbb{A}^{12}$ are shown below.



By Proposition 5.4.2, it follows that $(\mathbb{A}^d, \preceq)$ is also isomorphic to $(\mathbb{S}^d, \succeq)$. To make this explicit, recall that $V = (v_1, \ldots, v_s) \in \mathbb{W}^d$ corresponds to the spike

partition of d with $v_j - v_{j+1}$ parts equal to $2^j - 1$ for $1 \leq j \leq s$, where $v_{s+1} = 0$. Now $a_i = v_i + v_{i+1} + 2v_{i+2} + \cdots + 2^{s-i-1}v_s, = (v_i - v_{i+1}) + 2(v_{i+1} - v_{i+2}) + \cdots + 2^{s-i-1}(v_{s-1} - v_s) + 2^{s-i}v_s$. This gives a practical way to write down corresponding elements of $\mathbb{S}^d$, $\mathbb{W}^d$ and $\mathbb{A}^d$ using Ferrers blocks.

Proposition 5.7.3. *Let the spike $f \in P^d(n)$ have exponent vector $S(f) = (2^{\alpha_1} - 1, \ldots, 2^{\alpha_n} - 1) \in \mathbb{S}^d$, and let F be the corresponding Ferrers block, with row sum vector $\alpha(f) = (\alpha_1, \ldots, \alpha_n)$ and column sum vector $\omega(f) = \alpha(f)^t \in \mathbb{W}^d$. Then the admissible vector $A(f) \in \mathbb{A}^d$ corresponding to $S(f)$ and $\omega(f)$ is the column sum vector of the array given by replacing the ith row of F by $(2^{\alpha_i - 1}, \ldots, 2, 1)$ if $\alpha_i > 0$.* □

Example 5.7.4. Let F be the Ferrers block of Example 5.1.8, so that $f = x_1^{15}x_2^3x_3^3$ in $P^{21}(4)$. Then $S(f) = (15, 3, 3)$, $\alpha(f) = (4, 2, 2)$, $\omega(f) = (3, 3, 1, 1)$ and $A(f) = (12, 6, 2, 1)$ is the column sum vector of the array

$$\begin{array}{cccc} 8 & 4 & 2 & 1 \\ 2 & 1 & & \\ 2 & 1 & & \\ 0 & & & \end{array}.$$

By applying this to the minimum element $\omega^{\min}(d) \in \mathbb{W}^d$, we obtain an algorithm for finding the minimum admissible vector in $\mathbb{A}^d$.

Proposition 5.7.5. *For $d \geq 0$, let $\beta(d) = (d + \mu(d))/2$, and define the vector $A = (a_1, a_2, \ldots)$ recursively by $d_1 = d$ and, for $i \geq 1$, $a_i = \beta(d_i)$ and $d_{i+1} = d_i - a_i$. Then $A = A^{\min}(d)$ is the minimum element of $\mathbb{A}^d$, and Sq^A is the minimal admissible monomial in $\mathcal{A}_2^d$ in both the left and the right orders.*

Proof. Let $d = \sum_{i=1}^{\mu(d)}(2^{\alpha_i} - 1)$ be the minimal spike partition of d. Then $\beta(d) = (d + \mu(d))/2 = \sum_{i=1}^{\mu(d)} 2^{\alpha_i - 1}$ is the first column sum of the array for $A^{\min}(d)$ given by Proposition 5.7.3, and so its first entry $a_1 = \beta(d)$. The result follows by induction on d. The last statement follows from Proposition 5.2.3. □

Example 5.7.6. Let $d = 42$. The 'greedy algorithm' of Proposition 5.5.2 gives $d = 31 + 7 + 3 + 1$, so $\mu(d) = 4$ and $\beta(d) = 23$. Then $d_2 = 19 = 15 + 3 + 1$, so $\mu(d_2) = 3$ and $\beta(d_2) = 11$. Continuing in this way, we obtain $A = (23, 11, 5, 2, 1)$. In practice it is quicker to calculate $\omega^{\min}(d) = (4, 3, 2, 1, 1)$ and $A^{\min}(d)$ from the Ferrers block of the spike for $n = \mu(d)$, as in Example 5.7.4.

The isomorphism between the posets $(\mathbb{A}^d, \preceq)$ and $(\mathbb{S}^d, \succeq)$ is easily described using Milnor vectors. Recall from Definition 3.3.10 that the Milnor vector of the admissible vector $A = (a_1, \ldots, a_s)$ is the vector $R = (r_1, \ldots, r_s)$, where $r_i = a_i - 2a_{i+1}$. Since equation (5.2) or (5.3) gives $a_i - 2a_{i+1} = v_i - v_{i+1}$, we can write $R = (v_1 - v_2, \ldots, v_{s-1} - v_s, v_s)$ in terms of the element $V \in \mathbb{W}^d$ corresponding

to A. Reversing this vector gives the exponent vector $(v_s, v_{s-1} - v_s, \ldots, v_1 - v_2)$ of the spike partition $S = ((2^s - 1)^{v_s} \cdots 3^{v_2 - v_3} 1^{v_1 - v_2})$ corresponding to A. Thus $|R| = \mathrm{len}(S) = v_1$. For example, the admissible vector $A = (8, 3, 1)$ has Milnor vector $R = (2, 1, 1)$ and corresponds to $S = (731^2)$ and $V = (4, 2, 1)$.

It is convenient to regard the set of Milnor vectors corresponding to $\mathbb{A}^d$ or to $\mathbb{S}^d$ as a poset isomorphic to these.

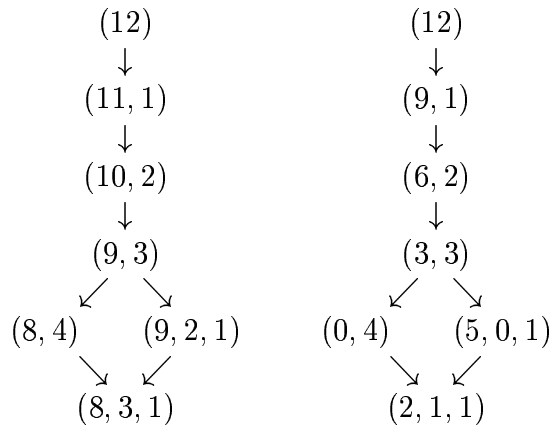
Definition 5.7.7. For $d \geq 0$, let $\mathbb{R}^d$ be the poset of vectors R indexing the Milnor basis of $\mathcal{A}_2^d$, with the partial order relation $R \geq R'$ if and only if $S \preceq S'$, where S and S' are the spike partitions of d corresponding to R and R' respectively.

The partial order on $\mathbb{R}^d$ can be defined more directly as follows. Recall that if $R \in \mathbb{R}^d$ then $\deg(Sq(R)) = d = \sum_{j \geq 1}(2^j - 1)r_j$.

Proposition 5.7.8. *Let $R = (r_1, \ldots, r_k)$ and $S = (s_1, \ldots, s_k)$ be vectors indexing Milnor basis elements in $\mathcal{A}_2^d$, where $k = \max(\mathrm{len}(R), \mathrm{len}(S))$. Then $R \geq S$ in $\mathbb{R}^d$ if and only if $\deg(Sq(r_j, \ldots, r_k)) \leq \deg(Sq(s_j, \ldots, s_k))$ for $1 < j \leq k$.*

Proof. Let A and $B \in \mathbb{A}^d$ be the admissible vectors corresponding to R and S respectively, so that by definition $R \geq S$ if and only if $A \succeq B$, i.e. $a_1 + \cdots + a_j \geq b_1 + \cdots + b_j$ for $1 \leq j \leq k$. Since $|A| = |B| = d$, equality holds for $j = 1$, and so these inequalities are equivalent to $a_j + \cdots + a_k \leq b_j + \cdots + b_k$ for $1 < j \leq k$. The result follows since $a_j + \cdots + a_k = \deg(Sq^{(a_j, \ldots, a_k)}) = \deg(Sq(r_j, \ldots, r_k))$. □

Example 5.7.9. The posets $\mathbb{A}^{12}$ and $\mathbb{R}^{12}$ are shown below.



If $(\mathbb{X}, \geq)$ is a poset, then a function ε from $\mathbb{X}$ to the integers ≥ 0 is **increasing** if $X \geq Y$ implies $\varepsilon(X) \geq \varepsilon(Y)$ for all $X, Y \in \mathbb{X}$.

Proposition 5.7.10. *If the vector $V \in \mathbb{W}^d$, the spike partition $S \in \mathbb{S}^d$, the admissible vector $A \in \mathbb{A}^d$ and the Milnor vector $R \in \mathbb{R}^d$ are corresponding elements of these posets, then $v_1 = \mathrm{len}(S) = 2a_1 - d = |R|$, and this function is increasing in each case.*

Proof. It is clear that these functions on the four posets correspond under the bijections defined above. To check that the functions are increasing, we need only check any one of the four cases. In the cases of $\mathbb{W}^d$ and $\mathbb{A}^d$, this is immediate by definition of the partial order. □

In the next section, we shall see that this function on $\mathbb{A}^d$ and $\mathbb{R}^d$ has a natural interpretation in terms of the action of the corresponding elements of $\mathcal{A}_2$ on $P(n)$.

5.8 The excess function

In this section we introduce the important **excess** function on $\mathcal{A}_2$, and determine it for Milnor basis elements and admissible monomials.

Definition 5.8.1. The **excess** $\mathrm{ex}(\theta)$ of a nonzero element θ of $\mathcal{A}_2$ is the minimum $n \geq 0$ such that $\theta(c(n)) \neq 0$, where $c(n) = x_1 \cdots x_n$ for $n \geq 1$ and $c(0) = 1$.

Proposition 5.8.2. *Let $\theta \in \mathcal{A}_2^+$ and let $f \in P^d(n)$ where $d < \mathrm{ex}(\theta)$. Then $\theta(f) = 0$.*

Proof. Recall from Proposition 1.2.3 that θ commutes with all linear substitutions in $M(n)$, and in particular with specializations of the variables. The result follows since any monomial of degree $d > 0$ can be obtained from the product $c(d)$ by such a specialization. □

Proposition 5.8.3. *The Milnor basis element $Sq(R)$ has excess $|R|$.*

Proof. Let $R = (r_1, r_2, \ldots) \in \mathbb{V}$. By Theorem 3.4.1 $Sq(R)c(n) = c(R^+)$ when n is sufficiently large, and in particular for $n \geq \sum_{i\geq 1}(2^i - 1)r_i$, where $R^+ = (n - |R|, r_1, r_2, \ldots)$.

By Proposition 1.3.5, $Sq(R)$ commutes with the partial differentiation operator $\partial/\partial x_i$ on $P(n)$ for $1 \leq i \leq n$. If all exponents in a monomial $f \in P(n)$ are 2-powers, then $\partial f/\partial x_i = f/x_i$ if the exponent of x_i in f is 1, and $\partial f/\partial x_i = 0$ otherwise. Hence $\partial c(R^+)/\partial x_n = c((n - |R| - 1, r_1, r_2, \ldots)$ if $n - 1 \geq |R|$, and otherwise $\partial c(R^+)/\partial x_n = 0$. However $\partial c(n)/\partial x_n = c(n-1)$ for all $n \geq 1$. It follows that $Sq(R)c(n) = c(R^+)$ for all $n \geq |R|$, and $Sq(R)c(n) = 0$ when $n < |R|$. □

Our next objective is to determine the excess of an admissible monomial $Sq^A \in \mathcal{A}_2^d$. To do this, we express Sq^A in the Milnor basis. First we relate the Milnor product formula to the partial order on $\mathbb{R}^d$.

Proposition 5.8.4. *Let $S = (s_1, \ldots, s_k) \in \mathbb{V}$, let $r \geq \sum_{j=1}^k 2^j s_j$, and let $S^+ = (s_0, s_1, \ldots, s_k)$ where $s_0 = r - \sum_{j=1}^k 2^j s_j$. Then $Sq^r Sq(S) = Sq(S^+) + \sum_{R^+} Sq(R^+)$ in $\mathcal{A}_2^d$, where all terms in the sum satisfy $R^+ > S^+$ in $\mathbb{R}^d$.*

Proof. We use Milnor's product formula 4.1.2 to evaluate $Sq^r Sq(S)$. Since $r \geq \sum_{j=1}^{k} 2^j s_j$, the Milnor matrix

$$X = \begin{array}{c|cccc} & 0 & 0 & \cdots & 0 \\ \hline s_0 & s_1 & s_2 & \cdots & s_k \end{array}$$

gives the term $Sq(S^+)$ in the product. A general Milnor matrix has the form

$$Y = \begin{array}{c|cccc} & x_1 & x_2 & \cdots & x_k \\ \hline y_0 & y_1 & y_2 & \cdots & y_k \end{array}$$

where $x_i + y_i = s_i$ for $1 \leq i \leq k$, and Y gives a term $b(Y)Sq(R^+)$, where the coefficient $b(Y) \in \mathbb{F}_2$ and $R^+ = (r_0, r_1, \ldots, r_k)$, with $r_i = x_{i+1} + y_i$ ($x_{k+1} = 0$) for $0 \leq i \leq k$. By Proposition 5.7.8, $R^+ \geq S^+$ in $\mathbb{R}^d$ if and only if $\deg(Sq(r_j, \ldots, r_k)) \leq \deg(Sq(s_j, \ldots, s_k))$ for $1 \leq j \leq k$. Then

$$\begin{aligned} \deg(Sq(r_j, \ldots, r_k)) &= \deg(Sq(x_{j+1}, \ldots, x_k)) + \deg(Sq(y_j, \ldots, y_k)), \\ \deg(Sq(s_j, \ldots, s_k)) &= \deg(Sq(x_j, \ldots, x_k)) + \deg(Sq(y_j, \ldots, y_k)), \end{aligned}$$

and $\deg(Sq(x_{j+1}, \ldots, x_k)) \leq \deg(Sq(x_j, \ldots, x_k))$, with equality if and only if $x_i = 0$ for $i \leq j \leq k$. Hence $R^+ > S^+$ if $Y \neq X$. □

This result allows us to strengthen Proposition 3.4.1 as follows.

Proposition 5.8.5. *Let A be an admissible vector with corresponding Milnor vector R. Then $Sq^A = Sq(R) + \sum_U Sq(U)$, where $U > R$ for all terms in the sum.*

Proof. We argue by induction on $\mathrm{len}(A)$. Since $Sq^a = Sq(a)$ for all $a \geq 0$, the statement is true if $\mathrm{len}(A) = 0$ or 1. Thus let $A = (a_1, a_2, \ldots, a_k)$ where $a_k > 0$ and $k \geq 2$. Let $B = (a_2, \ldots, a_k)$, and let $R = (r_1, r_2, \ldots, r_k)$ and $S = (r_2, \ldots, r_k)$ be the Milnor vectors of A and B. Then by the induction hypothesis $Sq^B = Sq(S) + \sum_T Sq(T)$, where $T > S$ for all terms in the sum. Then $Sq^A = Sq^{a_1} Sq^B = Sq^{a_1} Sq(S) + \sum_T Sq^{a_1} Sq(T)$.

By Proposition 5.8.4, $Sq^{a_1} Sq(S) = Sq(R) + \sum_U Sq(U)$, where $U > R$ for all terms in the sum. Thus let $T > S$, where $T = (t_2, \ldots, t_k)$ and let

$$Y = \begin{array}{c|ccc} & x_2 & \cdots & x_k \\ \hline y_1 & y_2 & \cdots & y_k \end{array}$$

be a Milnor matrix yielding a term Sq^U in the product $Sq^{a_1} Sq^T$, where $U = (u_1, \ldots, u_k)$. Thus $x_j + y_j = t_j$ for $2 \leq j \leq k$, and $u_j = x_{j+1} + y_j$ for $1 \leq j \leq k$, where $x_{k+1} = 0$. To complete the inductive step, we must prove that $U > R$.

