Symbolic dynamics and rotation symmetric Boolean functions

Alexandru Chirvasitu and Thomas Cusick

Abstract

We identify the weights $wt(f_n)$ of a family $\{f_n\}$ of rotation symmetric Boolean functions with the cardinalities of the sets of n-periodic points of a finite-type shift, recovering the second author's result that said weights satisfy a linear recurrence. Similarly, the weights of idempotent functions f_n defined on finite fields can be recovered as the cardinalities of curves over those fields and hence satisfy a linear recurrence as a consequence of the rationality of curves' zeta functions. Weil's Riemann hypothesis for curves then provides additional information about $wt(f_n)$. We apply our results to the case of quadratic functions and considerably extend the results in an earlier paper of ours.

Key words: shift, subshift, finite type, Weil conjectures, Riemann hypothesis for algebraic varieties, Boolean function, weight, finite field, cyclotomic polynomial, cyclotomic field, Galois group

MSC 2010: 06E30; 37B50; 11G20

Introduction

Boolean functions $f : \{0,1\}^n \to \{0,1\}$ have long held the interest of the cryptographic community due to their many applications to that field; see e.g. [23, 15, 4, 22, 25, 1] (to give just a few examples that could not possibly do the subject justice) or the discussion and numerous references cited in [12].

Among all Boolean functions, the ones with the best cryptographic properties tend to be *balanced*, i.e. take the values $0, 1$ equally many times. For that reason, one is more generally interested in the *weight* $wt(f)$ of a Boolean function, meaning the cardinality of the preimage $f^{-1}(1)$.

We are concerned here with Boolean functions that are *rotation symmetric* in the sense of [24]: those f as above that are invariant under permuting the n variables cyclically. Such a function is expressible as

$$f_n(x_0, \cdots, x_{n-1}) = \sum_{i \bmod n} x_i x_{i+a_1} \cdots x_{i+a_{d-1}},\ x_i \in \{0,1\}. \tag{0-1}$$

In fact, having fixed the a_j, the formula (0-1) gives rise to a *family* of rotation symmetric functions f_n in n variables respectively (see §1.1 below). The starting point for the current paper is the phenomenon constituting the main theorem of [10] (which in turn builds on earlier work in the same direction [11, 6, 3]) whereby the weights $wt(f_n)$ attached to a family of rotation symmetric Boolean functions satisfy a linear recurrence of the form

$$wt(f_{n+N}) = a_{N-1} wt(f_{n+N-1}) + \cdots + a_1 wt(f_n),\ \forall \text{ sufficiently large } n.$$

The motivation here was a desire to understand that recurrence phenomenon in light of other analogous results in the literature to the effect that sequences tracking the sizes of various meaningful sets are linearly recurrent. The general paradigm is that said sequences N_n are collected into

a single mathematical object

$$\zeta(s) := \exp\left(\sum_{n\geq 1} \frac{N_n}{n} s^n\right).$$

called a *zeta function* and the desired recurrence follows from the rationality of that power series since said rationality, in fact, will say even more:

$$N_n = \sum_i \alpha_i^n - \sum_j \beta_j^n$$

for algebraic integers α_i and β_j known as the *characteristic values* of the zeta function.

We consider two instances of this setup, both shedding light on Boolean functions in slightly different ways:

- zeta functions of dynamical systems [5], where N_n is the number of n-periodic points under the iterations of a continuous self-map of a compact space, and
- zeta functions of algebraic varieties [14], with N_n being the number of points of a fixed algebraic variety over the field $GF(q^n)$ with q^n elements for a fixed prime power q.

Section 1 gathers the needed background material on Boolean functions, symbolic dynamics and algebraic geometry. We also describe the irreducible factors of polynomials of the form $x^{2t} - 2^t$ (Proposition 1.8) for later use in Section 4.

In Section 2 we recast rotation symmetric Boolean functions as particular instances of well behaved dynamical systems known as *finite-type shifts*: closed subsets of a Cartesian power $\mathcal{A}^{\mathbb{Z}}$ of a finite alphabet, invariant under the leftward shift of bi-infinite sequences. These are well studied objects with a rich theory, and in particular it is a fact that their zeta functions are rational. Our main result in that section (Theorem 2.7) can be paraphrased as

Theorem 0.1 *For every family of rotation symmetric Boolean functions f_n as in (0-1) there is a finite-type shift with $2^{n+1} - 2wt(f_n)$ n-periodic points for each n.*

In particular, $wt(f_n)$ satisfies a linear recurrence. ■

Section 3 revolves around close cousins of rotation symmetric Boolean functions, definable in Galois-theoretic terms: having fixed a polynomial P with coefficients in the field $GF(2)$, one can consider the family of functions

$$f_n : GF(2^n) \to GF(2),\ \ f_n(x) = \mathrm{Tr}(P(x)). \tag{0-2}$$

These are introduced in [7] and studied there as Boolean functions (which is what they are, having identified $GF(2^n)$ with $GF(2)^n$). The analogue of Theorem 0.1 in this case is almost immediate (Corollary 3.3):

Theorem 0.2 *For a family (0-2) of trace functions there is a plane curve X defined over $GF(2)$ such that $2^{n+1} - 2wt(f_n)$ is the number of points of X over $GF(2^n)$.*

In particular, since zeta functions of algebraic varieties are rational [14], $wt(f_n)$ again satisfies a linear recurrence. ■

We give more precise information on the weights $wt(f_n)$ and the general shape of the recurrence they satisfy in Corollaries 3.6 and 3.7 by computing the genus of the curve X of Theorem 0.2 through successive blowups.

Finally, in Section 4 we go back to the quadratic case analyzed closely in [8]. In that setup Theorem 4.2 gives a close connection between the Boolean and trace sides of the picture. Furthermore, for *monomial* (quadratic, rotation symmetric) functions

$$f_{n,t}(x_i) = \sum_{i \bmod n} x_i x_{i+t}$$

we show in Theorem 4.4 that the characteristic values resulting as in Section 2 from the general theory of finite-type shifts precisely coincide with the eigenvalues (including multiplicities) of the recurrence matrix $R(t)$ for $wt(f_{n,t})$ constructed in [10]. This is a curious instance of consilience, given how different the methods of [10] and Section 2 are.

We hope that the methods of the present paper will not only provide a conceptual explanation for the weight recurrence phenomena so prevalent in the Boolean function literature, but also highlight connections to different areas (symbolic dynamics, algebraic geometry) by bringing to bear tools specific to those fields.

Acknowledgements

A.C. was partially supported by NSF grant DMS-1801011.

1 Preliminaries

Throughout, $GF(q)$ denotes the finite field with q elements. We focus primarily on characteristic-two fields, i.e. $q = 2^n$.

1.1 Boolean functions

We will work with functions defined on either

- tuples of Boolean variables. i.e. elements of $V_n = GF(2)^n$, or
- single finite fields $GF(2^n)$.

We will see that there are strong analogies between these two setups. Specifically, we will construct (following [7]) infinite families of functions f_n of the two types (indexed by the respective n).

To that end, consider a finite collection $\mathcal{C}$ of tuples

$$0 < a_1 < \cdots < a_{d-1}. \tag{1-1}$$

of positive integers for various d.

We then write $f_{\mathcal{C}}$ as a collective label for the functions $f_{\mathcal{C},n}$ defined in either of the two following ways (to be distinguished contextually in the sequel):

Definition 1.1 In *rotation symmetric (or RS) context* $f_{\mathcal{C},n}$ is the rotation symmetric Boolean function $f_{\mathcal{C},n} : V_n \to GF(2)$ obtained as the sum of the *monomial RS (or MRS) functions*

$$(0, a_1, \cdots, a_{d-1}) := \sum_{i \bmod n} x_i x_{i+a_1} \cdots x_{i+a_{d-1}} \tag{1-2}$$

as the tuples (1-1) range over $\mathcal{C}$.

Similarly, in *trace context* $f_{\mathcal{C},n}$ is the function $f_{\mathcal{C},n} : GF(2^n) \to GF(2)$ obtained as the sum of the *monomial trace functions*

$$GF(2^n) \ni x \mapsto \mathrm{Tr}\left(x^{1+2^{a_1}+\cdots+2^{a_{d-1}}}\right)$$

where once more the tuples (1-1) range over $\mathcal{C}$ and Tr denotes the trace $\mathrm{Tr}_n : GF(2^n) \to GF(2)$.♦

Having fixed $\mathcal{C}$, there is a close relationship between $f_{\mathcal{C}}$ in RS and trace context: if $f_n = f_{\mathcal{C},n}$ in RS context then the trace context counterpart $g_n = g_{\mathcal{C},n}$ is denoted in [7, Definition 4.1] by f'_n and can be obtained from f by

$$GF(2^n) \ni x \mapsto f_n(x, x^2, \cdots, x^{2^{n-1}}) \in GF(2).$$

1.2 Symbolic dynamics

For background on the topic we refer to [21, Chapters 1-3,6]. The central notion is

Definition 1.2 Let $\mathcal{A}$ be a finite set (the *alphabet*) and equip the space $\mathcal{A}^{\mathbb{Z}}$ of bi-infinite $\mathcal{A}$-valued sequences

$$\cdots,\ x_{-1},\ x_0,\ x_1,\ \cdots \in \mathcal{A}$$

with its compact Hausdorff product topology. A *shift over* $\mathcal{A}$ is a closed subset $X \subseteq \mathcal{A}^{\mathbb{Z}}$ preserved by the shift operator

$$\sigma : \mathcal{A}^{\mathbb{Z}} \to \mathcal{A}^{\mathbb{Z}}$$

defined by $\sigma(\mathbf{x})_i = x_{i+1}$, where

$$\mathbf{x} = (\cdots, x_{-1}, x_0, x_1, \cdots) \in \mathcal{A}^{\mathbb{Z}}.$$

We often write (X, σ) for a shift, to indicate that we are equipping X with the restriction of the shift map σ.