Since $T > S$, $\deg(Sq(r_j, \ldots, r_k)) \geq \deg(Sq(t_j, \ldots, t_k))$ for $2 \leq j \leq k$, where at least one of these inequalities is strict. Then for $2 \leq j \leq k$,

$$\begin{aligned} \deg(Sq(t_j, \ldots, t_k)) &= \deg(Sq(x_j, \ldots, x_k)) + \deg(Sq(y_j, \ldots, y_k)) \\ &\geq \deg(Sq(x_{j+1}, \ldots, x_k)) + \deg(Sq(y_j, \ldots, y_k)) \\ &= \deg(Sq(u_j, \ldots, u_k)). \end{aligned}$$

Hence $\deg(Sq(r_j, \ldots, r_k)) \geq \deg(Sq(u_j, \ldots, u_k))$ for $2 \leq j \leq k$, where at least one of these inequalities is strict. By Proposition 5.7.8, it follows that $U > R$. □

Proposition 5.8.6. *The admissible monomial $Sq^A \in \mathcal{A}_2^d$ has excess $2a_1 - d$.*

Proof. By Proposition 5.7.10, the function $|R|$ is increasing on the poset $\mathbb{R}^d$. By definition of excess, the excess of an element $\theta \in \mathcal{A}_2^d$ expressed in the Milnor basis is the minimum excess of the terms $Sq(R)$ appearing in the sum. By Proposition 5.8.5, it follows that the excess of the admissible monomial Sq^A is $|R|$, where R is the Milnor vector corresponding to A. By Proposition 5.7.10, $|R| = 2a_1 - d$. □

The next result follows from the two preceding ones and Proposition 5.7.10.

Proposition 5.8.7. *The correspondence between admissible monomials Sq^A and Milnor basis elements $Sq(R)$ preserves the excess.* □

Proposition 5.8.8. *For $d \geq 0$, $\mu(d)$ is the minimum excess of an admissible monomial in $\mathcal{A}_2^d$. In particular, the minimal element $A^{\min}(d)$ of $\mathbb{A}^d$ has excess $\mu(d)$.*

Proof. Since $\mu(d)$ is the minimum length of an element of $\mathbb{S}^d$, this follows from Proposition 5.7.10 and the isomorphism between $\mathbb{A}^d$ and $\mathbb{S}^d$. □

Proposition 5.8.9. *For $d \geq 0$, the excess of $Xq^d = \chi(Sq^d)$ is $\mu(d)$.*

Proof. By Proposition 3.4.5 Xq^d is the sum of all Milnor basis elements in $\mathcal{A}_2^d$. Hence $\mathrm{ex}(Xq^d)$ is the minimum excess of a Milnor basis element in $\mathcal{A}_2^d$. By Proposition 5.7.10, this is the minimum length $\mu(d)$ of an element of $\mathbb{S}^d$. □

5.9 Remarks

For general information about partially ordered sets and lattices we refer to [138, Chapter 3]. There is a substantial literature on binary partitions (see e.g. [84]), in contrast to spike partitions [15, 52]. In particular, the cardinality of the set of vectors $\mathbb{V}^d$, equivalently the set of ω-vectors of monomials in P^d, is also the cardinality of $\mathbb{B}^d$, namely the number $b(d)$ of binary partitions of d, and can be computed recursively from the relations $b(0) = 1$ and $b(2k+1) = b(2k) = b(2k-1) + b(k)$ for $k \geq 0$, or from the generating function $\sum_{d=0}^{\infty} b(d)x^d = \prod_{j=0}^{\infty} 1/(1 - x^{2^j})$. For

$n \geq 1$ the corresponding generating function for the bounded case $\mathbb{V}^d(n) \equiv \mathbb{B}^d(n)$ is $\sum_{d=0}^{\infty} b(n,d)x^d = \prod_{j=0}^{\infty}(1 - x^{2^j(n+1)})/(1 - x^{2^j})$. Similarly the generating function for the cardinality of the set of decreasing vectors $\mathbb{W}^d$, equivalently the spike partitions $\mathbb{S}^d$, is $\sum_{d=0}^{\infty} s(d)x^d = \prod_{j=1}^{\infty} 1/(1-x^{2^j-1})$. In the bounded case the cardinality of $\mathbb{W}^d(n) \equiv \mathbb{S}^d(n)$ is given by the coefficient of $x^d y^n$ in $\prod_{j=0}^{\infty} 1/(1-x^{2^j-1}y)$.

The ω-vector is known under different names in the theory of games. Appendix C explains the winning strategy in the game of NIM in terms of the ω-vector of a block.

Proposition 5.8.4 strengthens a result of K. G. Monks [104] which proves the corresponding result for the right order $<_r$.

Chapter 6

Filtrations on $Q^d(n)$

6.0 Introduction

From Chapter 3 we may interpret the hit polynomials $H(n)$ of Chapter 1 as $\mathcal{A}_2^+P(n)$, where $\mathcal{A}_2^+ = \sum_{k>0}\mathcal{A}_2^k$. Recall from Chapter 1 that a vector space basis for the 'cohits' $Q(n) = P(n)/H(n)$ lifts to a minimal set of polynomials which generate $P(n)$ as an $\mathcal{A}_2$-module.

In Section 6.1 we gather a number of items which explain, in terms of the combinatorics of blocks, how Steenrod squares and matrices act on monomials. We recall that an n-block is a matrix with n infinite rows having entries the integers 0 or 1 and only finitely many entries 1. The n-blocks F are in bijective correspondence with monomials f in $P(n)$. The α-vector $\alpha(F)$ and the ω-vector $\omega(F)$ are the row sum and column sum vectors of F respectively, consistent with the notation $\alpha(f)$ in Definitions 3.3.2 and 3.3.6. The degree of F, or of the monomial f, is determined in terms of $W = \omega(F)$ by $\deg(F) = \sum_{i>0} 2^{i-1}w_i$. As before we write the ith entry of a vector V as v_i, and the (i,j)th entry of a block B as $b_{i,j}$. If $W = \omega(F)$, we write $w_i = \omega_i(F)$ to emphasize the functional dependence of w_i on F, and similarly for $\alpha(F)$.

In Section 6.2 we consider the left and right orders on the set $\mathbb{V}^d(n)$ of n-bounded vectors of degree d, and introduce the notion of **reducibility** of polynomials with respect to each of them. In each case, this leads to a filtration of $Q^d(n)$ whose associated graded spaces $Q^W(n)$ give a vector space decomposition $Q^d(n) \cong \bigoplus_W Q^W(n)$, where the sum is over $W \in \mathbb{V}^d(n)$. This reduces the 'global' problem of finding $\dim Q(n)$ to the 'local' problems of finding $\dim Q^W(n)$. The spaces $Q^W(n)$ inherit a matrix action in a natural way, but these decompositions are not as direct sums of $\mathbb{F}_2GL(n)$-modules.

In Section 6.3 we introduce **concatenation** of n-blocks. This requires a treatment of **finite blocks**. In Section 6.4 we explain the technique called **splicing** of a block which is a useful method of generating hit equations. We show that $Q^W(n) = 0$ if $W <_l \omega^{\min}(d)$. In Section 6.5 we generalize the Kameko and du-

plication maps introduced in Chapter 1. In general, it is hard to describe $Q^W(n)$ as a $GL(n)$-module, but there are two extreme cases of vectors W for which we can do this fairly easily. We call these **head**-vectors $(n-1,\dots,n-1)$ in degree $(2^t-1)(n-1)$ and **tail**-vectors $(1,\dots,1)$ in degree 2^t-1, where t is the length of the vector. In Section 6.6 we deal with head vectors, leaving tail vectors until later.

6.1 Steenrod and matrix actions on blocks

A polynomial f is represented by a formal sum of blocks $F = B_1 + \cdots + B_k$, where no two terms are equal, so that the sum is irredundant over $\mathbb{F}_2$. We write $Sq^k(F)$ for the sum of blocks representing $Sq(f)$.

Example 6.1.1. Let F and G be 1-blocks, where F has an initial section of contiguous digits 1 followed by a zero in the last position, and G is obtained by interchanging the digits 0 and 1 in F, i.e. $F = 1-1\ 0$, $G = 0-0\ 1$. Then F represents x^{2^a-1}, where $a = \alpha(F)$, G represents x^{2^a}, and $Sq^1(F) = G$ by Proposition 1.1.7. Viewing F and G as reversed binary expansions of 2^a-1 and 2^a we see that G is obtained by adding the number 1 arithmetically to F.

For a general 1-block F in degree d and $k > 0$, by Proposition 1.5.3 $Sq^k(F)$ is zero unless $\mathrm{bin}(k) \subseteq \mathrm{bin}(d)$. This leads to the following combinatorial description of how Sq^k acts on F.

Proposition 6.1.2. *Let F be a 1-block representing x^d, so that F is the reversed binary expansion of d. Then $Sq^k(F)$, for $k > 0$, is either zero or is the block $F(1)$ obtained from $F = F(t+1)$ by working through the elements $2^{j_1},\dots,2^{j_t}$ of $\mathrm{bin}(k)$ in decreasing order and successively forming the blocks $F(i)$, for $1 \le i \le t$, by adding 1 arithmetically at position j_i+1 to $F(i+1)$. Equivalently $Sq^k(F) = Sq^{2^{j_t}}\cdots Sq^{2^{j_1}}(F)$. Further, for every block B in the expansion of $Sq^k(F)$, $\omega(B) <_{l,r} \omega(F)$ and $\alpha(B) \le \alpha(F)$.*

Proof. The combinatorial process described above is simply a way of adding the numbers d and k in reversed binary arithmetic. The result is proved by induction on t using Propositions 1.1.7 and 1.5.3. Under the condition $\mathrm{bin}(k) \subseteq \mathrm{bin}(d)$, at each step some digit 1 in F is moved to the right, no digit 1 is moved left, and the total number of digits 1 does not increase. Thus $\omega(F(i)) < \omega(F(i+1))$ in both left and right order, and $\alpha(F(i)) \le \alpha(F(i+1))$. □

Example 6.1.3. Let $F = F(3) = 0\ 1\ 1\ 1\ 1\ 1$ and $k = 10$, so that $\mathrm{bin}(k) = \{2,8\} \subset \{2,4,8,16,32\}$. Starting in position 4, we form $F(2) = Sq^8(F) = 0\,1\,1\,0\,0\,0\,1$. Then at position 2 we form $F(1) = Sq^2(F(2)) = 0\,0\,0\,1\,0\,0\,1$, giving $Sq^{10}(F) = Sq^2Sq^8(F)$.

More generally, we can evaluate a Steenrod operation Sq^k on a block.

Example 6.1.4. The equation $Sq^2(x^3yz) = x^5yz + x^4y^2z + x^4yz^2 + x^3y^2z^2$ is illustrated in terms of blocks by

$$Sq^2 \begin{matrix} 1 & 1 \\ 1 & 0 \\ 1 & 0 \end{matrix} = \begin{matrix} 1 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{matrix} + \begin{matrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{matrix} + \begin{matrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{matrix} + \begin{matrix} 1 & 1 \\ 0 & 1 \\ 0 & 1 \end{matrix}.$$

Proposition 6.1.5. *Let $f \in P^d(n)$ be a monomial and let g be a monomial appearing in $\theta(f)$, where $\theta \in \mathcal{A}_2^+$. Then $\omega(g) <_{l,r} \omega(f)$, and $\alpha_i(g) \leq \alpha_i(f)$ for $1 \leq i \leq n$.*

Proof. It is enough to prove the result for $\theta = Sq^k$. Then the case $n = 1$ is treated in Proposition 6.1.2, and the general case follows from the extended Cartan formula 1.1.8. □

The above observations concerning addition in reversed binary arithmetic are also involved in explaining how matrix substitution works on blocks. We first collect some elementary results concerning reversed binary arithmetic, the usual component-wise addition of vectors and the left and right order relations. Clearly $V + W \geq_l V$ and $V + W \leq_r V$ for all $V, W \in \mathbb{V}$, with equality if and only if W is the zero vector.

Proposition 6.1.6. *We have $\omega(x^{r+s}) \leq_{l,r} \omega(x^r) + \omega(x^s)$, with equality if and only if* $\mathrm{bin}(r)$ *and* $\mathrm{bin}(s)$ *are disjoint. Further $\omega(x^{r+s}) \leq_r \omega(x^r)$, with equality if and only if $s = 0$.*

Proof. Since the entries of $\omega(x^r)$ are the digits 0 or 1 in the reversed binary expansion of r, $\omega(x^{r+s})$ is obtained from the binary addition of r and s. Consider the least i such that $\omega_i(x^r) = \omega_i(x^s) = 1$. If no such i exists, then $\mathrm{bin}(r)$ and $\mathrm{bin}(s)$ are disjoint and clearly $\omega(x^{r+s}) = \omega(x^r) + \omega(x^s)$. Otherwise $\omega_i(x^{r+s}) = 0$, whereas $\omega_i(x^r) + \omega_i(x^s) = 2$. On the other hand, $\omega_j(x^{r+s}) = \omega_j(x^r) + \omega_j(x^s)$ for $j < i$. Thus $\omega(x^{r+s}) <_l \omega(x^r) + \omega(x^s)$. A similar argument, working from the right, proves the result for the right order $<_r$. Finally $\omega(x^r) + \omega(x^s) \leq_r \omega(x^r)$, with equality if and only if $\omega(x^s) = 0$. □

There is no analogue of the last statement of Proposition 6.1.6 for the left order, as we see from the examples $\omega(x^2) <_l \omega(x)$ and $\omega(x^3) >_l \omega(x)$, but the following more general result holds for the right order.

Proposition 6.1.7. *For all monomials f, g we have $\omega(fg) \leq_r \omega(f)$, with equality if and only if $g = 1$.*

Proof. Let $f = x_1^{r_1} \cdots x_n^{r_n}$ and $g = x_1^{s_1} \cdots x_n^{s_n}$. Then by Proposition 6.1.6 $\omega(fg) = \sum_{i=1}^n \omega(x_i^{r_i+s_i}) \leq_r \sum_{i=1}^n \omega(x_i^{r_i}) = \omega(f)$, with equality if and only if all $s_i = 0$. □

Definition 6.1.8. The **standard transvection** $T \in GL(n)$ is given by $x_1 \cdot T = x_1 + x_2$ and $x_i \cdot T = x_i$ for $i > 1$.

In the 2-variable case, the transvection T was introduced in Example 1.2.2. We next explain how to expand $(x+y)^a y^b$ in terms of blocks, or equivalently how to evaluate the action of T on a block.

Proposition 6.1.9. *Let F denote the 2-block representing the monomial $x^a y^b$. For each subset $S \subseteq \mathrm{bin}(a)$ we construct a 2-block F^S in the following way. The first row of F is obtained by putting $f_{(1,j+1)} = 0$ for $2^j \in S$ and leaving other entries unchanged. Let the first row of F^S be $\omega(r)$. The second row of F^S is then defined to be $\omega(s)$, where $s = a+b-r$. For the standard transvection T we have $x^a y^b \cdot T = \sum_S F^S$, where the sum is taken over all S as constructed above. Furthermore $\omega(F^S) \leq_{l,r} \omega(F)$ for all S.*

Proof. The expansion of $(x+y)^a y^b$ is the sum of terms $\binom{a}{r} x^r y^s$ where $r+s = a+b$ and $\binom{a}{r}$ is odd. By Proposition 1.5.3, each value of r corresponds to a set S as described in the proposition. The block F^S then corresponds to $x^r y^s$. Now by definition $\omega(F^S) = \omega(x^r y^s) = \omega(x^r) + \omega(y^s) = \omega(x^r) + \omega(x^s)$. Since $s = a - r + b$, Proposition 6.1.6 shows that $\omega(F^S) \leq_{l,r} \omega(x^r) + \omega(x^{a-r}) + \omega(x^b)$. By construction of the set S, $\omega(x^r)$ and $\omega(x^{a-r})$ are disjoint. Hence $\omega(x^r) + \omega(x^{a-r}) = \omega(a)$ and $\omega(F^S) \leq_{l,r} \omega(x^a) + \omega(y^b) = \omega(F)$ as required. □

Example 6.1.10. Let

$$F = \begin{matrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{matrix}, \quad B = \begin{matrix} 0 & 0 & 1 \\ 0 & 0 & 1 \end{matrix}, \quad B' = \begin{matrix} 1 & 0 & 0 \\ 1 & 1 & 1 \end{matrix}, \quad B'' = \begin{matrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{matrix},$$

Then $F \cdot T = F + B + B' + B''$. With reference to Proposition 6.1.9, $d = 5$, $\mathrm{bin}(d) = \{1, 4\}$ and F, B, B', B'' correspond to the subsets $\{\}, \{1\}, \{4\}, \{1, 4\}$ of $\mathrm{bin}(d)$.

We can extend the last statement of Proposition 6.1.9 to an arbitrary matrix.

Proposition 6.1.11. *Let $M \in M(n, \mathbb{F}_2)$ be a matrix and let F be a n-block. Then $\omega(B) \leq_{l,r} \omega(F)$ for every block B in the expansion of $F \cdot M$.*

Proof. The general linear group $GL(n)$ is generated by permutation matrices and the standard transvection T. Clearly, permutations of the rows of a block do not alter the ω-vector. The result for $GL(n)$ then follows from Proposition 6.1.9. Now every $n \times n$ matrix of rank $k > 0$ for can be written in the form $M' E_k M$ where M', M are in $GL(n)$ and $E_k = \begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix}$, where I_k is the identity matrix of order k. Applying E_k either sends a monomial to zero or leaves it fixed. The result is then clear for E_k and the proof is completed by composition of the actions of M', E_k and M. □

6.2 Reducibility and the spaces $Q^W(n)$

The results of the previous section show that the left and right orders on ω-vectors are valuable tools for studying both the action of the Steenrod algebra and that of linear substitutions on the polynomial algebra $P(n)$. We use these orderings to extend the definition of ω-vector from monomials to polynomials.

Definition 6.2.1. Let $<$ denote either the left or right order on $\mathbb{V}(n)$, and let $f \in P^d(n)$. Then the **ω-vector** $\omega(f)$ is the maximum of the vectors $\omega(g)$ in $\mathbb{V}(n)$ for monomials g appearing in f.

It is important to note that this definition depends on the choice of linear order. In this chapter, the notation $<$ will denote either the left order $<_l$ or the right order $<_r$ on $\mathbb{V}(n)$.

Example 6.2.2. Let $f = x^5yz + x^3y^2z^2 \in P^7(3)$. Then $\omega(f) = \omega(x^5yz) = (3,0,1)$ for the left order , but $\omega(f) = \omega(x^3y^2z^2) = (1,3)$ for the right order.

Propositions 6.1.5 and 6.1.11 give the following result.

Proposition 6.2.3. *Let $f \in P(n)$ be a polynomial. Then $\omega(\theta(f)) < \omega(f)$ for any operation $\theta \in \mathcal{A}_2^+$, and $\omega(f \cdot M) \leq \omega(f)$ for any matrix $M \in M(n, \mathbb{F}_2)$.* □

Given an ω-vector $W \in \mathbb{V}^d(n)$, we denote by $P^W(n)$ the vector subspace of $P^d(n)$ spanned by monomials g with $\omega(g) = W$. For each linear ordering $<$ we extend this notation, and denote by $P^{<W}(n)$ and $P^{\leq W}(n)$ the subspaces of $P^d(n)$ spanned by monomials g with $\omega(g) < W$ and $\omega(g) \leq W$ respectively. Thus for a homogeneous polynomial $f \in P^d(n)$, $f \in P^{<W}(n)$ if and only if $\omega(f) < W$, and $f \in P^{\leq W}(n)$ if and only if $\omega(f) \leq W$. By Proposition 6.1.11, $P^{<_l W}(n)$ and $P^{<_r W}(n)$ are $\mathbb{F}_2 M(n, \mathbb{F}_2)$-submodules of $P^d(n)$.

Given a choice of (left or right) ordering $<$ and an ω-vector W, we have a 'relative' version of the hit problem, as follows.

Definition 6.2.4. A polynomial $f \in P^W(n)$ is (**left** or **right**) **reducible** if $f = h + e$ where h is hit and $\omega(e) < W$. We write $f \approx g$ if $f, g \in P^W(n)$ and $f - g$ is reducible. The set of reducibility equivalence classes in $P^W(n)$ is denoted by $Q^W(n)$.

We use $f \approx_l g$ and $f \approx_r g$ when referring to the left and right orders. We recall that $f \sim g$ means that $f - g$ is hit.

Example 6.2.5. The monomial $x^5y^5z^5$ is left reducible. To see this, consider the hit equation

$$Sq^2(x^3y^5z^5) = x^5y^5z^5 + x^4y^6z^5 + x^4y^5z^6 + x^3y^6z^6.$$

The last three monomials have ω-vectors $(1,1,3), (1,1,3), (1,3,2)$, which are left lower than $\omega(x^5y^5z^5) = (3,0,3)$. Note that $\omega(x^3y^6z^6) = (1,3,2)$ is not right lower than $(3,0,3)$.

Example 6.1.10 shows that $P^W(n)$ is not in general closed under the action of $GL(n)$. However, by Proposition 6.2.3 the vector space isomorphism $P^W(n) \cong P^{\leq W}(n)/P^{<W}(n)$ may be used to define a matrix action on $P^W(n)$ for either the left or the right order.

The spaces $Q^W(n)$ can be obtained formally as the graded vector spaces of a filtration on $Q^d(n)$. The order $<$ induces a filtration on $P^d(n)$ by the subspaces $P^{\leq W}(n)$. If W is the highest element in $\mathbb{V}^d(n)$ then $P^{\leq W}(n) = P^d(n)$. Intersecting with the hit elements gives a filtration $P^{\leq W}(n) \cap H^d(n)$ of $H^d(n)$ and the cohits $Q^d(n)$ are then filtered by the quotients $P^{\leq W}(n)/(P^{\leq W}(n) \cap H^d(n))$. The equations which follow arise from standard isomorphisms of vector space theory. (For clarity, we drop the n in $P(n)$ and similar notation.) We have $P^{\leq W}/(P^{\leq W} \cap H^d) \cong (P^{\leq W} + H^d)/H^d$. Also $P^{<W}/(P^{<W} \cap H^d) \cong (P^{<W} + H^d)/H^d$. Hence the associated graded spaces of the filtration of $Q^d(n)$ are given by $P^{\leq W}/(P^{\leq W} \cap H^d)/(P^{<W}/(P^{<W} \cap H^d) \cong (P^{\leq W} + H^d)/(P^{<W} + H^d) \cong (P^W + P^{<W} + H^d)/(P^{<W} + H^d) \cong P^W/(P^{<W} + H^d) \cap P^W$. This last quotient is the set of equivalence classes of elements of P^W under the reduction relation $\approx$ of Definition 6.2.4, in other words

$$Q^W(n) = P^W(n)/(P^{<W}(n) + H^d(n)) \cap P^W(n).$$

Taking the direct sum of the associated graded spaces of the filtration, we deduce the following result.

Proposition 6.2.6. *For the left or right order on $\mathbb{V}^d(n)$ there is a direct sum decomposition of vector spaces*

$$Q^d(n) \cong \bigoplus_{W \in \mathbb{V}^d(n)} Q^W(n).$$

□

This is not in general a $\mathbb{F}_2GL(n)$-module decomposition, and $Q^W(n)$ is not in general a subspace or a quotient of $Q^d(n)$ in a natural way.

Example 6.2.7. Let $W = (k)$ where $1 \leq k \leq n$. Then W is the highest element in $\mathbb{V}^k(n)$ in both the left and right orders, and $P^W(n)$ is the subspace of $P^k(n)$ spanned by the k-fold products of the variables. Clearly $Q^W(n) = P^W(n)$. Since $P^{<W}(n)$ is spanned by monomials divisible by x_i^2 for some i, $P^{<W}(n)$ is a $GL(n)$-submodule of $P^k(n)$ and $P^W(n)$ is the corresponding quotient module. As a $\mathbb{F}_2GL(n)$-module, $Q^W(n) = P^W(n)$ is isomorphic to the kth exterior power $E^k(n)$ of the 'natural' module $P^1(n)$ by mapping $x_{i_1}x_{i_2}\cdots x_{i_k}$ to $x_{i_1} \wedge x_{i_2} \wedge \cdots \wedge x_{i_k}$.

The notation and result of Proposition 1.4.6 carry over immediately to $Q^W(n)$. For a subset $Y \subseteq \{1, \ldots, n\}$ we define $Q^W(Y)$ to be the subspace of Q^W spanned by monomials divisible exactly by the variables x_i for $i \in Y$. In particular we define $Q^W[k] = Q^W(Y_k)$ for the particular subset $Y_k = \{1, \ldots, k\}$.

Proposition 6.2.8. *As vector spaces* $Q^W(n) \cong \bigoplus_Y Q^W(Y)$, *where the sum is over the subsets* Y *of* $\{1, \ldots, n\}$. *Consequently*

$$\dim Q^W(n) = \sum_{k=1}^{n} \binom{n}{k} \dim Q^W[k].$$

Proof. The vector spaces $Q^W(Y)$ are independent for different choices of Y, and a suitable permutation of variables induces an isomorphism $Q^W(Y) \cong Q^W[k]$, where $|Y| = k$. There are $\binom{n}{k}$ choices of Y with cardinality k. □

Example 6.2.9. There are two members of $\mathbb{V}^3(3)$, $W = (3)$ and $W' = (1,1)$, both decreasing. Following the calculations in Example 1.4.7 we have $\dim Q^W(3) = 1$, and $Q^W(3)$ is generated by xyz. The calculation of $Q^{W'}(3)$ reduces to the cases $n = 1, 2$ because a monomial in $P^{W'}(3)$ cannot be divisible by all three variables. With reference to Proposition 6.2.8 we have

$$\dim Q^{W'}(3) \quad = \quad 3 \dim Q^3[2] + 3 \dim Q^3[1].$$

From Example 1.6.5 and 1.5.4 we obtain $\dim Q^{W'}(3) = 6$. Hence $\dim Q^3(3) = 7$. A monomial basis for $Q^3(3)$ is given by $\{xyz, x^2y, y^2z, z^2x, x^3, y^3, z^3\}$.

6.3 Concatenation of blocks

The main purpose of introducing blocks representing monomials is to facilitate the construction of hit equations by exploiting combinatorial devices which are natural for blocks. One of these devices is **concatenation**, but it requires a refinement of the definition of block. By an (n, c)-matrix we mean, as usual, a matrix with n rows and c columns. An n-matrix has n rows, infinitely many columns but only a finite number of nonzero entries. The term 'matrix' refers to either type and we shall call a (n, c)-matrix a **finite matrix**. A **finite block** is a finite matrix whose entries are the integers 0 or 1. A **finite vector** is a $(1, c)$-matrix. We are not allowed to ignore trailing zeros in finite matrices or vectors. An (n, c)-matrix gives rise to a unique n-matrix by adding trailing zeros. As in Definition 3.1.7, The **length** of a finite vector $V = (v_1, \ldots, v_c)$ is maximum j such that $v_i > 0$. We refer to c simply as the number of entries of V. A monomial $f \in P(n)$ is represented by an (n, c)-block, and its ω-vector $\omega(f)$ by a finite vector with c entries, if and only if all exponents in f are $< 2^c$. The modulus $|V|$, degree $\deg(V)$ and the left and right orders are defined for finite vectors in the obvious way.

Definition 6.3.1. Let A be a finite matrix and B a matrix with the same number of rows. The **concatenation** of A and B is the matrix $A|B$ formed by the columns of A followed by the columns of B.

In particular we can concatenate two finite blocks having the same number of rows to form a finite block, or a finite block with a block to form a block, but we can not concatenate an n-block with another block. Under the same proviso we can concatenate vectors, which are matrices with a single row, and the following result is obvious.

Proposition 6.3.2. *For a finite block F and block G with the same number of rows we have $\omega(F|G) = \omega(F)|\omega(G)$.* □

By definition of left order it is clear that for finite vectors V, V', if $V' <_l V$, then $V'|W' <_l V|W$ for any vectors W, W'. For the right order we have $V'|W' <_r V|W$ if V, V' have the same number of columns and $W' <_r W$.

If F is an (n, c)-block representing the monomial $f \in P(n)$ and G is an n-block representing $g \in P(n)$, then $F|G$ represents fg^{2^c}. Conversely, if no exponent of f exceeds $2^c - 1$, then fg^{2^c} is represented by $F|G$.

Concatenation may be extended to an (n, c)-block F and a sum of n-blocks $G = B_1 + \cdots + B_k$, representing a polynomial g, by $F|G = A|B_1 + \cdots + A|B_k$, which then represents fg^{2^c}. However, it is not useful to allow F to be a sum of finite blocks, even of the same degree, unless they have the same number of columns because the degree of $F|G$ depends on the number of columns in F.

Proposition 6.3.3. *Let $f \in P^V(n)$, where V is a finite vector with c entries, and let $g \approx_l g' \in P^W(n)$. Then $fg^{2^c} \approx_l fg'^{2^c}$ in $P^{V|W}(n)$. In particular, fg^{2^c} is left reducible if g is left reducible.*

Proof. The hypothesis on f means that all exponents of all monomials which appear in f are $< 2^c$. By Definition 6.2.4, $g = g' + h + e$ where h is hit and $\omega(e) <_l W$. Hence $fg^{2^c} = fg'^{2^c} + fh^{2^c} + fe^{2^c}$. There is a hit equation $h = \sum_{k>0} Sq^k(h_k)$. By Proposition 1.3.2 it follows that $h^{2^c} = \sum_{k>0} Sq^{k2^c}(h_k^{2^c})$. Then by the χ-trick 2.5.2

$$fh^{2^c} = \sum_{k>0} fSq^{k2^c}(h_k^{2^c}) \sim \sum_{k>0} Xq^{k2^c}(f)h_k^{2^c}.$$

By Proposition 6.1.5, $\omega(Xq^{k2^c}(f)) <_l V$. It follows that every monomial appearing in $Xq^{k2^c}(f)h_k^{2^c}$ has an ω-vector of the form $V'|W'$, where V' is a vector with c entries and $V' <_l V$. This implies that $V'|W' <_l V|W$. Hence we can write $fh^{2^c} = h' + e'$ where h' is hit and $\omega(e') <_l V|W$. Also, writing $e'' = fe^{2^c}$, we have $\omega(e'') = \omega(f)|\omega(e) <_l V|W$ by the assumption $\omega(e) <_l W$. Finally we obtain $fg^{2^c} = fg'^{2^c} + h' + e' + e''$, where h' is hit and $\omega(e' + e'') <_l V|W$. Hence $fg^{2^c} \approx_l fg'^{2^c}$. For the last statement we put $g' = 0$. □

In terms of blocks we can paraphrase the above proposition as follows.

Proposition 6.3.4. *Let F, G, G' be blocks with the same number of rows and suppose F is finite. If $G \approx_l G'$ then $F|G \approx_l F|G'$. If G is left reducible, in particular if G is hit, then $F|G$ is left reducible.* □

Example 6.3.5. It is not true in general that fg^{2^c} is hit if g is hit. For example, let $f = xyz$, $g = x^2$ and $c = 1$, so that $fg^{2^c} = x^5yz$. Then $\omega(f) = (3)$ and g is hit. However, x^5yz has degree 7 and so by Proposition 1.6.2 it is not hit. It is however left reducible, and an explicit hit equation appears in Example 6.1.4.

Proposition 6.3.3 has the following analogue for the right order.