A *subshift* $(Y, \sigma) \subseteq (X, \sigma)$ is a closed subset $Y \subseteq X$ invariant under σ. ♦

This is equivalent to [21, Definition 1.2.1]. One particular class of shifts we will be interested in is described in [5, Introduction] or [21, Definition 2.1.1].

Before recalling the definition we introduce the following piece of notation: for a finite word

$$w \in \mathcal{A}^* := \text{possibly-empty words with letters in } \mathcal{A}$$

we write X_w for the set of elements in $\mathcal{A}^{\mathbb{Z}}$ that do not contain w as a subword. More generally, for a set $\mathcal{S}$ of words we write

$$X_{\mathcal{S}} := \bigcap_{w \in \mathcal{S}} X_w = \text{sequences containing no element of } \mathcal{S} \text{ as a subword.}$$

It is clear that $X_{\mathcal{S}}$ is invariant under σ and is thus the underlying space of a shift. With this in hand we have

Definition 1.3 A shift (X, σ) over $\mathcal{A}$ is *of finite type* if there is a finite set $\mathcal{S}$ such that $X = X_{\mathcal{S}}$.♦

In other words, the finite-type shifts are those describable by requiring that the sequences in question avoid finitely many patterns (or words) over $\mathcal{A}$.

We also need the following concept (see [5, Introduction] or [21, Definition 6.4.1]).

Definition 1.4 Let (X, σ) be a shift over the alphabet $\mathcal{A}$. For each $n \geq 1$ denote by

$$N_n = N_n(X, \sigma)$$

the number of elements of X left invariant by σ^n (i.e. the sequences in X that are n-periodic).

The *zeta function* of (X, σ) is

$$\zeta(s) = \zeta_{X,\sigma}(s) := \exp\left(\sum_{n\geq 1} \frac{N_n}{n} s^n\right). \qquad \blacklozenge$$

One of the important results on zeta functions is [5, Theorem 1] (see also [21, Theorem 6.4.6]):

Theorem 1.5 *The zeta function of a finite-type shift is of the form*

$$\zeta(s) = \frac{1}{\det(1 - sA)}$$

for some square integer-entry matrix A.

1.3 The Weil conjectures

A good introduction for this is [17, Appendix C].

Let X be an algebraic variety (typically affine or projective) defined over a finite field $F = GF(q)$ for some prime power q. We write $N_n = N_n(X)$ for the number of points of X defined over $GF(q^n)$. Recall ([17, Appendix C.1]):

Definition 1.6 The *zeta function of X* is

$$\zeta(s) = \zeta_X(s) := \exp\left(\sum_{n\geq 1} \frac{N_n}{n} s^n\right). \qquad \blacklozenge$$

Note the analogy to Definition 1.4. The *Weil conjectures* are a series of statements regarding $\zeta_X(s)$ for *smooth* projective varieties X (which thus provide information about the numbers $N_n(X)$). The 'conjecture' moniker is preserved for historical reasons: posed in [28] and resolved for curves in [27], the most difficult of the statements was settled completely in [13], so the "conjectures" are, in fact, theorems. We refer to [17, Appendix C.2] for a more complete historical account.

Since we are concerned primarily with possibly-singular curves X, we phrase the results in the more complete form covered in [2]. Moreover, we focus on the numbers $N_n(X)$ themselves (rather than the zeta function). With this in mind, the relevant statement is [2, Corollary 2.4]:

Theorem 1.7 *Let X be a projective curve over a finite field $GF(q)$, $\widetilde{X} \to X$ the normalization of X, g the genus of $\widetilde{X}$ and Δ the number*

$$\left|\widetilde{X}(\overline{GF(q)}) - X(\overline{GF(q)})\right|.$$

Then, there are Galois-invariant multisets of algebraic integers

- *α_i, $1 \leq i \leq 2g$ with $|\alpha_i| = \sqrt{q}$;*
- *β_j, $1 \leq j \leq \Delta$ with $|\beta_j| = 1$*

such that

$$N_n(X) = q^n + 1 - \sum_{i=1}^{2g} \alpha_i^n - \sum_{j=1}^{\Delta} \beta_j^n. \tag{1-3}$$

1.4 A remark on scaled roots of unity

In the discussion below we will need to analyze the spectrum of a unitary matrix with minimal polynomial $x^{2t} - 2^t$ for $t \geq 2$. To that end, we have to understand the factorization of that polynomial over the integers.

It will be convenient to work with the following polynomials: for a positive integer d, $\Theta_d(x)$ is obtained from the d^{th} cyclotomic polynomial Φ_d by

- substituting x^2 for x: $\Phi_d(x) \mapsto \Phi_d(x^2)$;
- scaling all of the resulting roots by $\sqrt{2}$, i.e. applying the transformation

$$P(x) \mapsto 2^{\frac{\deg P}{2}} P\left(\frac{x}{\sqrt{2}}\right)$$

to $P(x) = \Phi_d(x^2)$.

More generally, we denote the procedure applied here to Φ_d (i.e. the two steps above, in succession) by α. In other words,

$$(\alpha P)(x) = 2^{\deg P} P\left(\frac{x^2}{2}\right) \tag{1-4}$$

and $\alpha\Phi_d = \Theta_d$.

Since $x^{2t} - 2^t$ is nothing but $\alpha(x^t - 1)$, it decomposes as

$$x^{2t} - 2^t = \prod_{d|t} \Theta_d(x).$$

This makes the following result relevant.

Proposition 1.8 *The polynomial Θ_d is irreducible except when the exact power of 2 dividing d is 4, in which case its irreducible factor decomposition is*

$$\Theta_d(x) = P(x)P(-x)$$

for some irreducible polynomial P.

Proof Let Δ_d be the set of primitive d^{th} roots of unity and Δ_d^{-2} its preimage through squaring (i.e. Δ_d^{-2} is the set of roots of $\Phi_d(x^2)$). Let also G be the absolute Galois group $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. We have to argue that $\sqrt{2}\Delta_d^{-2}$

- breaks up into two G-orbits when $d = 4(2e+1)$;
- is a single G-orbit otherwise.

The situation is qualitatively different depending on the parity of d:

Case 1: odd d. We then have the following disjoint unions

$$\Delta_d^{-2} = \Delta_d \sqcup \Delta_{2d} = \Delta_d \sqcup -\Delta_d$$

and the conclusion follows from the fact that the fields $\mathbb{Q}(\sqrt{2})$ and $\mathbb{Q}(\Delta_d \cup -\Delta_d)$ are linearly disjoint and hence the Galois group of their compositum is simply the product of their respective Galois groups. This affords us the choice to send a fixed primitive d^{th} root of unity to any other such root and $\sqrt{2}$ to $\pm\sqrt{2}$.

Case 2: even d. This time around

$$\Delta_d^{-2} = \Delta_{2d}.$$

We have $\sqrt{2} \in \mathbb{Q}(\Delta_8)$. Since $\mathbb{Q}(\Delta_e)$ and $\mathbb{Q}(\Delta_f)$ are linearly disjoint when e and f are coprime (e.g. [20, §IV.1, Theorem 2]), we have $\sqrt{2} \notin \mathbb{Q}(\Delta_{2d})$ unless d is divisible by 4 and we can then repeat the argument used in Case 1.

It thus remains to treat the case $4|d$. The linear disjointness of cyclotomic fields generated by roots of unity of coprime orders allows us to restrict our attention to the case when $d = 2^u$ for some $u \geq 2$ (and hence the field whose Galois group we are interested in is $\mathbb{Q}(\Delta_{2^{u+1}})$).

When $u = 2$ (corresponding to the case when the exact power of 2 dividing the original d was 4) one checks immediately that the four-element set $\sqrt{2}\Delta_8$ decomposes into two Galois orbits, namely

$$\{1 \pm i\} \text{ and } \{-1 \pm i\}.$$

When $u \geq 3$ (and hence 8 divides $d = 2^u$) the difference from the preceding discussion is that now $\sqrt{2}$ is a sum of *even* powers of a fixed primitive d^{th} root of unity ζ, so the image of $\sqrt{2}$ through a Galois group element $\zeta \mapsto \pm\zeta^b$ depends only on b (and not on the sign). If $\zeta \mapsto \zeta^b$ fixes $\sqrt{2}$ then that same Galois group element maps $\sqrt{2}\zeta \mapsto \sqrt{2}\zeta^b$. If, on the other hand, we have

$$\zeta \mapsto \zeta^b, \ \sqrt{2} \mapsto -\sqrt{2}$$

then the *other* Galois group element $\zeta \mapsto -\zeta^b$ will map

$$\sqrt{2}\zeta \mapsto \sqrt{2}\zeta^b. \tag{1-5}$$

Either way, the two arbitrary elements of $\sqrt{2}\Delta_{2d}$ appearing in (1-5) are in the same Galois orbit.■

2 Rotation-symmetric functions as dynamical systems

Let $f = f_{\mathcal{C}}$ be a family of Boolean RS functions associated to a collection $\mathcal{C}$ of tuples (1-1) as in Definition 1.1. As before, we write f_n for f specialized to the n-dimensional vector space $V_n \cong GF(2)^n$ over $GF(2)$.