Proposition 6.3.6. *Let $f \approx_r f'$ in $P^V(n)$ where V is a finite vector V with c entries, and let $g \in P^W(n)$. Then $fg^{2^c} \approx_r f'g^{2^c}$ in $P^{V|W}(n)$.*

Proof. The proof follows the same line as that of Proposition 6.3.3. By Definition 6.2.4 we can write $f = f' + h + e$, where h is hit and $\omega(e) <_r V$. Then $fg^{2^c} = f'g^{2^c} + hg^{2^c} + eg^{2^c}$ and there is a hit equation $h = \sum_{k>0} Sq^k(h_k)$. By the χ-trick

$$hg^{2^c} = \sum_{k>0} Sq^k(h_k)g^{2^c} \sim \sum_{k>0} h_k Xq^k(g^{2^c}).$$

By Proposition 1.3.1, $Xq^k(g^{2^c})$ is either zero or is a sum of elements of the form $\theta(g)^{2^c}$ for some $\theta \in \mathcal{A}_2^+$. Using 6.1.7 we obtain $\omega(hg^{2^c}) <_r V|W$. Also $\omega(eg^{2^c}) <_r V|W$ because $\omega(e) <_r V$. Hence $fg^{2^c} \approx_r f'g^{2^c}$ as required. □

In general $f \approx_l f'$ does not imply $fg^{2^c} \approx_l f'g^{2^c}$. In Example 6.3.9 we shall see a case where f is hit but fg^{2^c} is not hit. However, there is a useful variation of Proposition 6.3.3 obtained by imposing a restricted hit equation.

Proposition 6.3.7. *Let $f, f' \in P^V(n)$ and $g \in P^W(n)$ where V has c entries. Suppose $f \approx_l f'$ in such a way that $f = f' + h + e$, where $e \in P^{<_l W}(n)$ and h satisfies a restricted hit equation of the form $h = \sum_{k=1}^{2^c-1} Sq^k(h_k)$. Then $fg^{2^c} \approx_l f'g^{2^c}$ in $P^{V|W}(n)$ In particular, if f is hit then so is fg^{2^c}.*

Proof. Following the proof of Proposition 6.3.3, we see by the χ-trick that

$$fg^{2^c} = f'g^{2^c} + \sum_{k=1}^{2^c-1} Sq^k(h_k)g^{2^c} + eg^{2^c} \sim f'g^{2^c} + \sum_{k=1}^{2^c-1} h_k Xq^k(g^{2^c}) + eg^{2^c}.$$

Since $k < 2^c$, Xq^k does not involve factors Sq^t for $t \geq 2^c$. Hence $Xq^k(g^{2^c}) = 0$ by Proposition 1.3.2, and $\omega(eg^{2^c}) <_l V|W$ because $\omega(e) <_l V$ by assumption. It follows that $fg^{2^c} \approx_l f'g^{2^r}$ as required. Clearly, if f is hit then so is fg^{2^c}. □

We can paraphrase this result in terms of blocks as follows.

Proposition 6.3.8. *Let F, F' be sums of (n, c)-blocks with the same ω-vector such that $F \approx_l F'$ by a restricted hit equation (i.e. not involving Sq^k for $k \geq 2^c$). Then $F|G \approx_l F'|G$ for any n-block G.* □

The following example shows the need for the restriction on the hit equation.

Example 6.3.9. Consider the block $B = F|G$, split as shown below between columns 2 and 3, so that $c = 2$.

$$B = \left[\begin{array}{cc|c} 1 & 1 & 1 \\ 1 & 1 & \\ 0 & 1 & \\ 0 & 1 & \\ 0 & 1 & \end{array}\right] \qquad F = \begin{array}{cc} 1 & 1 \\ 1 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{array} \qquad G = \begin{array}{c} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{array}$$

Let $W = \omega(F) = (2, 5)$. Then $\mu(w_2) = \mu(5) = 3$ and $w_1(F) = 2$, so by Proposition 2.5.3 F is hit. However, B is not hit. This can be seen by specializing the variables in the corresponding monomial $b = x_1^7 x_2^3 x_3^2 x_4^2 x_5^2$. Setting $x_5 = x_4 = x_3 = x_2$, we obtain the monomial $x_1^7 x_2^9$, which is not hit (see the diagrams following Theorem 1.8.2). By Proposition 1.6.1 it follows that b is not hit. In this case Proposition 6.3.8 fails to apply, because there is no hit equation for F involving only Sq^1 and Sq^2.

We can combine Propositions 6.3.4 and 6.3.8 to give a useful mechanism for manipulating concatenated blocks under restricted hit equations.

Proposition 6.3.10. *Let F, F' be sums of (n, c)-blocks with ω-vector V such that $F \approx_l F'$ by a restricted hit equation. Let G, G' be n-blocks with ω-vector W such that $G \approx_l G'$. Then $F|G \approx_l F'|G'$ in $P^{V|W}$.* □

We may therefore manipulate the blocks F and G individually in $F|G$ up to equivalence without altering the equivalence class of the concatenated block in $Q^{W|W'}(n)$ providing the relevant hit equation is restricted.

We complete this section with one case where $Q^W(n)$ can be determined. From Chapter 5 we recall the notation $\omega^{\min}(d)$ for the lowest vector in $\mathbb{V}^d(n)$ in the left order and the corresponding Ferrers block $F^{\min}(n, d)$.

Proposition 6.3.11. *If B is a block in $P^d(n)$ such that $\omega(B) <_l \omega^{\min}(d)$, then B is hit, i.e. in the left order, $Q^W(n) = 0$ for all $W <_l \omega^{\min}(d)$.*

Proof. Let $W = \omega(B)$ and $W' = \omega^{\min}(d)$, so that $W <_l W'$. Hence, for some k, $w_j = w'_j$ for $1 \leq j < k$ and $w_k < w'_k$. Let B be split as $B = A|G|H$ where G is column k of B. Then $\mu(\deg H) > \deg(G) = w_k$, since otherwise we could make a Ferrers block lower than $F^{\min}(n, d)$. It follows from Proposition 2.5.3 that $F = G|H$ is hit. Then by Proposition 6.3.4 we know that $B = A|F$ is left reducible. We can therefore write $B \approx_l \sum_i B^i$ where $\omega(B^i) <_l W <_l \omega^{\min}(d)$ for all i. Since $\mathbb{V}^d(n)$ is a finite set it follows by iteration of the procedure on the blocks B^i that B is hit. □

6.4 Splicing

Splicing is a process which generalizes Example 6.1.1 and enables the construction of hit equations to which the results of Section 6.3 can be applied.

Definition 6.4.1. Let $F = (f_{i,j})$ denote a block such that for some row r and columns $u < v$ we have $f_{r,t} = 0$ for $u \leq t < v$ and $f_{r,v} = 1$. Let H denote the block obtained from F by defining $h_{r,t} = 1$ for $u \leq t < v$ and $h_{r,v} = 0$, leaving other entries of F unchanged. Then there is an equation of the form $F = G + Sq^{2^{u-1}}(H) + E$, where E and G are sums of blocks such that each block in G arises by the action of $Sq^{2^{u-1}}$ on a row of H other than row r, at column u, and each block in E arises from Steenrod operations acting at columns prior to u. The process of forming H, E, G is called **splicing** the block F at (r, u).

Example 6.4.2. The following illustrates splicing the block F at position $(2, 3)$.

$$F = \begin{matrix} 1 & 1 & 1 & 1 & & \\ 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & & & \end{matrix}, \quad H = \begin{matrix} 1 & 1 & 1 & 1 & & \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & & & \end{matrix}, \quad Sq^4(H) = F + E + G,$$

where, from the Cartan formula and Example 6.1.1, we have

$$G = \begin{matrix} 1 & 1 & 0 & 0 & 1 & & \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & & & \end{matrix} + \begin{matrix} 1 & 1 & 1 & 1 & & \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & & \end{matrix},$$

$$E = \begin{matrix} 1 & 0 & 0 & 0 & 1 & & \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & & \end{matrix} + \begin{matrix} 0 & 0 & 0 & 0 & 1 & & \\ 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & & \end{matrix} + \begin{matrix} 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & & & \end{matrix}.$$

Since the blocks in E arise from the action of Steenrod squares at columns $< u$, $\omega(E) <_l \omega(F)$. We regard these blocks as 'error' terms. On the other hand, blocks in G may have ω-vector $>_l \omega(F)$.

We have better control in the case $v = u + 1$. The splicing process then consists of moving a digit 1 in row r of F back one place previously occupied by 0 to form H. If $f_{r',u} = 1$ where $r' \neq r$, the effect of $Sq^{2^{u-1}}$ in row r' of H is to add the number 1 arithmetically at position (r', u). If $f_{r',u} = 1$, the resulting block G has $\omega(G) <_l \omega(F)$ and $\alpha_{r'}(G) < \alpha_{r'}(F)$. On the other hand, if $f_{r',u} = 0$, then neither the ω-vector nor the α value change. We shall refer to the case $v = u+1$ as 1-**back splicing**. We summarize these observations as follows.

Proposition 6.4.3. *In the case of* 1*-back splicing on an* n*-block* F *at position* (r, u) *let* H *be the block formed from* F *by moving the digit* $f_{r,u+1} = 1$ *back one place. Let* G *be the sum of blocks arising from the application of* $Sq^{2^{u-1}}$ *to rows of* H *other than row* r*. Then* $F \approx_l G$*. Hence* F *and* G *represent the same element of* $Q^W(n)$*, where* $W = \omega(F)$*.* □

Example 6.4.4. In the example below, $F \approx_l G$ by 1-back splicing F at $(1,2)$.

$$F = \begin{matrix} 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & & & \\ 1 & 1 & & & \end{matrix}, \; G = \begin{matrix} 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & & \\ 1 & 1 & & & \end{matrix} + \begin{matrix} 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & & & \\ 1 & 0 & 1 & & \end{matrix}.$$

By iterated 1-back splicing in a specified nonzero row of a block F, it is clear that we can find a sum of blocks $B = \sum_i B_i$ such that $B \approx_l F$ and, for all i, all digits 1 in that row of B_i are contiguous. This gives the following result.

Proposition 6.4.5. *For $1 \leq i \leq n$, $Q^W[n]$ has a spanning set consisting of monomials in which the exponent of x_i is an integer of the form $2^j - 1$.* □

The basis for $Q(2)$ shown in (1.1) satisfies this condition for the variable y.

Proposition 6.4.6. *A block with a non-trailing zero column is left reducible.*

Proof. Suppose column u of a block F is zero and $f_{r,u+1} = 1$. Then 1-back splicing at position (r,u) gives $G = 0$. By Proposition 6.4.3, F is left reducible. □

6.5 The Kameko and duplication maps

We begin by discussing two ways of generalizing the Kameko map of Section 1.7. The first applies to $Q^d(n)$, while the second is a local version which applies to $Q^W(n)$. In this section, the order used to define the cohit module $Q^W(n)$ and the ω-vector of a polynomial is always the left order $<_l$. For a vector W in $\mathbb{V}^d(n)$, we denote by W^+ the vector in $\mathbb{V}^{2d+n}(n)$ obtained by prefixing n, i.e. $W^+ = (n)|W$.

Definition 6.5.1. The 'up' Kameko map $\kappa : P^d(n) \to P^{2d+n}(n)$ sends $f \in P^d(n)$ to $g = mf^2 \in P^{2d+n}(n)$, where $m = x_1 \cdots x_n$. Given $W \in \mathbb{V}^d(n)$, κ restricts to a map $\kappa : P^W(n) \to P^{W^+}(n)$. The 'down' Kameko map $\kappa' : P^{2d+n}(n) \to P^d(n)$ sends $g \in P^{2d+n}(n)$ to 0 unless $g = mf^2$. In this case $\kappa'(g) = f$, and κ' restricts to a map $\kappa' : P^{W^+}(n) \to P^W(n)$.

In the language of blocks, if the $(n,1)$-block M represents m and F represents f then $\kappa(F) = M|F = G$. In the other direction $\kappa'(G) = 0$ unless $G = M|F$ in which case $\kappa'(G) = F$.

We shall frequently use the fact that a polynomial in $P^d(n)$, where d and n have the same parity, can be written uniquely in the form $mf^2 + e$ where $\omega_1(e) < n - 1$.

Proposition 6.5.2. *The map $\kappa' : P^{2d+n}(n) \to P^d(n)$ is a $\mathbb{F}_2GL(n)$-module map and induces a $\mathbb{F}_2GL(n)$-module map $\kappa' : Q^{2d+n}(n) \to Q^d(n)$.*

Proof. We first show that κ' sends hit elements to hit elements. With the notation as above, it is enough to show that if mf^2 is hit then so is f. A hit equation for mf^2 may be written in the form

$$mf^2 = Sq^1(e) + \sum_{k>0} Sq^{2k}(mh_k^2 + e_k),$$

where $\omega_1(e), \omega_1(e_k) < n$ for $k > 0$. Now $Sq^{2k}(mh_k^2) = mSq^{2k}(h_k^2) + e'_k$, where $\omega_1(e'_k) < n$. The same is clearly true for $Sq^1(e)$ and $Sq^{2k}(e_k)$. Comparing terms with $\omega_1 = n$ on each side of the equation gives $mf^2 = \sum_{k>0} m(Sq^k(h_k))^2 = m\sum_{k>0}(Sq^k(h_k))^2$. Hence $f = \sum_{k>0} Sq^k(h_k)$ as required.

Next we show that κ' is a $\mathbb{F}_2GL(n)$-map. The result is obviously true for permutation matrices. For the standard transvection T we obtain $(mf^2)\cdot T = m(f\cdot T)^2 + h$, where $\omega_1(h) < n$. Hence, by definition of κ', we have $\kappa'(h) = 0$. It follows that $\kappa'((mf^2)\cdot T) = \kappa'(m(f\cdot T))^2 = f\cdot T = (\kappa'(mf^2))\cdot T$ as required. □

In general, the up Kameko map $\kappa : P^d(n) \to P^{2d+n}(n)$ does not send hit elements to hit elements. For example, $x^2 \in P^2(3)$ is hit but $\kappa(x^2) = x^5yz$ is not (see Example 6.3.5). However $\kappa : Q^d(n) \to Q^{2d+n}(n)$ is defined for certain degrees d, and is the inverse of κ'.

Proposition 6.5.3. *If $\mu(2d+n) = n$, then κ induces a map $\kappa : Q^d(n) \to Q^{2d+n}(n)$ which is the inverse of κ'. In particular, $Q^{2d+n}(n) \cong Q^d(n)$ as $\mathbb{F}_2GL(n)$-modules.*

Proof. We need to show that κ sends hit elements to hit elements. Suppose that $f \in P^d(n)$ is hit. By the χ-trick, we show that mf^2 is reducible to an element g of $P^{2d+n}(n)$ with $\omega_1(g) < n$ as follows. Given a hit equation $f = \sum_{k>0} Sq^k(h_k)$, we have $mf^2 = \sum_{k>0} mSq^{2k}(h_k^2) \sim \sum_{k>0} Xq^{2k}(m)(h_k^2) = g$. Now $\omega_1(Xq^{2k}(m)) < n$. Hence $\omega_1(g) < n$. In our present situation we know from Proposition 5.5.6 that $\omega_1^{\min}(2d+n) = n$ because $\mu(2d+n) = n$. It follows that $\omega(g) <_l \omega^{\min}(2d+n)$, and so g is hit by Proposition 6.3.11. Hence $\kappa : Q^d(n) \to Q^{2d+n}(n)$ is well defined, and clearly it is the inverse of κ'. □

Now we look at the local version of the up Kameko map κ.

Proposition 6.5.4. *For $W \in \mathbb{V}^d(n)$, $\kappa : P^d(n) \to P^{2d+n}(n)$ induces a $\mathbb{F}_2GL(n)$-module isomorphism $\kappa : Q^W(n) \to Q^{W^+}(n)$, where $W^+ = (n)|W$, with inverse κ'.*

Proof. We first demonstrate that $\kappa : Q^W(n) \to Q^{W^+}(n)$ is well defined by showing that if $f \approx_l 0$ in $P^W(n)$ then $mf^2 \approx_l 0$ in $P^{W^+}(n)$. By Definition 6.2.4 we can write $f = h + e$ where h is hit and $\omega(e) <_l W$. Then $\omega(me^2) <_l W^+$ and, by the same argument as in the proof of Proposition 6.5.3, $mh^2 \approx_l 0$ in $P^{W^+}(n)$. Hence $mf^2 \approx_l 0$ as required. Clearly κ is surjective and the down Kameko map induces the inverse of κ. □

Example 6.5.5. Using Proposition 6.5.4, $Q^{(3,0,1)}(3) \cong Q^{(0,1)}(3) = 0$ for the left order. Although x^5yz is not hit (see Example 6.3.5), it is left reducible by Example 6.1.4, and so it represents 0 in $Q^{(3,0,1)}(3)$.

The next proposition makes a minor simplification of the hit problem which is useful for small values of n.

Proposition 6.5.6. *For all $d' \geq 0$ and $W' \in \mathbb{V}^{d'}(n)$, $Q^{d'}(n) \cong Q^d(n)$ and $Q^{W'}(n) \cong Q^W(n)$ for some d with $\mu(d) < n$ and some $W \in \mathbb{V}^d(n)$.*

Proof. If $\mu(d') = n$ then d' and n have the same parity, and we can write $d' = 2d + n$. We can then apply Proposition 6.5.3 iteratively until $\mu(d) < n$. Using Proposition 6.5.4, a similar argument applies to the local case. □

We now look at the situation where $\mu(d) = n-1$ and generalize the duplication map of Section 1.8. This is more subtle than the Kameko map. We do not have a down duplication map analogous to κ', and the local version of the up duplication map is not in general an isomorphism. In terms of blocks, the duplication map repeats the first column of F when this column contains exactly one entry 0. If f is a spike in $P^d(n)$, then $\delta(f)$ is a spike in $P^{2d+n-1}(n)$.

Definition 6.5.7. Let $m = x_1 \cdots x_n$ and let $m_i = m/x_i$ for $1 \leq i \leq n$. The **duplication map** $\delta : P^d(n) \to P^{2d+n-1}(n)$ is the linear map defined by $\delta(m_i g^2) = m_i^3 g^4$ for $1 \leq i \leq n$ and all monomials g, and $\delta(f) = 0$ if $\omega_1(f) < n-1$.

The duplication map δ matches degrees with opposite parity to n whereas, after possibly one iteration, the Kameko map matches degrees of the same parity as n. We shall frequently use the fact that, if d and n have opposite parities, then $f \in P^d(n)$ has a unique expression of the form $f = \sum_{i=1}^n m_i f_i^2 + e$, where $\omega_1(e) < n-1$. The next two propositions state criteria for a polynomial f in $P^d(n)$ to be hit or reducible when d and n have opposite parity.

Proposition 6.5.8. *Let $n > 1$ and let n and d have opposite parity. Let $f = \sum_{i=1}^n m_i f_i^2 + e$ be a polynomial in $P^d(n)$, where $f_i \in P^{(d-n+1)/2}(n)$ for $1 \leq i \leq n$, and $\omega_1(e) < n-1$. If f is hit, then there is a polynomial $h \in P^{(d-n-1)/2}(n)$ such that $f_i \sim x_i h$ for all i. Conversely, if all monomials in $P^d(n)$ with $\omega_1 < n-1$ are hit, and if $f_i \sim x_i h$ for all i, then f is hit.*

Proposition 6.5.9. *For $n > 1$ let W be a vector in $\mathbb{V}^d(n)$ with $\omega_1 = n-1$ and $g = \sum_{i=1}^n m_i g_i^2$ an element in $P^W(n)$. Then $g \approx_l 0$ if and only if there exists $h \in P^{(d-n-1)/2}$ such that $g_i \approx_l x_i h$ for $1 \leq i \leq n$.*

Proof. We prove the two propositions together starting with Proposition 6.5.9. By assumption we can write $W = (n-1)|W'$ for some vector W'. Now let $g = \sum_{i=1}^n m_i g_i^2$ be left reducible. Then by definition $g = f + e'$ where f is hit and

$e' \in P^{<_l W}(n)$. Hence we can write $e' = \sum_{i=1}^n m_i c_i^2 + c$, where $\omega_1(c) < n-1$ and $\omega(c_i) <_l W'$ for each i. A hit equation for f may be taken in the form

$$f = Sq^1(a) + \sum_{k>0} Sq^{2k}(a_k) \tag{6.1}$$

where $a \in P^{d-1}(n)$ and $a_k \in P^{d-2k}(n)$. Since $d = n-1 \bmod 2$, $\deg(a) = n \bmod 2$ and $\deg(a_k) = n-1 \bmod 2$. Hence we may write

$$a = mh^2 + b, \quad a_k = \sum_{i=1}^n m_i h_{i,k}^2 + b_k,$$

where $\omega_1(b)$, $\omega_1(b_k) < n-1$. The same is true for $Sq^1(b)$ and $Sq^{2k}(b_k)$. Now $Sq^1(m) = \sum_{i=1}^n m_i x_i^2$. Hence $Sq^1(mh^2) = \sum_{i=1}^n m_i (x_i h)^2$ and $\omega_1(Sq^1(b)) < n-1$. Also $Sq^{2k}(m_i h_{i,k}^2) = m_i (Sq^k(h_{i,k}))^2 + b_{i,k}$ where $\omega_1(b_{i,k}) < n-1$. It follows that

$$g = \sum_{i=1}^n m_i g_i^2 = \sum_{i=1}^n m_i (x_i h)^2 + \sum_{i=1}^n m_i \sum_{k>0} (Sq^k(h_{i,k}))^2 + \sum_{i=1}^n m_i c_i^2,$$

modulo monomials which have $\omega_1 < n-1$. Equating terms in m_i gives $g_i = x_i h + \sum_{k>0} Sq^k(h_{i,k}) + c_i$. In other words $g_i \approx_l x_i h$, as required in Proposition 6.5.9.

To obtain the proof of Proposition 6.5.8 we assume $f \in P^d(n)$ is hit. Then $f = \sum_{i=1}^n m_i f_i^2 + e$, where $\omega_1(e) < n-1$. The above analysis applies to f and we have

$$f = \sum_{i=1}^n m_i f_i^2 = \sum_{i=1}^n m_i (x_i h)^2 + \sum_{i=1}^n m_i \sum_{k>0} (Sq^k(h_{i,k}))^2$$

modulo monomials with $\omega_1 < n-1$. Equating terms in m_i gives $f_i = x_i h + \sum_{k>0} Sq^k(h_{i,k})$. In other words $f_i \sim x_i h$, as required in one direction for Proposition 6.5.8. The proof in the other direction of both propositions follows by reversing the steps in the above argument, on the assumption in Proposition 6.5.8 that all monomials in $P^d(n)$ with $\omega_1(h) < n-1$ are hit. □

We can now state the global and local results for the duplication maps.

Proposition 6.5.10. *If $\mu(d) = n-1$, then the duplication map induces a map $\delta : Q^d(n) \to Q^{2d+n-1}(n)$ of $\mathbb{F}_2 GL(n)$-modules.*

Proposition 6.5.11. *Let $W \in \mathbb{V}^d(n)$ with $w_1 = n-1$, Then the duplication map induces a map $\delta : Q^W(n) \to Q^{W'}(n)$ of $\mathbb{F}_2 GL(n)$-modules, where $W' = (n-1)|W \in \mathbb{V}^{2d+n-1}(n)$.*

Proof. We prove the two propositions together. Starting with Proposition 6.5.10 we show that δ sends hit elements to hit elements. Since $\mu(d) = n-1$, we have $\omega_1^{\min}(d) = n-1$. It follows from Proposition 6.3.11 that any monomial in $P^d(n)$ with $\omega_1 < n-1$ is hit. The same applies to $P^{2d+n-1}(n)$ because $\mu(2d+n-1) = n-1$, as we see from Proposition 5.5.4. In particular, if $g \sim 0$ for some $g \in P^d(n)$, then $m_i g^2 \sim 0$ in $P^{2d+n-1}(n)$. To see this, we consider a hit equation $g = \sum_{k>0} Sq^k(h_k)$. Then $m_i g^2 = \sum_{k>0} m_i Sq^{2k}(h_k^2) = \sum_{k>0} Sq^{2k}(m_i h_k^2)$, modulo monomials with $\omega_1 < n-1$, which are hit. A similar argument shows that if $g \approx_l 0 \in P^W(n)$ then $m_i g^2 \approx_l 0 \in P^{W'}(n)$.

For $f = \sum_{i=1}^n m_i f_i^2 + e$ in $P^d(n)$, where $\omega_1(e) < n-1$, we have by definition $\delta(f) = \sum_{i=1}^n m_i (m_i f_i^2)^2$. If f is hit, then by Proposition 6.5.8 $f_i \sim x_i h$. It follows from the preceding argument that $m_i f_i^2 \sim m_i x_i^2 h^2 = x_i m h^2$. The factor mh^2 plays the role of h in Proposition 6.5.8 applied to $P^{2d+n-1}(n)$. Hence $\delta(f)$ is hit. Similarly, if f is reducible in $P^W(n)$ then $f_i \approx_l x_i h$. Then $m_i f_i^2 \approx_l x_i m h^2$, which shows by Proposition 6.5.9 that $\delta(f)$ is reducible in $P^{W'}(n)$. This proves that δ is well defined in Proposition 6.5.11.

The proof that δ is a map of $\mathbb{F}_2 GL(n)$-modules follows the familiar pattern. It is clear that duplication commutes with the action of permutation matrices. For the standard transvection T we may assume $n > 2$, since the 2-variable case was treated in Chapter 1. For both propositions it is enough to consider $f = \sum_{i=1}^n m_i f_i^2$. We note that $m_1 \cdot T = m_1, \quad m_2 \cdot T = m_1 + m_2, \quad m_i \cdot T = m_i + e_i$, for $i > 2$, where $\omega_1(e_i) < n-1$. Hence $f \cdot T = \sum_{i=1}^n (m_i \cdot T)(f_i \cdot T)^2 = m_1 (f_1 \cdot T)^2 + (m_1 + m_2)(f_2 \cdot T)^2 + \sum_{i=3}^n m_i (f_i \cdot T)^2 + e''$, where $\omega_1(e'') < n-1$. Then, since $\delta(e'') = 0$, we have

$$\delta(f \cdot T) \quad = m_1^3 (f_1 \cdot T)^4 + (m_1^3 + m_2^3)(f_2 \cdot T)^4 + \sum_{i=3}^n m_i^3 (f_i \cdot T)^4.$$

On the other hand

$$\delta(f) \cdot T = \sum_{i=1}^n m_i^3 \cdot T (f_i \cdot T)^4 = (m_1)^3 (f_1 \cdot T)^4 + (m_1 + m_2)^3 (f_2 \cdot T)^4 + \sum_{i=3}^n m_i^3 (f_i \cdot T)^4.$$

The difference between $\delta(f \cdot T)$ and $\delta(f) \cdot T$ is $b = (m_1^2 m_2 + m_1 m_2^2)(f_2 \cdot T)^4$. Let $c = (x_1 x_2 x_3 \cdot x_n^3)(f_2 \cdot T)^4$. Then $Sq^1(c) = b + h$, where $w_2(h) < n-1$. In particular $h \in P^{<_l W'}(n)$. It follows that $\delta(f \cdot T) \approx_l \delta(f) \cdot T$ as required in Proposition 6.5.9 and $\delta(f \cdot T) \sim \delta(f) \cdot T$ in Proposition 6.5.8. □

6.6 Head ω-vectors

In this section we identify a composition series for $Q^W(n)$ in the left order, where $W = (n-1, \ldots, n-1) \in \mathbb{V}^d(n)$ is a **head** vector of length t, so that $d = (n-1)(2^t - 1)$. By Proposition 5.6.2, W is the maximum element in $\mathbb{V}^d(n)$. Blocks with head ω-vectors are called head blocks.

Proposition 6.6.1. *Let $F, F' \in P^W(n)$ be head blocks such that the columns of F' are a permutation of the columns of F. Then, for any vector W' and block $G \in P^{W'}(n)$, the blocks $F|G$ and $F'|G$ represent the same element of $Q^{W|W'}(n)$, i.e. $F|G \approx_l F'|G$.*

Proof. Since every permutation is a product of transpositions of adjacent elements, we may assume that F' is obtained from F by interchanging columns j and $j+1$ for some j. If the zero entry in column j is in row r, then by Proposition 6.4.3 splicing $F|G$ in position (r, j) gives $F|G \approx_l F'|G$. □

Proposition 6.6.2. *Let $W \in \mathbb{V}(n)$ be a head vector of length $\geq n$, and let $V = (n-1)|W$. Then the duplication map $\delta : Q^{W|W'}(n) \to Q^{V|W'}(n)$ is surjective for any $W' \in \mathbb{V}(n)$.*

Proof. Since $\mathrm{len}(V) > n$, every block in $P^V(n)$ has at least one row with at least two zeros. By Proposition 6.6.1 we may assume this row has zeros in the first two columns. The result follows. □

Given n and t, we classify head (n, t)-blocks by the rows which contain zeros.

Definition 6.6.3. Let $Y \subseteq \{1, \ldots, n\}$ be a non-empty subset. A head (n, t)-block is in the **class** Y if it has at least one entry 0 in each row $i \in Y$ and rows $i \notin Y$ have all entries 1.

Example 6.6.4. For $n = 5$, and $t = 6$, the class $Y = \{1, 3, 5\}$ contains the block

$$\begin{array}{cccccc} 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 1 \end{array} .$$

Definition 6.6.5. For given n, t let $Y = \{i_1, \ldots, i_k\} \subseteq \{1, \ldots, n\}$ with entries in increasing order. The **canonical block** in class Y is the (n, t)-block C defined by $c_{i_r, r} = 0$ for $1 \leq r \leq k$ and $c_{i_k, j} = 0$ for $k < j \leq t$, with all other entries 1.

The canonical block in case $n = 3$, $t = 4$ and $Y = \{1, 3\}$ is

$$C = \begin{array}{cccc} 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \end{array} .$$

If $k = 1$ then C is a spike. If $Y = \{1, \ldots, k\}$, the zero entries of C are the diagonal entries $c_{i,i}$ for $1 \leq i \leq k$ and the entries $c_{k,j}$ for $j > k$, for example when

$n = 5$, $t = 6$ and $k = 3$,

$$C = \left[\begin{array}{ccc|ccc} 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{array}\right].$$

We fix Y and use splicing operations to study head blocks of class Y. Splicing in a row in Y will generally produce terms which arise from the action of a Steenrod square on rows not in Y. Since all entries in these rows are 1, such terms have ω-vector $<_l W$, and may therefore be ignored up to reducibility. Thus these rows play no essential part in the argument.

Proposition 6.6.6. *Let $W \in \mathbb{V}(n)$ be a head vector and $Y \subseteq \{1, \ldots, n\}$. Then all blocks in $P^W(n)$ of class Y represent the same element in $Q^W(n)$.*

Proof. We shall use splicing operations to bring a block B in class Y to canonical form. As observed above, we may ignore rows not in Y, and so we may assume that $Y = \{1, \ldots, n\}$ and $t \geq n$.