To f we can also associate polynomial functions $P_{f,n} : V_n \to V_n$ defined for $f = (0, a_1, \cdots, a_{d-1})$ by

$$V_n \ni (x_0, \cdots, x_{n-1}) \mapsto (x_0 x_{a_1} \cdots x_{a_{d-1}}, \ x_1 x_{a_1+1} \cdots x_{a_{d-1}+1}, \ \cdots) \in V_n \tag{2-1}$$

and in general by extending this definition additively. If we now denote the coordinate-sum map

$$V_n \ni (x_0, \cdots, x_{n-1}) \mapsto \sum x_i \in GF(2)$$

by $\mathrm{Tr} = \mathrm{Tr}_n$ then we have

$$f_n = \mathrm{Tr}_n \circ P_{f,n}. \tag{2-2}$$

We make note of the following elementary linear algebra fact whose proof we omit.

Lemma 2.1 *Let k be a field and n a positive integer. A vector $\mathbf{x} \in k^n$ has vanishing sum of coordinates if and only if $x = y - \sigma y$ for some $y \in k^n$, where σ is the rotation operator on k^n defined by (2-4).*

Remark 2.2 Lemma 2.1 is an analogue of the celebrated *Hilbert theorem 90*, in its additive version: if $K \subset L$ is a Galois extension with cyclic Galois group $\langle\sigma\rangle$ then an element $x \in L$ has vanishing trace if and only if

$$x = y - \sigma y$$

for some $y \in L$. ♦

Lemma 2.3 *For $\mathbf{x} \in V_n$ we have $f_n(\mathbf{x}) = 0$ if and only if*

$$P_{f,n}(\mathbf{x}) = \mathbf{y} - \sigma\mathbf{y} \tag{2-3}$$

for some $\mathbf{y} \in V_n$.

Proof This is immediate from Lemma 2.1 and (2-2). ■

Corollary 2.4 *The number of zeros of f_n (i.e. $2^n - wt(f_n)$) is half the number of solutions $(\mathbf{x}, \mathbf{y}) \in V_n^2$ to the equation (2-3).*

Proof This is immediate from Lemma 2.3, since having fixed $\mathbf{x}$, a solution $\mathbf{y}$ to (2-3) is uniquely determined up to translation by the all-1 vector $\mathbf{1} = (1, 1 \cdots)$. ■

Consider embeddings $V_n \to V_{dn}$ for all positive integers d, n defined by

$$\iota = \iota_{n,dn} : V_n \ni (x_0, \cdots, x_{n-1}) \mapsto (x_0, x_1, \cdots, x_{n-1}, x_0, x_1, \cdots) \in V_{dn}.$$

If we equip each V_n with its rotation operator

$$\sigma : (x_0, \cdots, x_{n-1}) \mapsto (x_1, \cdots, x_{n-1}, x_0) \tag{2-4}$$

then the ι embeddings intertwine the respective rotations.

The spaces V_n with connecting maps $\iota_{n,dn}$ form a diagram in the category of vector spaces whose colimit (i.e. union) we denote by V_∞. The latter is nothing but the space of bi-infinite periodic sequences over $GF(2)$. Furthermore, the rotation operators σ on the various V_n lift precisely to the shift on V_∞ in the sense of §1.2 (again denoted by σ).

Remark 2.5 The analogy noted in Remark 2.2 extends further: the $\iota : V_n \to V_{dn}$ parallel the inclusions $GF(2^n) \subset GF(2^{dn})$ of finite fields, the shift operator σ on V_∞ is similar in spirit to the Frobenius automorphism $x \mapsto x^2$ of the algebraic closure $\overline{GF(2)}$, etc. ♦

The polynomial functions $P_{f,n} : V_n \to V_n$ fit into commutative diagrams

$$\begin{array}{ccccc} & \overset{\iota}{\nearrow} & V_{dn} & \overset{P_{f,dn}}{\searrow} & \\ V_n & & & & V_{dn} \\ & \underset{P_{f,n}}{\searrow} & V_n & \underset{\iota}{\nearrow} & \end{array}$$

and hence give rise to a map $P_f : V_\infty \to V_\infty$. This is significant because it will allow us, in a sense, to lift the zero-counting for f_n from V_n to V_∞ via Lemma 2.3: while the trace map Tr does not make sense on V_∞, the equation

$$P_f(\mathbf{x}) = \mathbf{y} - \sigma\mathbf{y} \tag{2-5}$$

does. In fact, we can do more: the definition of P_f extends in the obvious fashion to a self-map of the entire sequence space $\Sigma := GF(2)^{\mathbb{Z}}$ by mimicking the definition in (2-1) for monomials and then extending additively, as before. Note that P_f will then be a shift intertwiner, in the sense that

$$\begin{array}{ccccc} & \overset{\sigma}{\nearrow} & \Sigma & \overset{P_f}{\searrow} & \\ \Sigma & & & & \Sigma \\ & \underset{P_f}{\searrow} & \Sigma & \underset{\sigma}{\nearrow} & \end{array}$$

commutes.

We are now in a position to associate a shift (X_f, σ) to each RS function f: define X_f to be the subspace of

$$(GF(2) \times GF(2))^{\mathbb{Z}} \cong GF(2)^{\mathbb{Z}} \times GF(2)^{\mathbb{Z}}$$

consisting of the $(\mathbf{x}, \mathbf{y})$ satisfying (2-5). The shift map σ on X_f will simply be the restriction of diagonal (σ, σ) to

$$X_f \subset GF(2)^{\mathbb{Z}} \times GF(2)^{\mathbb{Z}}.$$

The following remark captures the relationship between the zeros of f_n and the shift (X_f, σ).

Lemma 2.6 *With the notation above we have*

$$2^{n+1} - 2wt(f_n) = N_n(X_f, \sigma)$$

with N_n denoting number of fixed points of σ^n, as in §1.2.

Proof This follows from Corollary 2.4 after noting that

- $V_\infty \subset GF(2)^{\mathbb{Z}}$ consists precisely of the periodic sequences, so the elements $(\mathbf{x}, \mathbf{y})$ contributing to N_n belong to $V_\infty \times V_\infty$.
- V_n is identifiable with the fixed-point set of σ^n in V^∞. ■

We can now analyze the shift (X_f, σ) for the purpose of extracting interesting properties for the function $n \mapsto wt(f_n)$ via Lemma 2.6.

Theorem 2.7 *For any RS function f the associated shift (X_f, σ) is of finite type.*

Proof Recall that by definition,

$$X_f \subset (GF(2) \times GF(2))^{\mathbb{Z}}$$

consists of those pairs of elements $\mathbf{x}$, $\mathbf{y}$ in $GF(2)^{\mathbb{Z}}$ satisfying (2-5), paraphrased here as

$$P_f(\mathbf{x}) - (\mathbf{y} - \sigma\mathbf{y}) = \mathbf{0} \in GF(2)^{\mathbb{Z}}. \tag{2-6}$$

The left hand side of (2-6) constitutes a shift-equivariant polynomial map

$$(GF(2) \times GF(2))^{\mathbb{Z}} \ni (\mathbf{x}, \mathbf{y}) \mapsto Q(\mathbf{x}, \mathbf{y}) \in GF(2)^{\mathbb{Z}},$$

in the sense that

- there is some finite interval $I \subset \mathbb{Z}$ such that

  $$Q(\mathbf{x}, \mathbf{y})_0 = \text{polynomial } R(x_i, y_j) \text{ for } i, j \in I$$

  (that justifies the term "polynomial") and
- $Q(\sigma\mathbf{x}, \sigma\mathbf{y}) = \sigma Q(\mathbf{x}, \mathbf{y})$ (i.e. "shift-equivariant").

In other words, X_f consists precisely of those sequences of elements in $GF(2) \times GF(2)$ which do not contain, as subwords, the finitely many non-solutions to

$$R(x_i, y_j) = 0, \ i, j \in I.$$

This makes it clear that the shift is indeed of finite type. ■

In particular, Theorem 1.5 and Lemma 2.6 then proves

Corollary 2.8 *Let f be an RS Boolean function and set*

$$N_n = 2^{n+1} - 2wt(f_n) = 2^n + W_f(\mathbf{0}).$$

Then, we have

$$\exp\left(\sum_{n\geq 1} \frac{N_n}{n} s^n\right) = \frac{1}{\det(1 - sA)}$$

for some integer square matrix A.