We argue by induction on n, the number of rows. There is nothing to prove if $n = 1$, so let $n > 1$ and assume the result for head blocks with $n - 1$ rows. By permuting columns using Proposition 6.6.1, $B \approx_l B_1$ where in B_1 all entries 0 in the first row are contiguous on the left, for example

$$B = \begin{matrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 \end{matrix}, \quad B_1 = \begin{matrix} 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \end{matrix}.$$

Next we splice B_1 at $(1, 1)$. Apart from terms with ω-vector $<_l W$, the result is a single head block B_2 with one entry 0 in its first row. Using Proposition 6.6.1 again, we can permute columns of B_2 to place the single digit 0 in this row in the first column, giving B_3. For example

$$B_2 = \begin{matrix} 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \end{matrix}, \quad B_3 = \begin{matrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \end{matrix}.$$

Thus $B \approx_l B_3$. Let B_3' be the subblock of B_3 obtained by deleting the first row and the first column. By induction hypothesis, B_3' can be brought to canonical form by splicing processes. When the same processes are applied to the whole block B_3, error terms E arising from Steenrod squares acting on the first row or column have $\omega(E) <_l W$. Hence $B_3 \approx_l C$, the canonical block. This completes the inductive step. □

Proposition 6.6.7. *For a head vector $W \in \mathbb{V}(n)$ of length t the canonical blocks $C(Y)$ associated with subsets $Y \subseteq \{1, \ldots, n\}$ of cardinality $1 \leq |Y| \leq t$ form a vector space basis of $Q^W(n)$. Thus $\dim Q^W(n) = \sum_{i=1}^{t} \binom{n}{i}$ if $t \leq n$ and $\dim Q^W(n) = 2^n - 1$ if $t \geq n$.*

Proof. By Proposition 6.6.6, the canonical blocks $C(Y)$ span $Q^W(n)$. We first prove that the corresponding monomials are not hit, and then extend the argument to prove that they represent linearly independent elements of $Q^W(n)$.

Step 1. Let $Y = \{i_1, \ldots, i_k\}$ with entries in increasing order. For the first step, the method is to apply a sequence of transvections to C to obtain a spike modulo blocks with lower ω-vectors and then appeal to Proposition 1.6.1. Proposition 6.1.9 explains how to perform a transvection in terms of two rows of a block by removing a subset S of digits 1 from one row and adding them arithmetically to the other. In general this will produce a sum of blocks including the original one (taking S to be empty). In our present situation we form the sequence of blocks $F_0 = C, F_1, \ldots, F_{n-1}$, where F_j is obtained from F_{j-1}, for $1 \leq j < n$, by arithmetically adding row i_k of F_{j-1} to row i_{n-j}, discarding blocks with ω-vector lower than W and F_{k-1} itself. These are the only types of block produced by the transvections in this case. The final block F_{n-1} is a spike. For example, when $n = 3$, $t = 5$ and $Y = \{1, 2, 3\}$,

$$F_0 = \begin{matrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 \end{matrix}, \quad F_1 = \begin{matrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{matrix}, \quad F_2 = \begin{matrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{matrix}.$$

Step 2. Suppose that a sum of canonical blocks $C(Y)$ is reducible, and let $C(Y)$ be a term in the sum for which Y has minimum cardinality. The procedure in Step 1 affects only the rows in Y, and cannot produce a spike out of any other block in the sum, since such a block must have a zero entry in some row not in Y. This contradicts reducibility of the sum. □

From Propositions 6.6.6 and 6.6.7, any choice of head block of class Y for each Y yields a basis of $Q^W(n)$. The duplication map preserves the class of a block and sends basis elements to basis elements. If $t > n$, then some row of any block must have at least two zero entries. By Proposition 6.6.1, we may assume that these are in the first two columns, giving the following result.

Proposition 6.6.8. *Let $W = (n-1, \ldots, n-1)$ be a head vector of length $t \geq 1$, and let $W' = (n-1)|W$. Then the duplication map $\delta : Q^W(n) \to Q^{W'}(n)$ is an injective map of $\mathbb{F}_2 GL(n)$-modules, and is an isomorphism if $t \geq n$.* □

Identifying $Q^W(n)$ as a submodule of $Q^{W'}(n)$ via δ, we obtain a filtration

$$Q^{(n-1)}(n) \subset Q^{(n-1,n-1)}(n) \subset \cdots \subset Q^{(n-1,\ldots,n-1)}(n), \tag{6.2}$$

where the last vector $(n-1,\dots,n-1)$ has length n. As in Example 6.2.7, for $0 \leq k \leq n$ we denote by $E^k(n)$ the kth exterior power of the natural $\mathbb{F}_2GL(n)$-module $P^1(n)$. It is a standard fact of representation theory that $E^k(n)$ is a simple module, i.e. it has no non-trivial submodules. Recall from Example 6.2.7 that $Q^{(n-1)}(n) \cong E^{n-1}(n)$. We shall prove that the quotient modules of the filtration (6.2) are exterior powers of $P^1(n)$, so that (6.2) is a composition series.

Proposition 6.6.9. *Let $W = (n-1,\dots,n-1)$ have length $k-1$, $2 \leq k \leq n$ and let $W' = (n-1)|W$. Then as $\mathbb{F}_2GL(n)$-modules $Q^{W'}(n)/Q^W(n) \cong E^{n-k}(n)$.*

Proof. Since $k \leq n$, the quotient $Q^{W'}(n)/Q^W(n)$ has a vector space basis consisting of the canonical blocks $C(Y)$ for $|Y| = k$ having a single entry 0 in rows in Y. Let the rows not in Y, where all entries are 1, be rows $i_1,\dots,i_{n-k}$ in increasing order. Then the map $\phi : Q^{W'}(n)/Q^W(n) \to E^{n-k}(n)$ which assigns to $C = C(Y)$ the element $x_{i_1} \wedge x_{i_2} \wedge \cdots \wedge x_{i_{n-k}} \in E^{n-k}(n)$ is a vector space isomorphism. It is clear that ϕ commutes with the action of permutation matrices. Since $GL(n)$ is generated by permutation matrices and the standard transvection T, it suffices to check that ϕ commutes with the action of T. This fixes $u = x_{i_1} \wedge x_{i_2} \wedge \cdots \wedge x_{i_{n-k}}$ unless $i_1 = 1$ and $i_2 > 2$, when it maps u to $u+v$ where v is obtained by replacing x_1 by x_2 in u. In the first case, if $i_1 = 1$ and $i_2 = 2$ or if $i_1 = 2$ then all digits of row 2 are 1. By Proposition 6.1.9 $C \cdot T = C + E$ where $E \in P^{<W}(n)$. Otherwise $i_1 > 2$ and $C \cdot T = C + D + E$ where D is a duplicate block whose first two columns are the same as the first column of C, and $E \in P^{<W}(n)$. Thus D is in the submodule $Q^W(n)$ of $Q^{W'}(n)$. In the second case, where $i_1 = 1$ and $i_2 > 2$, all digits in row 1 of C are 1 and $c_{2,1} = 0$. By Proposition 6.1.9, $C \cdot T = C + C' + E$ where C' is the block obtained from C by exchanging the entries $c_{1,1}$ and $c_{2,1}$, and $E \in P^{<W}(n)$. Since $\phi(C') = v$, this completes the proof. □

6.7 Remarks

Reducibility techniques for manipulating blocks have been natural ingredients in work on the hit problem since its inception, for example in Kameko's thesis [76]. The fact that blocks below the minimal spike are hit is due to Singer [134]. The duplication map first appeared in [158].

Appendix A

Differential operators

There is a natural explanation of the action of Steenrod squares on polynomials in terms of differential operators with variable coefficients. Let $D = \sum_{i\geq 1} x_i^2 \frac{\partial}{\partial x_i}$. Although the differential operator is a formally infinite sum, its action on a polynomial is finite and it is easy to verify that $Sq^1(f) = D(f)$ mod 2 for any $f \in P(n)$. The **wedge** product of two differential operators allows the first operator to pass the second without acting, in other words all symbols commute. For example

$$(\sum_{i\geq 1} x_i^2 \frac{\partial}{\partial x_i}) \vee (\sum_{i\geq 1} x_i^2 \frac{\partial}{\partial x_i}) = (\sum_{i\geq 1} x_i^2 \frac{\partial}{\partial x_i})^{\vee 2} = \sum_{i\geq 1, j\geq 1} x_i^2 x_j^2 \frac{\partial^2}{\partial x_i \partial x_j}.$$

This particular operator is divisible by 2 and we can write $\frac{D^{\vee 2}}{2}$ as an integral operator. More generally $\frac{D^{\vee k}}{k!}$ is an integral operator and we have $Sq^k = \frac{D^{\vee k}}{k!}$ mod 2. The total Steenrod square may then be expressed by an exponential generating function $Sq = \exp_{\vee}(D) = \sum_{k\geq 0} \frac{D^{\vee k}}{k!}$. For example, if $x \in P^1(n)$ then $\exp_{vee}(D)(x) = x + x^2$. In this context the Cartan formula is the Leibniz formula $\exp_{vee}(D)(fg) = \exp_{vee}(D)(f)\exp_{vee}(D)(g)$ for differentiating a product of functions. Further development of the Steenrod algebra from the differential operator point of view can be found in [166] and [167].

The study of the hit problem is motivated by several problems in topology and algebra. In 1987 Frank Peterson posed and solved the hit problem for $n = 2$ [118], as far as the dimension of the space $Q^d(2)$ is concerned. He conjectured that $Q^d(n) = 0$ unless $\alpha(d+n) \leq n$. An equivalent statement of the conjecture is that f is hit if there are no spikes with the same degree and in the same variables as f. Peterson's reason for studying the hit problem was to do with the following conjecture: if M is a smooth manifold of dimension d such that all products of length n of Stiefel-Whitney classes of its normal bundle vanish, then either $\alpha(d) \leq n$ or M is cobordant to zero. These conjectures were subsequently proved in [163, 164, 119].

Appendix B

Catalan numbers and Xq^k

There is an alternative proof of Proposition 2.4.2, based on an exercise on page 154 of [122], which links the conjugate Steenrod square to Catalan numbers. Our aim is to prove the following statement.

Proposition B.0.1. *$Xq^k(x^d) = \binom{d+2k}{k}x^{d+k}$, where the binomial coefficients are taken mod 2.*

This surprisingly simple formula does not seem to be well known. The proof is based on an integral relationship between the formal power series

$$C = \sum_{k\geq 0} \binom{2k}{k}\frac{x^k}{(k+1)}, \quad F(d) = \sum_{k\geq 0}\binom{d+2k}{k}x^k,$$

where C is the generating function for the Catalan numbers.

Proposition B.0.2. *There is an identity $F(d+1) = CF(d)$.*

Proof. ¿From Pascal's triangle we have

$$\binom{d+2k}{k} = \binom{d-1+2k}{k} + \binom{d-1+2k}{k-1}$$

and this immediately gives rise to the recurrence relation

$$xF(d+1) = F(d) - F(d-1).$$

We now argue by induction on d assuming that the result of the proposition is proved up to some $d \geq 1$. Then $xF(d+1) = F(d) - F(d-1)$ and multiplying through by C gives $xCF(d+1) = CF(d) - CF(d-1) = F(d+1) - F(d)$ by the inductive assumption. Hence $xCF(d+1) = xF(d+2)$ by the recurrence relation and $CF(d+1) = F(d+2)$ as required in the inductive step. It remains to verify the statement of the proposition for $d = 0, 1$. For this we need to know the following standard fact about the Catalan series [138]

Proposition B.0.3. *The Catalan series satisfies the identity*

$$xC^2 = C - 1.$$

Multiplying through the identity by $F(0)$ and assuming $F(1) = CF(0)$ gives

$$xCF(1) = xC^2F(0) = CF(0) - F(0) = F(1) - F(0) = xF(2),$$

by the recurrence relation. Hence $F(2) = CF(1)$ and it remains to show that $F(1) = CF(0)$. From the definition of $F(d)$ it clear that

$$F(0) = C + x\frac{dC}{dx}, \quad F(1) = C + 2x\frac{dC}{dx}.$$

Now we need to show that

$$C^2 + xC\frac{dC}{dx} = C + 2x\frac{dC}{dx}.$$

Differentiating the identity of Proposition B.0.3 we have

$$C^2 + 2xC\frac{dC}{dx} = \frac{dC}{dx}.$$

Solving for $\frac{dC}{dx}$, the problem reduces to showing that

$$\frac{C^2 - C}{2x - xC} = \frac{C^2}{1 - 2xC}.$$

the rest is elementary algebra using the identity of Proposition B.0.3 again. □

It follows by induction on d in Proposition B.0.2 that

$$F(d) = F(0)C^d.$$

We can now prove Proposition B.0.1. We recall that the total conjugate Steenrod operation Xq is multiplicative by definition and

$$Xq(x) = x + x^2 + \cdots + x^{2^k} + \cdots.$$

It is known that the Catalan numbers c_k are even except when k has the form $2^r - 1$, so that working mod 2 we have

$$C = 1 + x + x^3 + \cdots + x^{2^r - 1} + \cdots.$$

Hence $Xq(x) = xC$ mod 2. Also the the binomial coefficients $\binom{2k}{k}$ are even for $k > 0$. Hence $F(0) = 1$ mod 2 and we obtain

$$Xq(x^d) = Xq(x)^d = (xC)^d = x^dC^d = x^dF(d)$$

mod 2. Selecting the terms of grading $k + d$ establishes Proposition B.0.1.

Appendix C

Blocks and ω-vectors

Blocks and ω-vectors (under various names) are well known in the literature on game theory. For example, Hardy and Wright [54] explain the winning strategy in the game of NIM in essentially the following manner. First we recall that the game is between two players and starts with at least two non-empty piles of objects. A move consists in taking away at least one object from a single pile of the player's choice. The players take it in turn to make a move. NIM is usually played as a misère game in which the loser is the player forced to remove the last object but in the normal game a player wins by taking all the objects when only one pile remains. We asssume that the the game is normal. The state of the game at any time may be represented by the nonzero block B whose rows are the reversed binary expansions of the numbers of objects in the piles. We call B a losing block if all entries of $\omega(B)$ are even, otherwise B is a winning block. For example a 1-block is a winning block and a 2-block with equal rows is a losing block.

Proposition C.0.4. *Any move on a losing block produces a winning block.*

Proof. Let $r > 0$ be a number with $\mathrm{len}(\omega(r)) = l$. Then for any number $d \geq r$ we have $\omega(d-r)_l = 1 - \omega(d)_l$. Hence removing r objects from a pile of d objects changes the parity of column l of the corresponding block. In particular a losing block becomes a winning block. □

Proposition C.0.5. *There is a move on a winning block which either wins the game or produces a losing block.*

Proof. Let B be a winning block. If B has only one row then removing all the objects wins the game. Suppose on the other hand that B has at least two (nonzero) rows. Let $\omega(B)$ have its last odd entry in column c. Then $B_{i,c} = 1$ for some row position i. Let B' be the block obtained from B by putting $B'_{i,j} = 1 - B_{i,j}$ for each j where where $\omega(B)_j$ is odd. and leaving the other entries of B unchanged. Then B' is a losing block. Furthermore $B'_{i,j} = B_{i,j}$ if $j > c$ and $B'_{i,c} = 0$. Hence, if d, d' are the numbers represented by the rows i of B and

B', then $d' < d$. It follows that B' is obtained from B by removing the positive number $d - d'$ of objects from the ith pile. □

¿From the above propositions we see that that a knowledgeable player presented at some stage in the game with a winning block has a winning strategy. For example, in a game with two piles, the strategy is to equalize the numbers of objects in the piles so presenting the opposing player with a losing block.

lock and a 2-block with equal rows is a losing block.

Proposition C.0.6. *Any move on a losing block produces a winning block.*

Proof. Let $r > 0$ be a number with $\text{len}(\omega(r)) = l$. Then for any number $d \geq r$ we have $\omega(d - r)_l = 1 - \omega(d)_l$. Hence removing r objects from a pile of d objects changes the parity of column l of the corresponding block. In particular a losing block becomes a winning block. □

Proposition C.0.7. *There is a move on a winning block which either wins the game or produces a losing block.*

Proof. Let B be a winning block. If B has only one row then removing all the objects wins the game. Suppose on the other hand that B has at least two (nonzero) rows. Let $\omega(B)$ have its last odd entry in column c. Then $B_{i,c} = 1$ for some row position i. Let B' be the block obtained from B by putting $B'_{i,j} = 1 - B_{i,j}$ for each j where where $\omega(B)_j$ is odd. and leaving the other entries of B unchanged. Then B' is a losing block. Furthermore $B'_{i,j} = B_{i,j}$ if $j > c$ and $B'_{i,c} = 0$. Hence, if d, d' are the numbers represented by the rows i of B and B', then $d' < d$. It follows that B' is obtained from B by removing the positive number $d - d'$ of objects from the ith pile. □

¿From the above propositions we see that that a knowledgeable player presented at some stage in the game with a winning block has a winning strategy. For example, in a game with two piles, the strategy is to equalize the numbers of objects in the piles so presenting the opposing player with a losing block.

Appendix D

Topological applications

In the 1970s topologists began to investigate the problem of splitting suspensions of classifying spaces of Lie groups into wedge sums, so giving rise to a splitting of their cohomology into summands invariant under the action of the Steenrod operations. In the 1980s the work of Mitchell and Priddy [99] highlighted the role played by the Steinberg idempotent in splitting off certain summands of the suspended classifying space $\Sigma(BV)$, where V is an n-dimensional vector space over $\mathbb{F}_2$, regarded as an elementary abelian 2-group. The cohomology $H^*(BV)$ over $\mathbb{F}_2$ may be identified with $P(n)$. Adams, Gunawardena and Miller [2] showed that the only graded linear transformations of the $P(n)$ which commute with the Steenrod operations are the ones given by matrix substitution. It was proved in [55] that the complete decomposition of the $\mathcal{A}_2$-module $P(n)$ into indecomposable summands is obtained from a maximal set of orthogonal idempotents in $\mathbb{F}_2 M(n)$. In theory the hit problem can be treated one summand at a time, but in practice it is difficult to gain a sufficiently good grip on the idempotents, apart from the Steinberg case, to make this approach workable. However, a less refined but useful decomposition of $P(n)$ can be obtained from the group algebras of certain subgroups of $GL(n)$, for example cyclic groups, and we shall look at examples later.

The $\mathbb{F}_2 GL(n)$-modules $Q^d(n)$ are of interest in representation theory. It is known that every irreducible representation of $GL(n)$ occurs as a composition factor in $P^d(n)$ for some d. We shall see that the same is true for $Q^d(n)$. For $n > 1$, $\dim(P^d(n)) \to \infty$ as $d \to \infty$, but we shall see that $\dim(Q^d(n))$ is bounded by a function of n independent of d. Various classical $\mathbb{F}_2 GL(n)$-modules such as flag modules, Weyl modules and the Steinberg module appear in the modules $Q^d(n)$.

The interplay of the Steenrod operations and the action of matrices in $P(n)$ leads naturally to a study of algebras of invariants of subgroups of $GL(n)$. This broadens the scope of the hit problem and justifies the purely algebraic approach to the study of modules over the Steenrod algebra which are not realisable as the cohomology of topological spaces [136]. A further motivation for the hit problem

arises from Singer's transfer map [133] linking the Adams spectral sequence at level n with the $GL(n)$-invariants of the cohits $Q(n)$.

Bibliography

[1] J. F. Adams, On the structure and applications of the Steenrod algebra, Comment. Math. Helv. **32** (1958), 180–214.

[2] J. F. Adams, J. Gunawardena and H. Miller, The Segal conjecture for elementary abelian 2-groups, Topology **24** (1985), 435–460.

[3] J. F. Adams and H. R. Margolis, Sub-Hopf algebras of the Steenrod algebra, Proc. Cambridge Philos. Soc. **76** (1974), 45–52.

[4] J. Adem, The iteration of Steenrod squares in algebraic topology, Proc. Nat. Acad. Sci. U.S.A. **38** (1952), 720–726.

[5] J. Adem, The relations on Steenrod powers of cohomology classes, in Algebraic Geometry and Topology, a symposium in honour of S. Lefschetz, 191–238, Princeton Univ. Press, Princeton, NJ 1957.

[6] J. L. Alperin and Rowen B. Bell, Groups and Representations, Graduate Texts in Mathematics 162, Springer-Verlag, New York, 1995.

[7] M. A. Alghamdi, M. C. Crabb and J. R. Hubbuck, Representations of the homology of BV and the Steenrod algebra I, Adams Memorial Symposium on Algebraic Topology vol. 2, London Math. Soc. Lecture Note Ser. **176**, Cambridge Univ. Press 1992, 217–234.

[8] D. J. Anick and F. P. Peterson, A^*-annihilated elements in $H_*(\Omega\Sigma(\mathbb{R}P^\infty))$, Proc. Amer. Math. Soc. **117** (1993), 243–250.

[9] D. Arnon, Monomial bases in the Steenrod algebra, Journal of Pure and Applied Algebra **96** (1994), 215–223.

[10] M. F. Atiyah and F. Hirzebruch, Cohomologie-Operationen und charakteristische Klassen, Math. Z. **77** (1961), 149–187.

[11] M. G. Barratt and H. Miller, On the anti-automorphism of the Steenrod algebra, Contemp. Math. **12** (1981), 47–52.

[12] D. Bausum, An expression for $\chi(Sq^m)$, Preprint, Minnesota University (1975).

[13] D. J. Benson, Representations and cohomology II: Cohomology of groups and modules, Cambridge Studies in Advanced Mathematics **31**, Cambridge University Press (1991).

[14] D. J. Benson and V. Franjou, Séries de compositions de modules instables et injectivité de la cohomologie du groupe $\mathbb{Z}/2$, Math. Zeit **208** (1991), 389–399.

[15] P. C. P. Bhatt, An interesting way to partition a number, Information Processing letters **71** (1999), 141–148.

[16] J. M. Boardman, Modular representations on the homology of powers of real projective spaces, Algebraic Topology, Oaxtepec 1991, Contemp. Math. **146** (1993), 49–70.

[17] Kenneth S. Brown, Buildings, Springer-Verlag, New York, 1989.

[18] R. R. Bruner, Lê M Hà, Nguyen H. V. Hung,, On the algebraic transfer, Trans. Amer. Math. Soc. **357** (2005), 473–487.

[19] S. R. Bullett and I. G. Macdonald, On the Adem relations, Topology **21** (1982), 329–332.

[20] H. E. A. Campbell and P. S. Selick, Polynomial algebras over the Steenrod algebra, Comment. Math. Helv. **65** (1990), 171–180.

[21] D. Carlisle, P. Eccles, S. Hilditch, N. Ray, L. Schwartz, G. Walker, R. Wood, Modular representations of $GL(n,p)$, splitting $\Sigma(\mathbb{C}P^\infty \times \ldots \times \mathbb{C}P^\infty)$, and the β-family as framed hypersurfaces, Math. Zeit. **189** (1985), 239–261.

[22] D. P. Carlisle and N. J. Kuhn, Subalgebras of the Steenrod algebra and the action of matrices on truncated polynomial algebras, Journal of Algebra **121** (1989), 370–387.

[23] D. P. Carlisle and N. J. Kuhn, Smash products of summands of $B(\mathbb{Z}/p)^n_+$, Contemp. Math. **96** (1989), 87–102.

[24] David P. Carlisle and Grant Walker, Poincaré series for the occurrence of certain modular representations of $GL(n,p)$ in the symmetric algebra, Proc. Roy. Soc. Edinburgh **113A** (1989), 27–41.

[25] D. P. Carlisle and R. M. W. Wood, The boundedness conjecture for the action of the Steenrod algebra on polynomials, Adams Memorial Symposium on Algebraic Topology, Vol. 2, London Math. Soc. Lecture Note Ser. **176**, Cambridge University Press, (1992), 203–216.

[26] D. P. Carlisle, G. Walker and R. M. W. Wood, The intersection of the admissible basis and the Milnor basis of the Steenrod algebra, Journal of Pure and Applied Algebra **128** (1998), 1–10.

[27] H. Cartan, Une théorie axiomatique des carrés de Steenrod, C.R. Acad. Sci. Paris **230** (1950), 425–427.

[28] H. Cartan, Sur l'itération des opérations de Steenrod, Comment. Math. Helv. **29** (1955), 40–58.

[29] R. W. Carter, Representation theory of the 0-Hecke algebra, J. of Algebra **104** (1986), 89–103.

[30] R. W. Carter and G. Lusztig, Modular representations of finite groups of Lie type, Proc. London Math. Soc. (3) **32** (1976), 347–384.

[31] D. E. Cohen, On the Adem relations, Proc. Camb. Phil. Soc. **57** (1961), 265–267.

[32] M. C. Crabb, M. D. Crossley and J. R. Hubbuck, K-theory and the anti-automorphism of the Steenrod algebra, Proc. Amer. Math. Soc. **124** (1996), 2275–2281.

[33] M. C. Crabb and J. R. Hubbuck, Representations of the homology of BV and the Steenrod algebra II, Algebraic Topology: new trends in localization and periodicity (Sant Feliu de Guixols, 1994) 143–154, Progr. Math. **136**, Birkhaüser, Basel 1996.

[34] M. D. Crossley, $\mathcal{A}_p$-annihilated elements of $H_*(\mathbb{C}P^\infty \times \mathbb{C}P^\infty)$, Math. Proc. Camb. Phil. Soc. **120** (1996), 441–453.

[35] M. D. Crossley, H^*V is of bounded type over $\mathcal{A}_p$, Group Representations: Cohomology, group actions, and topology (Seattle 1996), Proc. Sympos. Pure Math. **63**, Amer. Math. Soc. (1998), 183–190.

[36] M. D. Crossley, $\mathcal{A}(p)$ generators for H^*V and Singer's algebraic transfer, Math. Zeit. **230** (1999), No. 3, 401–411.

[37] M. D. Crossley, Monomial bases for $H^*(\mathbb{C}P^\infty \times \mathbb{C}P^\infty)$ over $\mathcal{A}(p)$, Trans. Amer. Math. Soc. **351** (1999), No. 1, 171–192.

[38] Charles W. Curtis and Irving Reiner, Representation theory of finite groups and associative algebras, Wiley, New York 1962.

[39] D. M. Davis, The antiautomorphism of the Steenrod algebra, Proc. Amer. Math. Soc. **44** (1974), 235–236.

[40] D. M. Davis, Some quotients of the Steenrod algebra, Proc. Amer. Math. Soc. **83** (1981), 616–618.

[41] J. Dieudonné, A history of algebraic and differential topology 1900–1960, Birkhäuser Basel, 1989.

[42] A. Dold, Über die Steenrodschen Kohomologieoperationen, Annals of Math. **73** (1961), 258–294.

[43] Stephen Doty and Grant Walker, The composition factors of $\mathbb{F}_p[x_1, x_2, x_3]$ as a $GL(3, \mathbb{F}_p$-module, J. of Algebra **147** (1992), 411–441.

[44] Stephen Doty and Grant Walker, Modular symmetric functions and irreducible modular representations of general linear groups, J. Pure App. Algebra **82** (1992), 1–26.

[45] Stephen Doty and Grant Walker, Truncated symmetric powers and modular representations of GL_n, Math. Proc. Cambridge Philos. Soc. **119** (1996), 231–242.

[46] V. Franjou and L. Schwartz, Reduced unstable A-modules and the modular representation theory of the symmetric groups, Ann. Scient. Ec. Norm. Sup. **23** (1990), 593–624.

[47] W. Fulton, Young Tableaux, London Math. Soc. Stud. Texts **35**, Cambridge Univ. Press, 1997.

[48] A. M. Gallant, Excess and conjugation in the Steenrod algebra, Proc. Amer. Math. Soc. **76** (1979), 161–166.

[49] V. Giambalvo, N. H. V. Hung and F. P. Peterson, $H^*(RP^\infty \times \cdots \times RP^\infty)$ as a module over the Steenrod algebra, Hilton Symposium 1993, Montreal, CRM Proc. Lecture Notes **6** , Amer. Math. Soc. Providence RI (1994), 133–140.

[50] V. Giambalvo and F. P. Peterson, On the height of Sq^{2^n}, Contemp. Math. **181** (1995), 183–186.

[51] V. Giambalvo and F. P. Peterson, The annihilator ideal of the action of the Steenrod algebra on $H^*(\mathbb{R}P^\infty)$, Topology Appl. **65** (1995), 105–122.

[52] M. Y. Goh, P. Hitczenko and Ali Shokoufandeh, s-partitions, Information Processing Letters **82** (2002), 327–329.

[53] B. Gray, Homotopy Theory, Academic Press, New York, 1975.

[54] G. H. Hardy and E. M. Wright, An introduction to the theory of numbers, Oxford at the Clarendon Press, 1979.

[55] J. C. Harris and N. J. Kuhn, Stable decomposition of classifying spaces of finite abelian p-groups, Math. Proc. Cambridge Philos. Soc. **103** (1988), 427–449.

[56] T. J. Hewett, Modular invariant theory of parabolic subgroups of $GL_n(\mathbf{F}_q)$ and the associated Steenrod modules, Duke Math. J. **82** (1996), 91–102.

[57] Florent Hivert and Nicolas M. Thiéry, The Hecke group algebra of a Coxeter group and its representation theory, J. Algebra bf 321, No. 8, (2009), 2230–2258.

[58] H. Hopf, Über die abbildungen von Sphären auf Sphären niedrigerer Dimension, Fund. Math. **25** (1935), 427–440.

[59] J. E. Humphreys, Modular Representations of Finite Groups of Lie Type, London Math. Soc. Lecture Note Ser. **326**, Cambridge Univ. Press 2005.

[60] N. H. V. Hung, The action of Steenrod squares on the modular invariants of linear groups, Proc. Amer. Math. Soc. **113** (1991), 1097–1104.

[61] N. H. V. Hung, The action of the mod p Steenrod operations on the modular invariants of linear groups, Vietnam J. Math. **23** (1995), 39–56.

[62] N. H. V. Hung, The cohomology of the Steenrod algebra and representations of the general linear groups, Trans. Amer. Math. Soc. **357** (2005), 4065–4089.

[63] N. H. V. Hung and Pham Anh Minh, The action of the mod p Steenrod operations on the modular invariants of linear groups, Vietnam J. Math. **23** (1995), 39–56.

[64] N. H. V. Hung and Tran Ngoc Nam, The hit problem for modular invariants of linear groups, J. Algebra **246** (2001), 367–384.

[65] N. H. V. Hung and Tran Ngoc Nam, The hit problem for the Dickson algebra, Trans. Amer. Math. Soc. **353** (2001), 5029–5040.

[66] N. H. V. Hung and F. P. Peterson, $\mathcal{A}$-generators for the Dickson algebra, Trans. Amer. Math. Soc. **347** (1995), 4687–4728.

[67] N. H. V. Hung and F. P. Peterson, Spherical classes and the Dickson algebra, Math. Proc. Camb. Phil. Soc. **124** (1998), 253–264.

[68] Masateru Inoue, $\mathcal{A}$-generators of the cohomology of the Steinberg summand $M(n)$, Contemp. Math. **293** (2002), 125–139.

[69] Masateru Inoue, Generators of the cohomology of $M(n)$ as a module over the odd primary Steenrod algebra, J. Lond. Math. Soc. **75**, No. 2, (2007), 317–329.

[70] G. D. James and A. Kerber, The representation theory of the symmetric group, Encyclopaedia of Mathematics, vol. **16**, Addison-Wesley, Reading, Mass. 1981.

[71] Ali Sarbaz Janfada, A criterion for a monomial in $P(3)$ to be hit, Math. Proc. Cambridge Phil. Soc. **145** (2008), 587–599.

[72] Ali Sarbaz Janfada, A note on the unstability conditions of the Steenrod squares on the polynomial algebra, J. Korean Math. Soc **46** (2009), No. 5, 907–918.