Or again:

Corollary 2.9 *For any RS Boolean function f there are algebraic integers α_i, $1 \leq i \leq r$ such that*

(a) the multiset $\{\alpha_i\}$ is closed under Galois conjugation over $\mathbb{Q}$, and

(b) we have

$$wt(f_n) = 2^n - \frac{\alpha_1^n + \cdots + \alpha_r^n}{2}.$$

Proof Let α_i, $1 \leq i \leq r$ be the eigenvalues (with multiplicity) of the integer matrix A. We have

$$\frac{1}{\det(1 - sA)} = \exp\left(\sum_{n\geq 1} \frac{\alpha_1^n + \cdots + \alpha_r^n}{n} s^n\right),$$

so

$$2^{n+1} - 2wt(f_n) = N_n = \alpha_1^n + \cdots + \alpha_r^n$$

for N_n as in Corollary 2.8. This completes the proof. ■

Finally, as an immediate consequence of Corollary 2.9 we obtain

Corollary 2.10 *The weights $wt(f_n)$ of an RS Boolean function f satisfy a linear recurrence with integer coefficients.*

This provides a new proof for the existence of the linear recurrences, which can be computed using the results in [9, 10].

3 Trace representations and the Weil conjectures

We now give a parallel treatment for trace-context functions $f_n = f_{\mathcal{C},n} : GF(2^n) \to GF(2)$ as in Definition 1.1. One is again interested in the weights $wt(f_n)$, i.e. the cardinalities of the sets $f_n^{-1}(1) \subset GF(2^n)$.

The analogue of Lemma 2.1 in the present setting is precisely the Hilbert theorem 90 recalled in Remark 2.2:

Lemma 3.1 *Let n be a positive integer. An element $x \in GF(2^n)$ has vanishing trace if and only if $x = y - \sigma y$, where σ is the Frobenius automorphism $y \mapsto y^2$ on $GF(2^n)$.*

As in Section 2, we introduce the polynomials

$$P_{f,n} : GF(2^n) \to GF(2^n)$$

defined for monomials (1-1) by

$$P_{f,n}(x) = x \cdot x^{2^{a_1}} \cdots x^{2^{a_{d-1}}}$$

and extended additively from this in general. These are restrictions to $GF(2^n)$ of a single polynomial P_f defined on the entire algebraic closure $\overline{GF(2)}$. We now have $f_n = \mathrm{Tr}_n \circ P_{f,n}$ (as in the RS case), hence the following versions of Lemma 2.3 and Corollary 2.4.

Lemma 3.2 *For $x \in GF(2^n)$ we have $f_n(x) = 0$ if and only if*

$$P_{f,n}(x) = y - y^2 \tag{3-1}$$

for some $y \in GF(2^n)$.

Corollary 3.3 *The number of zeros of f_n (i.e. $2^n - wt(f_n)$) is half the number of solutions $(x, y) \in GF(2^n)^2$ to the equation (3-1).*

We will now repurpose the notation from Section 2: X_f will denote the affine plane algebraic curve

$$X_f = \{(x, y) \in \overline{GF(2)}^2 \mid P_f(x) = y - y^2\}. \tag{3-2}$$

With this notation, Corollary 3.3 says that we have

$$2^{n+1} - 2wt(f_n) = N_n(X_f). \tag{3-3}$$

We would now like to apply the point count in Theorem 1.7 to the curve X_f with $q = 2$. The only slight obstacle is that theorem applies to *projective* curves, whereas X_f is affine. Its closure X'_f in the projective plane $\mathbb{P}^2$ over the algebraic closure $\overline{GF(2)}$ is given by the *homogenization* of the defining equation

$$P_f(x) = y - y^2$$

in (3-2):

$$X'_f = \{[x : y : z] \in \mathbb{P}^2 \mid \overline{P}_f(x, z) = yz^{e-1} - y^2 z^{e-2}\} \tag{3-4}$$

where

- e is the largest degree of a monomial in P_f, and
- $\overline{P}_f$ is the homogeneous degree-e polynomial in x, z obtained by multiplying each monomial of $P_f(x)$ by the appropriate power of z.

Remark 3.4 e is of the form

$$1 + 2^{a_0} + \cdots + 2^{a_{d-1}}$$

for a tuple (1-1) and is thus odd and ≥ 3. ♦

Now, note that the original affine curve X_f consists precisely of those points in its projective completion (3-4) with $z = 0$. Since exactly one of the monomials in $\overline{P}_f(x, z)$ is a power of x, we have

$$[x : y : z] \in X'_f,\ z = 0 \Rightarrow x = 0 \Rightarrow [x : y : z] = [0, 1, 0] =: p_0.$$

In other words, the affine curve is missing exactly one point of its completion:

$$|X'_f(GF(2^n))| - |X_f(GF(2^n))| = 1,\ \forall n \geq 1.$$

In other words, the version of Theorem 1.7 applicable to X_f simply omits the '+1' summand in that statement:

Theorem 3.5 *Let f_n, $n \geq 1$ be a family of trace functions $GF(2^n) \to GF(2)$ attached to a finite set of tuples (1-1). Then, there are Galois-invariant multisets of algebraic integers*

- *α_i, $1 \leq i \leq 2g$ with $|\alpha_i| = \sqrt{2}$;*
- *β_j, $1 \leq j \leq \Delta$ with $|\beta_j| = 1$*

such that

$$wt(f_n) = 2^{n-1} + \frac{\sum_{i=1}^{2g} \alpha_i^n}{2} + \frac{\sum_{j=1}^{\Delta} \beta_j^n}{2}. \tag{3-5}$$

Proof Simply apply Theorem 1.7 to the projective curve X'_f, omit the '+1' term in (1-3) as explained above, and use (3-3) to identify $N_n(X_f)$ with $2^{n+1} - 2wt(f_n)$. The rest is simple arithmetic.■

Theorem 1.7 makes it clear that the size Δ of the set of β_j depends on "how singular" the projective curve in question is. For that reason, it will be of interest to understand the singularities of our curve X'_f defined in (3-4). Writing

$$Q(x, y, z) = Q_f(x, y, z) := \overline{P}_f(x, z) - yz^{e-1} + y^2 z^{e-2}$$

for the homogeneous degree-e polynomial whose vanishing defines X'_f. The singularities of the latter are the points where

$$\frac{\partial Q}{\partial x} = \frac{\partial Q}{\partial y} = \frac{\partial Q}{\partial z} = 0.$$

The partial derivative $\frac{\partial Q}{\partial y}$ is nothing but z^{e-1} (because we are in characteristic 2 and hence the derivative of $y \mapsto y^2$ vanishes), so the singular set of X'_f is either empty or precisely

$$\{p_0\} = \{[0:1:0]\} = X'_f \setminus X_f.$$

As for whether or not p_0 is indeed singular, we first observe that the x and y partial derivatives do indeed vanish, leaving the question of whether $\frac{\partial Q}{\partial z}$ does. Recall from Remark 3.4 that e is odd and hence all powers of z appearing in $\overline{P}_f(x, z)$ are even. It follows that the z-partial derivative of $\overline{P}_f(x, z)$ vanishes, so

$$\frac{\partial Q}{\partial z}(p_0) = y^2 z^{e-3}.$$

This is zero (and hence the point is singular) when $e > 3$ and non-zero when $e = 3$. We thus have two possibilities:

(a) $e = 3$, in which case X'_f is an elliptic curve;

(b) $e > 3$, in which case X'_f is a projective plane curve with a single singularity at $[0:1:0]$.

We now focus on case (b), seeking to determine the discrepancy between $X' = X'_f$ and its desingularization. First, we focus attention on the affine portion C of X' corresponding to $y \neq 0$. Making the variable change

$$u = \frac{x}{y}, \ v = \frac{z}{y},$$

we can describe C as the curve in the u, v plane defined by the equation

$$\overline{P}_f(u, v) + v^{e-1} + v^{e-2} = 0 \tag{3-6}$$

(where we have dropped minus signs, since we are working in characteristic two). We now proceed to resolve the singularity $p_0 = (0,0)$ (in u, v coordinates) of C by the procedure described in [17, Theorem V.3.9 and surrounding discussion], of successive blowup.

The initial blowup of the curve $C \in \mathbb{A}^2$ (the affine plane) defined by (3-6) centered at the singularity $(0,0)$ is achieved as described on [17, pp.29-30]:

- Introduce coordinates α, β for the projective line $\mathbb{P}^1$ over $\overline{GF(2)}$.
- Consider the subvariety V of $\mathbb{A}^2 \times \mathbb{P}^1$ cut out by (3-6) and the equation

$$u\beta = v\alpha.$$

  V is the union of the distinguished projective line $E := \{(0,0)\} \times \mathbb{P}^1$ and the blowup C_1 of the original curve $C = C_0$.

- As in [17, Example I.4.9.1], we now cover E with the open affine patches $\alpha \neq 0$ and $\beta \neq 0$ and determine the intersection of C_1 with each open patch in order to determine the preimage of the singularity $(0,0)$ through the rational map $C_1 \to C$.

In this last step, assume first that $\alpha \neq 0$. By rescaling we can thus assume $v = u\beta$. Making this substitution in (3-6) we obtain

$$u^e Q(\beta) = u^{e-1}\beta^{e-1} + u^{e-2}\beta^{e-2} \tag{3-7}$$

for some polynomial in β with free term 1, so that $Q = 1 + R$ with $R(0) = 0$.

When $u = 0$ we have $v = 0$ as well, and the equations describe E. In order to determine its intersection with C_1 assume $u \neq 0$ in (3-7) and divide through by u^{e-2} to obtain

$$u^2(1 + R(\beta)) = u\beta^{e-1} + \beta^{e-2} = \beta^{e-2}(1 + u\beta). \tag{3-8}$$

The only solution to this equation with $u = 0$ is the point $[\alpha : \beta] = [1 : 0]$ on $E \cong \mathbb{P}^1$.

A similarly simple calculation shows that $C_1 \cap E$ contains *no* points in the open patch $\beta \neq 0$. In conclusion, the partial desingularization $C_1 \to C_0$ of $(0,0)$ provides a single singular point, obtained as $(0,0)$ on curve defined by (3-8) in the u, β plane.

Let A be the localization of the ring

$$\overline{GF(2)}[u, \beta]/(\text{equation } (3-8))$$

at the ideal (u, β) and $\widehat{A}$ its completion with respect to its maximal ideal. In other words, $\widehat{A}$ is the formal power series ring

$$\overline{GF(2)}[[u, \beta]]$$

modulo the equation (3-8).

Since we are in characteristic 2 and $e - 2$ is odd, $1 + u\beta$ and $1 + R(\beta)$ are both (invertible) $(e-2)^{nd}$ powers in $\widehat{A}$. This means that we can make a change of variables

$$u \mapsto u, \ \ \beta \mapsto \gamma = g(\beta)$$

in $\widehat{A}$ so as to transform (3-8) into

$$u^2 = \gamma^{e-2}. \tag{3-9}$$

In the language of [17, §I.5], the $(0,0)$ singularity of (3-8) is *analytically isomorphic* to the $(0,0)$ singularity of (3-9). But the singularities of the form (3-9) are analyzed in [17, Example V.3.9.5]: they are resolved through a sequence of blowups

$$C_{\frac{e-1}{2}} \to \cdots \to C_2 \to C_1, \tag{3-10}$$

with each C_i having a single singular point.

This analysis will allow us to sharpen Theorem 3.5 in two ways. First, since we have just established that the desingularization $\widetilde{X'_f} \to X'_f$ has a *unique* point mapping to the singularity of X'_f, Theorem 1.7 says that in fact $\Delta = 0$, i.e. there are no βs in (3-5):

Corollary 3.6 *Let f_n, $n \geq 1$ be a family of trace functions $GF(2^n) \to GF(2)$ attached to a finite set of tuples (1-1). Then, there is a Galois-invariant multiset of algebraic integers*

$$\alpha_i,\ 1 \leq i \leq 2g \text{ with } |\alpha_i| = \sqrt{2}$$

such that

$$wt(f_n) = 2^{n-1} + \frac{\sum_{i=1}^{2g} \alpha_i^n}{2}. \tag{3-11}$$

Secondly, we can determine the genus g of the desingularization $\widetilde{X'_f} \to X'_f$ (i.e. the g appearing in (3-11)). This will require stepping through the desingularization procedure by successive blowup sketched above, using the numerical information provided by [17, Example V.3.9.2].

The latter says that the genus g of the smooth curve $\widetilde{X'_f}$ is obtained from the arithmetic genus $p_a(X'_f)$ by subtracting

$$\sum_p \frac{r_p(r_p-1)}{2}$$

for all singular points appearing during the successive blowups, where r_p is the *multiplicity* of the singular point p.

We now assemble the ingredients:

- The arithmetic genus $p_a(X'_f)$ is
$$\frac{(e-1)(e-2)}{2},$$
since e is the degree of the plane curve $X'_f \subset \mathbb{P}^2$ ([17, Exercise I.7.2]).
- The multiplicity of the singularity $(0,0)$ on a plane curve is the smallest degree appearing in an expansion of its defining equation. It is thus $e-2$ for the initial singularity (3-6) and 2 for each of the subsequent $\frac{e-3}{2}$ desingularization steps in (3-10).
- In conclusion, the genus g is
$$\frac{(e-1)(e-2)}{2} - \frac{(e-2)(e-3)}{2} - \frac{e-3}{2} = \frac{e-1}{2}.$$

In short:

Corollary 3.7 *The number $2g$ of summands in (3-11) is $e-1$, where e is the degree*

$$1 + 2^{a_0} + \cdots + 2^{a_{d-1}}$$

of P_f.

4 Quadratic functions

By 'quadratic' we mean functions $f_{\mathcal{C},n}$ (in either the trace or RS setup) associated to collections $\mathcal{C}$ of tuples (1-1) with $d = 2$. In that case we refer to $\mathcal{C}$ itself (or to its members) as being quadratic. The functions $f_{\mathcal{C},n}$ for quadratic $\mathcal{C}$ form the focus of the present section.

First, it is well known that quadratic Boolean functions are *plateaued*: the weight of a quadratic Boolean function f_n either vanishes or is of the form $2^{n-1} \pm 2^{\frac{n+v}{2}-1}$ for some integer v of the same parity as n. the same applies in trace context, since trace functions as in Definition 1.1 can always be regarded as quadratic Boolean functions after choosing an appropriate basis for $GF(2^n)$ (see [7, Remark 3.2]).

Now let $\mathcal{C}$ be a finite collection of quadratic tuples (1-1) and $f_{\mathcal{C}}$, $g_{\mathcal{C}}$ the RS and trace function families attached to $\mathcal{C}$ respectively. With this in place, [7, Theorem 5.1] implies that f_n and g_n are plateaued for the same parameter $v = v(n)$.

We can in fact say more [8, Theorem 2.1]:

Theorem 4.1 *Let $\mathcal{C}$ be a quadratic family of tuples (1-1) and $f_{\mathcal{C},n}$, $g_{\mathcal{C},n}$ the RS and respectively trace functions attached to it. Then, we either have $wt(f_{\mathcal{C},n}) = 0 = wt(g_{\mathcal{C},n})$ or*

$$wt(f_{\mathcal{C},n}) = 2^{n-1} \pm 2^{\frac{n+v}{2}-1}$$
$$wt(g_{\mathcal{C},n}) = 2^{n-1} \pm 2^{\frac{n+v}{2}-1}$$

for the same $v = v(n)$ (but perhaps not the same sign).

We have

$$v(n) = \deg \gcd(x^n - 1, A_n(x))$$

where

$$A_n(x) = \sum_{(0,t)\in\mathcal{C}} (x^t + x^{n-t})$$

and the greatest common divisor is taken in the polynomial ring $GF(2)[x]$. It follows that $v(n)$ is periodic in n, and reaches its maximal value once per period: precisely when n is divisible by

$$N = N_{\mathcal{C}} = \min\{n \text{ such that } A(x) \mid x^n - 1\} \tag{4-1}$$

where

$$A(x) = \sum_{(0,t)\in\mathcal{C}} (x^t + x^{-t})$$

and divisibility takes place in the Laurent polynomial ring $GF(2)[x^{\pm 1}]$. For all of this we refer to [8, Theorem 5.2].

All of this additional information available in the quadratic case allows us to recast Corollaries 2.9 and 3.6 as follows.

Theorem 4.2 *Let $\mathcal{C}$ be a finite collection of quadratic tuples (1-1) and $f_{\mathcal{C},n}$, $g_{\mathcal{C},n}$ the RS and trace functions associated to $\mathcal{C}$ respectively. Then, there are Galois-invariant multisets*

$$|\alpha_i| = \sqrt{2} = |\gamma_j|,\ 1 \leq i, j \leq max_n 2^{\frac{v(n)}{2}}$$

such that

$$wt(f_n) = 2^{n-1} - \frac{\sum \alpha_i^n}{2}$$

and

$$wt(g_n) = 2^{n-1} + \frac{\sum \gamma_i^n}{2}.$$

Furthermore, the degree of the group generated by the roots of unity $\frac{\alpha_i}{\sqrt{2}}$ *(resp.* $\frac{\gamma_j}{\sqrt{2}}$*) is either the period* $N = N_{\mathcal{C}}$ *of* $v(n)$ *or* $2N$*.*

Proof To fix ideas, we focus on the trace functions $g_n = g_{\mathcal{C},n}$. The RS half of the statement will follow from this and [7, Theorem 5.1] (which says that $f_{\mathcal{C},n}$ and $g_{\mathcal{C},n}$ have the same nonlinearity) or Theorem 4.1.

On the one hand, we know that

$$wt(g_n) = 2^{n-1} \pm 2^{\frac{n+v}{2}-1} \text{ or } 0. \tag{4-2}$$

On the other hand, by Corollary 3.6 we have

$$wt(g_n) = 2^{n-1} + \frac{\sum \gamma_j^n}{2} \tag{4-3}$$

for a Galois-invariant multiset of algebraic integers γ_j of absolute value $\sqrt{2}$. We write

$$\chi_j = \frac{\gamma_j}{\sqrt{2}}$$

for the phases of γ_j.

The fact that the number of γ_j is $\max 2^{\frac{v}{2}}$ follows by comparing (4-2) and (4-3): the former says that the maximal absolute value of

$$\frac{wt(g_n) - 2^{n-1}}{2^{\frac{n}{2}-1}} \tag{4-4}$$

is $\max 2^{\frac{v}{2}}$ while the latter shows it is the size of the multiset $(\gamma_j)_j$.