[73] A. S. Janfada and R. M. W. Wood, The hit problem for symmetric polynomials over the Steenrod algebra, Math. Proc. Camb. Phil. Soc. **133** (2002), 295–303.

[74] A. S. Janfada and R. M. W. Wood, Generating $H^*(BO(3), \mathbb{F}_2)$ as a module over the Steenrod Algebra, Math. Proc. Camb. Phil. Soc. **134** (2003), 239–258.

[75] M. Kameko, Products of projective spaces as Steenrod modules, Ph.D. thesis, Johns Hopkins Univ., 1990.

[76] M. Kameko, Generators of the cohomology of BV_3, J. Math. Kyoto Univ. **38** (1998), 587–593.

[77] M. Kameko, Generators of the cohomology of BV_4, preprint, Toyama Univ. 2003.

[78] M. Kaneda, M. Shimada, M. Tezuka and N. Yagita, Representations of the Steenrod algebra, J. of Algebra **155** (1993), 435–454.

[79] C. Kassel, Quantum Groups, Graduate Texts in Mathematics **155**, Springer-Verlag 1995.

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

[81] L. Kristensen, On a Cartan formula for secondary cohomology operations, Math. Scand. **16** (1965), 97–115.

[82] N. J. Kuhn, Generic representations of the finite general linear groups and the Steenrod algebra: I, Amer. J. Math. **116** (1994), 327–360; II, K-theory **8** (1994), 395–428; III, K-theory **9** (1995), 273–303.

[83] N. J. Kuhn and S. A. Mitchell, The multiplicity of the Steinberg representation of $GL_n\mathbb{F}_q$ in the symmetric algebra, Proc. Amer. Math. Soc. **96** (1986), 1–6.

[84] M. Latapy, Partitions of an integer into powers, in Discrete Mathematics and Theoretical Computer Science Proceedings, Paris 2001, 215–228.

[85] Z. Li, Product formulas for Steenrod operations, Proc. Edinburgh Math. Soc. **38** (1995), 207–232.

[86] L. Lomonaco, The iterated total squaring operation, Proc. Amer. Math. Soc. **115** (1992), 1149–1155.

[87] I. G. Macdonald, Symmetric Functions and Hall Polynomials (second edition), Oxford mathematical monographs, Clarendon Press, Oxford 1995.

[88] H. Margolis, Spectra and the Steenrod algebra, North Holland Math Library, vol. 29, Elsevier, Amsterdam (1983).

[89] J. P. May, A general algebraic approach to Steenrod operations, The Steenrod Algebra and its Applications, Lecture Notes in Mathematics 168, Springer-Verlag (1970), 153–231.

[90] Dagmar M. Meyer, Stripping and conjugation in the Steenrod algebra and its dual, Homology, Homotopy and Applications **2** (2000), 1–16.

[91] Dagmar M. Meyer, Hit polynomials and excess in the mod p Steenrod algebra, Proc. Edinburgh Math. Soc. (2) **44** (2001), 323–350.

[92] Dagmar M. Meyer and Judith H. Silverman, Corrigendum to 'Hit polynomials and conjugation in the dual Steenrod algebra', Math. Proc. Cambridge Phil. Soc. **129** (2000), 277–289.

[93] J. Milnor, The Steenrod algebra and its dual, Annals of Math. **67** (1958), 150–171.

[94] J. Milnor and J. C. Moore, On the structure of Hopf algebras, Annals of Math. **81** (1965), 211–264.

[95] P. A. Minh and T. T. Tri, The first occurrence for the irreducible modules of the general linear groups in the polynomial algebra, Proc. Amer. Math. Soc. **128** (2000), 401–405.

[96] Pham Anh Minh and Grant Walker, Linking first occurrence polynomials over $\mathbb{F}_p$ by Steenrod operations, Algebr. Geom. Topol. **2** (2002), 563–590.

[97] S. A. Mitchell, Finite complexes with $\mathcal{A}(n)$-free cohomology, Topology **24** (1985), 227–248.

[98] S. A. Mitchell, Splitting $B(\mathbb{Z}/p)^n$ and BT^n via modular representation theory, Math. Zeit. **189** (1985), 285–298.

[99] S. A. Mitchell and S. B. Priddy, Stable splittings derived from the Steinberg module, Topology **22** (1983), 285–298.

[100] K. Mizuno and Y. Saito, Note on the relations on Steenrod squares, Proc. Jap. Acad. **35** (1959), 557–564.

[101] K. G. Monks, Nilpotence in the Steenrod algebra, Bol. Soc. Mat. Mex. **37** (1992), 401–416.

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

[103] K. G. Monks, The nilpotence height of P_t^s, Proc. Amer. Math. Soc. **124** (1996), 1296–1303.

[104] K. G. Monks, Change of basis, monomial relations, and the P_t^s bases for the Steenrod algebra, J. of Pure and Applied Algebra **125** (1998), 235–260.

[105] R. E. Mosher and M. C. Tangora, Cohomology operations and applications in homotopy theory, Harper and Row, New York 1968.

[106] M. F. Mothebe, Generators of the polynomial algebra $\mathbb{F}_2[x_1, \ldots, x_n]$ as a module over the Steenrod algebra, Communications in Algebra **30** (2002), 2213–2228.

[107] M. F. Mothebe, Dimensions of subspaces of the polynomial algebra $\mathbb{F}_2[x_1, \ldots, x_n]$ generated by spikes, Far East J. Math. Sci. **28** (2008), 417–430.

[108] H. Mui, Dickson invariants and Milnor basis of the Steenrod algebra, Topology, theory and application, Coll. Math. Soc. Janos Bolyai **41**, North Holland (1985), 345–355.

[109] Tran Ngoc Nam, $\mathcal{A}$-générateurs génériques pour l'algèbre polynomiale, Adv. Math. **186** (2004), 334–362.

[110] Tran Ngoc Nam, Transfert algébrique et action du groups linéaire sur les puissances divisées, Ann. Inst. Fourier (Grenoble) **58** (2008), 1785–1837.

[111] P. N. Norton, 0-Hecke algebras, J. Austral. Math. Soc. (Ser. A) **27** (1979), 337–357.

[112] S. Papastavridis, A formula for the obstruction to transversality, Topology **11** (1972), 415–416.

[113] David J. Pengelley, Franklin P. Peterson and Frank Williams, A global structure theorem for the mod 2 Dickson algebras, and unstable cyclic modules over the Steenrod and Kudo-Araki-May algebras, Math. Proc. Cambridge Phil. Soc. **129** (2000), 263–275.

[114] D. J. Pengelley and F. Williams, Sheared algebra maps and operation bialgebras for mod 2 homology and cohomolgy, Trans. Amer. Math. Soc. **352** (2000), No. 4, 1453–1492.

[115] D. J. Pengelley and F. Williams, Global Structure of the mod 2 symmetric algebra, $H^*(BO, \mathbb{F}_2)$ over the Steenrod algebra, Algebr. Geom. Topol. **3** (2003), 1119–1138.

[116] D. J. Pengelley and F. Williams, The global Structure of odd-primary Dickson algebras as algebras over the Steenrod algebra, Math. Proc. Cambridge Philos. Soc. **136** (2004), No. 1, 67–73.

[117] D. J. Pengelley and F. Williams, Beyond the hit problem: minimal presentations of odd-primary Steenrod modules, with application to $\mathbb{C}P^\infty$ and BU, Homology, Homotopy and Applications, **9**, No.2, (2007), 363–395.

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

[119] F. P. Peterson, $\mathcal{A}$-generators for certain polynomial algebras, Math. Proc. Camb. Phil. Soc. **105** (1989), 311–312.

[120] F. P. Peterson, Some formulas in the Steenrod algebra, Proc. Amer. Math. Soc. **45** (1974), 291–294.

[121] J. Repka and P. Selick, On the subalgebra of $H_*((\mathbb{R}P^\infty)^n; \mathbb{F}_2)$ annihilated by Steenrod operations, J. Pure Appl. Algebra **127** (1998), 273–288.

[122] J. Riordan, Combinatorial Identities, John Wiley & Sons, New York 1968.

[123] L. Schwartz, Unstable Modules over the Steenrod algebra and Sullivan's fixed point set conjecture, Chicago Lectures in Mathematics, University of Chicago Press, 1994.

[124] J. Segal, Notes on invariant rings of divided powers, CRM Proceedings and Lecture Notes **35**, Invariant Theory in All Characteristics, ed. H. E. A. Campbell and D. L. Wehlau, Amer. Math. Soc. 2004, 229–239.

[125] J.-P. Serre, Cohomologie modulo 2 des complexes d'Eilenberg-MacLane, Comment. Math. Helv. **27** (1953), 198–232.

[126] C. Shengmin and S. Xinyao, On the action of Steenrod powers on polynomial algebras, Proceedings of the Barcelona Conference on Algebraic Topology, Lecture Notes in Mathematics **1509**, Springer-Verlag (1991), 326–330.

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

[128] Judith H. Silverman, Multiplication and combinatorics in the Steenrod algebra, Preprint, University of Michigan (1994).

[129] Judith H. Silverman, Hit polynomials and the canonical antiautomorphism, Proc. Amer. Math. Soc. **123** (1995), 627–637.

[130] Judith H. Silverman, Stripping and conjugation in the Steenrod algebra, J. of Pure and Applied Algebra **121** (1997), 95–106.

[131] Judith H. Silverman, Hit polynomials and conjugation in the the dual Steenrod algebra, Proc. Cambridge Philos. Soc., **123** (1998), 531 – 547.

[132] Judith H. Silverman and William M. Singer, On the action of Steenrod squares on polynomial algebras II, J. Pure and Applied Algebra **98** (1995), 95–103.

[133] William M. Singer, The transfer in homological algebra, Math. Z. **202** (1989), 493–523.

[134] William M. Singer, On the action of Steenrod squares on polynomial algebras, Proc. Amer. Math. Soc. **111** (1991), 577–583.

[135] William M. Singer, Rings of symmetric functions as modules over the Steenrod algebra, Algebr. Geom. Topol. **8** (2008), 541–562.

[136] Larry Smith, Polynomial Invariants of Finite Groups, A. K. Peters, Wellesley, Mass. 1995.

[137] Larry Smith, An algebraic introduction to the Steenrod algebra, in: Proceedings of the School and Conference in Algebraic Topology (Hanoi, 2004), Geometry and Topology Monographs **11** (2007), 327–348.

[138] R. P. Stanley, Enumerative Combinatorics, vol. 2, Cambridge Studies in Advanced Mathematics **62**, Cambridge University Press, (1999).

[139] N. E. Steenrod, Products of cocycles and extensions of mappings, Ann. of Math. **48** (1947), 290–320.

[140] N. E. Steenrod, Reduced powers of cohomology classes, Ann. of Math. **56** (1952), 47–67.

[141] N. E. Steenrod, Homology groups of symmetric groups and reduced power operations, Proc. Nat. Acad. Sci. U.S.A. **39** (1953), 213–217.

[142] N. E. Steenrod and D. B. A. Epstein, Cohomology Operations, Annals of Math. Studies 50, Princeton University Press (1962).

[143] P. D. Straffin, Identities for conjugation in the Steenrod algebra, Proc. Amer. Math. Soc. **49** (1975), 253–255.

[144] N. Sum, On the action of the Steenrod-Milnor operations on the modular invariants of linear groups, Japan J. Math. **18** (1992), 115–137.

[145] N. Sum, On the action of the Steenrod algebra on the modular invariants of special linear group, Acta Math. Vietnam **18** (1993), 203–213.

[146] N. Sum, Steenrod operations on the modular invariants, Kodai Math. J. **17** (1994), 585–595.

[147] N. Sum, On the hit problem for the polynomial algebra in four variables, Preprint, University of Quynhon, Vietnam 2007.

[148] N. Sum, The negative answer to Kameko's conjecture on the hit problem, Preprint, University of Quynhon, Vietnam 2008.

[149] R. Thom, Une théorie intrinsèque des puissances de Steenrod, Colloque de Topologie de Strasbourg, Publication of the Math. Inst. University of Strasbourg (1951).

[150] R. Thom, Éspaces fibrés en sphères et carrés de Steenrod, Ann. Sci. Ec. Norm. Sup. **69** (1952), 109–182.

[151] R. Thom, Quelque propriétés globales des variétés différentiables, Comment. Math. Helv. **28** (1954), 17–86.

[152] Grant Walker, Modular Schur functions, Trans. Amer. Math. Soc. **346** (1994), 569–604.

[153] G. Walker and R. M. W. Wood, The nilpotence height of Sq^{2^n}, Proc. Amer. Math. Soc. **124** (1996), 1291–1295.

[154] G. Walker and R. M. W. Wood, The nilpotence height of P^{p^n}, Math. Proc. Camb. Phil. Soc. **123** (1998), 85–93.

[155] G. Walker and R. M. W. Wood, Linking first occurrence polynomials over $\mathbb{F}_2$ by Steenrod operations, Journal of Algebra **246** (2001), 739–760.

[156] G. Walker and R. M. W. Wood, Young tableaux and the Steenrod algebra, in: Proceedings of the School and Conference in Algebraic Topology (Hanoi, 2004), Geometry and Topology Monographs **11** (2007), 379–397.

[157] G. Walker and R. M. W. Wood, Weyl modules and the mod 2 Steenrod Algebra, J. Algebra **311** (2007), 840–858.

[158] G. Walker and R. M. W. Wood, Flag modules and the hit problem for the Steenrod algebra, Math. Proc. Cambridge Phil. Soc. **147** (2009), 143–171.

[159] C. T. C. Wall, Generators and relations for the Steenrod algebra, Annals of Math. **72** (1960), 429–444.

[160] C. Wilkerson, A primer on the Dickson invariants, Contemp. Math. **19** (1983), 421–434.

[161] R. M. W. Wood, Modular representations of $GL(n, \mathbf{F}_p)$ and homotopy theory, Algebraic topology Göttingen 1984, Lecture Notes in Mathematics **1172**, Springer-Verlag (1985), 188–203.

[162] R. M. W. Wood, Splitting $\Sigma(\mathbb{C}P^\infty \times \ldots \times \mathbb{C}P^\infty)$ and the action of Steenrod squares on the polynomial ring $\mathbf{F}_2[x_1, \ldots, x_n]$, Algebraic Topology Barcelona 1986, Lecture Notes in Mathematics **1298**, Springer-Verlag (1987), 237–255.

[163] R. M. W. Wood, Steenrod squares of Polynomials, Advances in homotopy theory, London Mathematical Society Lecture Notes 139, Cambridge University Press (1989), 173–177.

[164] R. M. W. Wood, Steenrod squares of polynomials and the Peterson conjecture, Math. Proc. Camb. Phil. Soc. **105** (1989), 307–309.

[165] R. M. W. Wood, A note on bases and relations in the Steenrod algebra, Bull. London Math. Soc. **27** (1995), 380–386.

[166] R. M. W. Wood, Differential operators and the Steenrod algebra, Proc. London Math. Soc. **75** (1997), 194–220.

[167] R. M. W. Wood, Problems in the Steenrod algebra, Bull. London Math. Soc. **30** (1998), 194–220.

[168] R. M. W. Wood, Hit problems and the Steenrod algebra, Proceedings of the Summer School 'Interactions between Algebraic Topology and Invariant Theory', a satellite conference of the third European Congress of Mathematics, Ioannina University, Greece (2000), 65–103.

[169] R. M. W. Wood, Invariants of linear groups as modules over the Steenrod algebra, Ingo2003, Invariant Theory and its interactions with related fields, University of Göttingen, (2003).

[170] R. M. W. Wood, The Peterson Conjecture for Algebras of Invariants, Invariant Theory in all characteristics, CRM Proceedings and Lecture Notes, vol 35, Amer. Math. Soc., Providence R.I. (2004), 275–280.

[171] W-T. Wu, Sur les puissances de Steenrod, Colloque de Topologie de Strasbourg, Publication of the Math. Inst. University of Strasbourg (1952).