We now observe that

- By (4-3), the (4-4) is maximized precisely when all γ_j^n are positive, i.e. n is divisible by

$$\text{ord}(\gamma_j, j) := |\text{group generated by the phases of } \gamma_j|\,.$$

- On the other hand, (4-2) shows that (4-4) is maximized in absolute value if and only if $wt(g_n) \neq 0$ and $v(n)$ is maximal, i.e. n is divisible by the period $N = N_{\mathcal{C}}$ defined in (4-1).
- (4-4) can be maximized in absolute value but negative only there is some n so that $\chi_j^n = -1$ for all j.

We now consider several possibilities.

(a) $wt(g_N) \neq 2^{N-1}$ **and (4-4) is positive for** $n = N$**.** In this case the remarks above show that (4-4) achieves its maximal value at $n = N$ and hence $\text{ord}(\gamma_j, j) = N$.

(b) $wt(g_N) \neq 2^{N-1}$ **and (4-4) is negative for** $n = N$**.** This means that (4-4) is maximal in absolute value but negative at $n = N$ and hence $\chi_j^N = -1$ for all j. But this then implies that $\text{ord}(\gamma_j, j) = 2N$.

(c) $wt(g_N) = 2^{N-1}$**.** We know from [8, Corollary 5.19] that g_{2N} is *not* balanced, i.e. $wt(g_{2N}) \neq 2^{2N-1}$. We can now reiterate the arguments in cases (a) and (b) with $2N$ in place of N to conclude that

$$\text{ord}(\gamma_j, j) = 2N \text{ or } 4N.$$

This finishes the proof of the theorem. ■

The following result is an offshoot of the proof of Theorem 4.2.

Theorem 4.3 *Under the hypotheses of Theorem 4.2 the weights $wt(f_n)$ and $wt(g_n)$ satisfy linear recurrences of orders $\leq 2N+1$, where N is the period of the sequence $v(n)$ of plateau parameters.*

Proof In the language of Theorem 4.2, consider the various possible values for $\mathrm{ord}(\gamma_j, j)$:

If it is $\leq 2N$ then we are done. Indeed, by (4-3) the weight $w(g_n)$ is a linear combination of n^{th} powers of algebraic integers satisfying the polynomial equation

$$(x-2)(x^{2N}-2^N)=0.$$

On the other hand, it follows from the proof of Theorem 4.2 that the case

$$\mathrm{ord}(\gamma_j, j) = 4N$$

occurs only when $\gamma_j^{2N} = -2^N$ for all j. This means we have recursion polynomial

$$(x-2)(x^{2N}+2^N)=0.$$

instead.

Either way, the minimal recursion polynomial will have degree $\leq 2N$. ■

It turns out that, for the quadratic MRS function $(0,t)_n = h_{t,n}$, say, in the RS context, the algebraic integers $\alpha_1, \ldots, \alpha_{2^t}$ from Corollary 2.9 and Theorem 3.5 can be taken to be the roots of the characteristic polynomial of the square matrix $R(t)$ associated with $h_{t,n}$ in [8, Section 3]. The next theorem proves this. Note that this count of the numbers α_i agrees with the count given in (3-11) and Corollary 3.7. The matrix $R(t)$ has 2^t rows and is given explicitly in [8, Theorem 3.1].

To describe the matrix we need the cyclic permutation μ which acts on vectors $(b_1, b_2, \ldots, b_k)$ of any length k by placing the last entry to the front, e.g.

$$\mu^2((0,1,0,1,0,0)) = (0,0,0,1,0,1).$$

We use 0_j to stand for a string of j consecutive entries equal to 0 and similarly for 1_j. Now $R(t)$ is the square matrix whose rows are the 2^{t-1} pairs

$$\mu^i((1, 0_{2^{t-1}-1}, 1, 0_{2^{t-1}-1})),\ \mu^i((1, 0_{2^{t-1}-1}, -1, 0_{2^{t-1}-1}))$$

for $i = 0, 1, \ldots, 2^{t-1}-1$ taken in order.

The minimal polynomial for $R(t)$ has degree $2t$ and is [8, Theorem 3.4]

$$x^{2t} - 2^t. \tag{4-5}$$

We let $\delta_1, \ldots, \delta_{2t}$ denote the roots of (4-5). These roots are obviously distinct. The characteristic polynomial, say $c_t(x)$, for $R(t)$ has degree 2^t and has the same roots δ_i, but in general some of these roots will occur multiple times. Thus the multiset of roots of $c_t(x)$, counted with multiplicities, will have size 2^t but only $2t$ distinct elements.

Theorem 4.4 *The recursion for the weights of* $(0,t)_n = h_{t,n}$, *extended backwards from* $n = 2t+1$ *to* $n = 1$, *generates a sequence* $w_1, w_2, \ldots$ *(with* $w_i = wt(h_{t,i})$ *for* $i \geq 2t+1$*) such that*

$$w_n = 2^{n-1} - \frac{1}{2}(\delta_1^n + \ldots + \delta_{2^t}^n),\ n = 1, 2, \ldots. \tag{4-6}$$

Here $\delta_1, \ldots, \delta_{2^t}$ *is the list of the* 2^t *roots of the characteristic polynomial* $c_t(x)$ *for the matrix* $R(t)$, *with the distinct roots* $\delta_1, \ldots, \delta_{2t}$ *of the minimal polynomial* $m_t(x) = x^{2t} - 2^t$ *for* $R(t)$ *listed first. The remaining roots are various duplicates of the first* $2t$ *roots.*

For monomial functions $(0,t)_n$ we have the following consequence of [19, Theorem 8]. Recall that the *Möbius function* [16, §16.3] is defined by

$$\mu(n) = \begin{cases} 1 & \text{if } n = 1 \\ (-1)^k & \text{if } n \text{ is a product of } k \text{ distinct prime factors} \\ 0 & \text{otherwise} \end{cases}$$

Theorem 4.5 *Let* $t = 2^\nu m$ *for an odd number* m *and* $\nu \geq 0$. *The weight* w_n *of* $(0,t)_n$ *is expressible as*

$$w_n = 2^{n-1} - \frac{1}{2}(\delta_1^n + \ldots + \delta_{2^t}^n),\ n = 1, 2, \ldots$$

where the multiset (δ_i) *is the union* $\sqrt{2}\Delta$ *of the multisets* $\sqrt{2}\Delta_d$ *indexed by divisors* $d|m$, *where* Δ_d *is the collection of* $2^{\nu+1}d^{th}$ *roots of unity, each with multiplicity.*

$$\frac{\sum_{d'|d} \mu\left(\frac{d}{d'}\right) 2^{2^\nu d'}}{2^{\nu+1} d} \tag{4-7}$$

where μ *is the Möbius function.*

Implicit in the statement of Theorem 4.5 is the remark that the numerator of (4-7) is divisible by its denominator. Since every summand $\pm 2^{2^\nu d'}$ of the numerator is a multiple of $2^{\nu+1}$, so is the numerator as a whole. On the other hand, divisibility by d follows from

Lemma 4.6 *Let* $q, d > 1$ *be coprime positive integers. Then,*

$$D(d,q) := \sum_{d'|d} \mu\left(\frac{d}{d'}\right) q^{d'}$$

is divisible by d.

Proof If q is a prime power then $\frac{D(d,q)}{d}$ is known to be the number of monic irreducible degree-d polynomials over the field $GF(q)$ with q elements [18, §4.13, Corollary 2]. In general, since q and d are assumed coprime Dirichlet's theorem on primes in arithmetic progressions ([26, §VI.4, Theorem 2]) ensures that there is some prime congruent to q modulo d, reducing the problem to the prime-q case. ■

We will also need the following remark.

Lemma 4.7 *Let* $\{\gamma_i\}$ *and* $\{\delta_j\}$ *be two finite sets of complex numbers and* $s_i, t_j \in \mathbb{C}$. *If*

$$\sum_i s_i \gamma_i^n = \sum_j t_j \delta_j^n \tag{4-8}$$

for all non-negative integers n *then the sets* $\{\gamma_i\}$ *and* $\{\delta_j\}$ *and, having identified their respective elements, the corresponding coefficients* s_i *and* t_j *also coincide.*

Proof Suppose the conclusion does *not* hold. Rewriting (4-8) as

$$\sum_i s_i \gamma_i^n = \sum_j t_j \delta_j^n$$

and aggregating the terms where some γ equals some δ, the failure of the conclusion means that we obtain equations

$$\sum_k u_k \eta_k^n = 0, \ \forall n$$

for some non-empty set of (distinct) η_k's (and non-zero u_k). But this means that the vector with components u_k is annihilated by the Vandermonde matrix with entries

$$u_{kl} := \eta_k^{\ell-1}.$$

This contradicts the fact that said matrix has non-zero determinant $\prod_{k>k'}(\eta_k - \eta_{k'})$ and is thus invertible. ■

Proof of Theorem 4.5 According to Lemma 4.7, it will be enough to show that

$$w_n = 2^{n-1} - 2^{\frac{n}{2}-1} \sum_{\delta \in \Delta} \delta^n.$$

Equivalently, by [19, Theorem 8] this amounts to

$$\sum_{\delta \in \Delta} \delta^n = \begin{cases} 2^{gcd(n,t)} & \text{if } \frac{n}{gcd(n,t)} \text{ is even} \\ 0 & \text{otherwise.} \end{cases} \tag{4-9}$$

The second branch is easily dispatched: $\frac{n}{gcd(n,t)}$ being odd is equivalent to n *not* being divisible by $2^{\nu+1}$. Since each Δ_d consists of the $2^{\nu+1}d^{th}$ roots of unity all with the same multiplicity, we have

$$\sum_{\delta \in \Delta_d} \delta^n = 0, \ \forall d | m.$$

It thus remains to treat the case when $2^{\nu+1}$ divides n, when the target equality (4-9) becomes

$$\sum_{\delta \in \Delta} \delta^n = 2^{gcd(n,t)} = 2^{2^\nu gcd(n,m)}. \tag{4-10}$$

Set $D = gcd(n, m)$ for brevity. All

$$\sum_{\delta \in \Delta_d} \delta^n, \ d \not| D$$

vanish, so we need only consider divisors $d|D$. Keeping this in mind (4-10) reads

$$\sum_{d'|d|D} \mu\left(\frac{d}{d'}\right) 2^{2^\nu d'} = 2^{2^\nu D}.$$

This, however, is nothing but an instance of the Möbius inversion formula [16, §16.4]. ■

By (4-9), proving Theorem 4.4 amounts to showing that for every n, we have

$$\text{tr } R(t)^n = \begin{cases} 2^{\frac{n}{2}+gcd(n,t)} & \text{if } \frac{n}{gcd(n,t)} \text{ is even} \\ 0 & \text{otherwise.} \end{cases} \tag{4-11}$$

Since by (4-5) the eigenvalues of $R(t)$ are $2t^{th}$ roots of unity rescaled by $\sqrt{2}$, it is enough to prove (4-11) for $1 \leq n \leq 2t-1$. It will thus be useful to describe $R(t)^n$ explicitly. To that end, we follow [8, §3] in denoting by $M(n)$ the $2^n \times 2^n$ matrix

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^{\otimes n}.$$

For $1 \leq n \leq t$ we also write $M(n,t)$ for the $2^n \times 2^t$ matrix obtained by inserting $2^{t-n}-1$ zero columns after each original column of $M(n)$. For a matrix M we write $\mu(M)$ (or μM) for the matrix obtained by rotating the rows of M rightward (this extends the above definition of the cyclic permutation μ on the individual rows). With all of this in place, the following is a simple computation achievable inductively by partitioning $R(t)$ into four $2^{t-1} \times 2^{t-1}$ block matrices.

Lemma 4.8 *Let $1 \leq n \leq 2t-1$. The power $R(t)^n$ can then be described as follows.*

(a) If $1 \leq n \leq t$ then

$$R(t)^n = \begin{pmatrix} M(n,t) \\ \mu M(n,t) \\ \vdots \\ \mu^{2^{t-n}-1} M(n,t) \end{pmatrix}.$$

(b) On the other hand, if $t \leq n \leq 2t-1$ then

$$R(t)^n = 2^{n-t} \left(R(t)^{2t-n}\right)^T,$$

where the T superscript denotes transposition. ■

In particular, part (b) of Lemma 4.8 proves (4-11) for $t \leq n \leq 2t-1$ provided it is known for $1 \leq n \leq t$. Even more robustly, it recovers (4-11) for a specific $t \leq n \leq 2t-1$ provided we know the analogue for the reflection $2t-n$ of n across t. In conclusion, it suffices to focus on the range $1 \leq n \leq t$. In turn, in those cases the trace of interest is computable as follows, numbering the rows and columns of all matrices starting at 0:

Lemma 4.9 *For $1 \leq n \leq t$ the trace $\text{tr } R(t)^n$ is the sum of the following elements of $M(n)$:*

- *the lower right hand corner $M(n)_{2^n-1,2^n-1}$;*
- *for each $0 \leq k \leq 2^n-2$ the entry with index $2^{t-n}k$ modulo 2^n-1 in the k^{th} column.* ■

Note that $2^{dn}-1$ is divisible by 2^n-1 for all $d \geq 0$, so in Lemma 4.9 is is enough to replace 2^{t-n} with the residue $t(\text{mod } n)$. We can now rephrase Lemma 4.9 as follows.

Lemma 4.10 *Let $1 \leq n < t$ and set $a = t(\text{mod } n)$ and $b = n-a$. Then, the trace $\text{tr } R(t)^n$ is the sum of the entries*

$$(q + r2^a,\ q2^b + r) \tag{4-12}$$

where $0 \leq q \leq 2^a-1$ and $0 \leq r \leq 2^b-1$. ■

We make note of the following "central symmetry" property of the Hadamard matrix $M(n)$.

Lemma 4.11 *Let $a+b=n$ be positive integers and denote by (k,ℓ) the coordinates (4-12) of an entry in $M(n)$ for some $0 \leq q \leq 2^a-1$ and $0 \leq r \leq 2^b-1$. Let also*

$$k' = 2^n - 1 - k = q' + r'2^a$$
$$\ell' = 2^n - 1 - \ell = q'2^b + r'$$

be the coordinates of the reflection of (k,ℓ) across the center of the matrix $M(n)$, where

$$q' + q = 2^a - 1, \quad r' + r = 2^b - 1.$$

Then,

$$M(n)_{k',\ell'} = \begin{cases} M(n)_{k,\ell} & \text{if } n \text{ is even} \\ -M(n)_{k,\ell} & \text{if } n \text{ is odd.} \end{cases}$$

Proof We can prove this by induction on n, using the decomposition

$$M(n+1) = \begin{pmatrix} M(n) & M(n) \\ M(n) & -M(n) \end{pmatrix}$$

and treating the separate possibilities for the placement of (k,ℓ) in one of the four quadrants. If, say, $M(n+1)_{k,\ell}$ is in the upper left hand $M(n)$ corner and hence

$$M(n+1)_{k,\ell} = M(n)_{k,\ell}$$

then $M(n+1)_{k',\ell'}$ is in the lower right hand $-M(n)$ quadrant and hence is *minus* the reflection of $M(n)_{k,\ell}$ across the center of $M(n)$. The inductive hypothesis implies the conclusion.

The argument is analogous in the other cases, and we leave it to the reader; in fact, there is only *one* other case: (k,ℓ) and (k',ℓ') play symmetric roles, so it is enough to assume (k,ℓ) is either in the top left or the top right quadrant. ■

We can now tackle the following particular case of (4-11).

Corollary 4.12 $\text{tr } R(t)^n = 0$ *when n is odd and hence (4-11) holds in that case.*

Proof Indeed, Lemmas 4.10 and 4.11 show that the trace is a sum of pairs ± 1 of entries of $M(n)$, each pair summing to zero. ■

Corollary 4.13 *The matrix $R(t)$ is conjugate to $-R(t)$.*

Proof Since the two matrices are unitary and hence diagonalizable over the complex numbers it is enough to argue that they have the same characteristic polynomial. The coefficients of the latter are algorithmically computable from the traces of the powers of the matrix, so it is enough to show that we have

$$\text{tr } R(t)^n = \text{tr } (-R(t))^n = (-1)^n \text{tr } R(t)^n, \ \forall n.$$

The two sides are obviously equal for even n, so the conclusion follows from Corollary 4.12, which shows that everything in sight vanishes for odd n. ■

Recall the polynomials Θ_d discussed in §1.4.

Lemma 4.14 *The characteristic polynomial of $R(t)$ is a product of factors Θ_d for divisors $d|t$.*

Proof We already know that $R(t)$ is annihilated by $x^{2t} - 2^t$, so its characteristic polynomial will be a product of irreducible factors of the latter. By Proposition 1.8, the conclusion follows from the fact that any two irreducible factors P, Q of the characteristic polynomial related by $Q(x) = P(-x)$ have equal exponents because $R(t)$ is conjugate to $-R(t)$ (Corollary 4.13). ■

The roots of Θ_d are simply those of $\Phi_d(x^2)$ scaled by $\sqrt{2}$. In turn, since d divides t, the sum of n^{th} powers of the roots of $\Phi_d(x^2)$ equals the sum of $gcd(n, 2t)^{th}$ powers. In conclusion:

Lemma 4.15 *It suffices to prove (4-11) for $n|2t$.* ■

There are thus two cases: n divides t or it doesn't, but $\frac{n}{2}$ does. The easy half is

Proposition 4.16 *If $n|t$ then* tr $R(t)^n = 0$ *and hence (4-11) holds.*

Proof Indeed, in that case n divides $t - n$ and hence

$$2^n - 1 \mid 2^{t-n} - 1.$$

Lemma 4.9 then says that tr $R(t)^n$ is precisely the trace of $M(n)$, which is zero, being the n^{th} power of the trace of $\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$. ■

As for the case $n \not| \ t$, we then have $gcd(n, t) = \frac{n}{2}$ (since at any rate we are assuming that n divides $2t$) and hence the desired conclusion (4-11) reads

$$\text{tr } R(t)^n = 2^n.$$

Lemma 4.9 then equates this to proving

Lemma 4.17 *Let $0 \leq n \leq t$ be a divisor of $2t$ but not t. Then, for $0 \leq k \leq 2^n - 2$, the entry with index $2^{t-n}k$ modulo $2^n - 1$ in the k^{th} column of $M(n)$ is 1.*

Proof The hypothesis ensures that $t - n$ is of the form $(2s + 1)\frac{n}{2}$ for some $s \geq 0$, and hence

$$2^{t-n}k = 2^{sn}2^{\frac{n}{2}}k = 2^{\frac{n}{2}}k \quad \text{modulo} \quad 2^n - 1$$

because $2^n - 1$ divides $2^{sn} - 1$. In short, it will be enough to assume that $t = \frac{3n}{2}$ thus substituting $2^{\frac{n}{2}}$ for 2^{t-n} in the statement.

We can now partition $M(n)$ into blocks M_{ij} of size $2^{\frac{n}{2}} \times 2^{\frac{n}{2}}$ for $0 \leq i, j \leq 2^{\frac{n}{2}} - 1$, each a copy of either $M(\frac{n}{2})$ or $-M(\frac{n}{2})$. The entries of interest in the first $2^{\frac{n}{2}}$ columns are

- the 0^{th} entry in the 0^{th} row of M_{00};
- the 1^{st} entry in the 0^{th} row of M_{10};
- $\cdots$;
- entry $(M_{2^{\frac{n}{2}}-1,0})_{0,2^{\frac{n}{2}}-1}$.

The pattern recurs: including the bottom right corner of $M(n)$, the entries we are after are precisely those of the form $(M_{ij})_{ji}$. That these are all 1 follows from the recursive construction of $M(n)$ giving

$$M(n) = M\left(\frac{n}{2}\right) \otimes M\left(\frac{n}{2}\right)$$

together with the fact that M is symmetric. ■

Proof of Theorem 4.4 As discussed above, the result amounts to (4-11). In turn, the latter is taken care of by Lemma 4.15 and proposition 4.16 and Lemma 4.17. ■

Theorem 4.4 states that the formula which gives the weights w_n for the MRS function $(0,t)_n$ in terms of powers of the roots of the characteristic polynomial has simple coefficients which are all $\pm\frac{1}{2}$. We say that the recursion for the weights of $(0,t)_n$ has *easy coefficients*. We believe this remains true for *any* quadratic RS function and so state the following conjecture:

Conjecture 4.18 (Easy Coefficients Conjecture) *The recursion for the weights of any rotation symmetric function has easy coefficients, attached to a multiset of algebraic integers. At least in the quadratic case, these algebraic integers are the roots of the characteristic polynomial of a matrix computable by the method of [10].*

The first sentence of the Easy Coefficients Conjecture is proved by Corollary 2.9. The second sentence for MRS quadratic functions is proved by Theorem 4.4. Since there is no nice formula like (4-5) for the minimal polynomial for the square matrix R (generalization of $R(t)$ for $(0,t)_n$) corresponding to a general quadratic RS function, the method of proof of Theorem 4.4 does not apply. However, we are confident that the result is true for general quadratic RS functions. In fact, many computations suggest that the second sentence of the Easy Coefficients Conjecture is true for *all* RS functions, of any degree.

Given a recursion relation of order r for the weights w_n of any RS Boolean function in n variables, the standard way to compute the weights is to compute the needed initial r weights and then use the recursion to find further desired weights. It is well known that the runtime to find w_n is $O(n2^n)$ (the extra n comes from the operations needed to compute each entry in the truth table), so finding the initial conditions in this way is an exponential computation. Given a function for which the Easy Coefficients Conjecture is true, a much quicker way to find the initial weights is to compute the roots of the characteristic polynomial and then use the analog of (4-6). The problem of computing all of the roots of a polynomial with integer coefficients has been studied for a long time. The runtime for doing that is known to be $O(n(log^k\ n))$ for some small integer k, so an exponential computation in n has been replaced by one that is nearly *linear* in n.

References

[1] Carlisle M. Adams. Constructing symmetric ciphers using the CAST design procedure. volume 12, pages 283–316. 1997. Selected areas in cryptography (Ottawa, ON, 1995).

[2] Yves Aubry and Marc Perret. A Weil theorem for singular curves. In *Arithmetic, geometry and coding theory (Luminy, 1993)*, pages 1–7. de Gruyter, Berlin, 1996.

[3] Maxwell L. Bileschi, Thomas W. Cusick, and Daniel Padgett. Weights of Boolean cubic monomial rotation symmetric functions. *Cryptogr. Commun.*, 4(2):105–130, 2012.

[4] Céline Blondeau and Kaisa Nyberg. Perfect nonlinear functions and cryptography. *Finite Fields Appl.*, 32:120–147, 2015.

[5] R. Bowen and O. E. Lanford, III. Zeta functions of restrictions of the shift transformation. In *Global Analysis (Proc. Sympos. Pure Math., Vol. XIV, Berkeley, Calif., 1968)*, pages 43–49. Amer. Math. Soc., Providence, R.I., 1970.

[6] Alyssa Brown and Thomas W. Cusick. Recursive weights for some Boolean functions. *J. Math. Cryptol.*, 6(2):105–135, 2012.

[7] Claude Carlet, Guangpu Gao, and Wenfen Liu. A secondary construction and a transformation on rotation symmetric functions, and their action on bent and semi-bent functions. *J. Combin. Theory Ser. A*, 127:161–175, 2014.

[8] Alexandru Chirvasitu and Thomas W. Cusick. Affine equivalence for quadratic rotation symmetric boolean functions, 2019. arXiv:1908.08448.

[9] Thomas W. Cusick. Weight recursions for any rotation symmetric boolean functions, 2017. arXiv:1701.06648.

[10] Thomas W. Cusick. Weight recursions for any rotation symmetric Boolean functions. *IEEE Trans. Inform. Theory*, 64(4, part 2):2962–2968, 2018.

[11] Thomas W. Cusick and Pantelimon Stănică. Fast evaluation, weights and nonlinearity of rotation-symmetric functions. *Discrete Math.*, 258(1-3):289–301, 2002.

[12] Thomas W. Cusick and Pantelimon Stănică. *Cryptographic Boolean functions and applications*. Elsevier/Academic Press, London, second edition, 2017.

[13] Pierre Deligne. La conjecture de Weil. I. *Inst. Hautes Études Sci. Publ. Math.*, (43):273–307, 1974.

[14] Bernard Dwork. On the rationality of the zeta function of an algebraic variety. *Amer. J. Math.*, 82:631–648, 1960.

[15] Réjane Forré. The strict avalanche criterion: spectral properties of Boolean functions and an extended definition. In *Advances in cryptology—CRYPTO '88 (Santa Barbara, CA, 1988)*, volume 403 of *Lecture Notes in Comput. Sci.*, pages 450–468. Springer, Berlin, 1990.

[16] G. H. Hardy and E. M. Wright. *An introduction to the theory of numbers*. Oxford University Press, Oxford, sixth edition, 2008. Revised by D. R. Heath-Brown and J. H. Silverman, With a foreword by Andrew Wiles.

[17] Robin Hartshorne. *Algebraic geometry*. Springer-Verlag, New York-Heidelberg, 1977. Graduate Texts in Mathematics, No. 52.

[18] Nathan Jacobson. *Basic algebra. I.* W. H. Freeman and Company, New York, second edition, 1985.

[19] Hyeonjin Kim, Sung-Mo Park, and Sang Geun Hahn. On the weight and nonlinearity of homogeneous rotation symmetric Boolean functions of degree 2. *Discrete Appl. Math.*, 157(2):428–432, 2009.

[20] Serge Lang. *Algebraic number theory*, volume 110 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1994.

[21] Douglas Lind and Brian Marcus. *An introduction to symbolic dynamics and coding*. Cambridge University Press, Cambridge, 1995.

[22] Kaisa Nyberg. Perfect nonlinear S-boxes. In *Advances in cryptology—EUROCRYPT '91 (Brighton, 1991)*, volume 547 of *Lecture Notes in Comput. Sci.*, pages 378–386. Springer, Berlin, 1991.

[23] John D. Olsen, Robert A. Scholtz, and Lloyd R. Welch. Bent-function sequences. *IEEE Trans. Inform. Theory*, 28(6):858–864, 1982.

[24] Josef Pieprzyk and Cheng Xin Qu. Fast hashing and rotation-symmetric functions. *J.UCS*, 5(1):20–31, 1999.

[25] Jennifer Seberry, Xian Mo Zhang, and Yuliang Zheng. Nonlinearity and propagation characteristics of balanced Boolean functions. *Inform. and Comput.*, 119(1):1–13, 1995.

[26] J.-P. Serre. *A course in arithmetic*. Springer-Verlag, New York-Heidelberg, 1973. Translated from the French, Graduate Texts in Mathematics, No. 7.

[27] André Weil. *Sur les courbes algébriques et les variétés qui s'en déduisent*. Actualités Sci. Ind., no. 1041 = Publ. Inst. Math. Univ. Strasbourg **7** (1945). Hermann et Cie., Paris, 1948.

[28] André Weil. Numbers of solutions of equations in finite fields. *Bull. Amer. Math. Soc.*, 55:497–508, 1949.

Department of Mathematics, University at Buffalo, Buffalo, NY 14260-2900, USA
E-mail address: achirvas@buffalo.edu

Department of Mathematics, University at Buffalo, Buffalo, NY 14260-2900, USA
E-mail address: cusick@buffalo.edu