APN Functions 3: Linearity

Gary McGuire

Claude Shannon Institute
www.shannoninstitute.ie
and
School of Mathematical Sciences
University College Dublin
Ireland

SORIA 2010

Talk 3

1. Linear Attack
2. Nonlinearity
3. Fourier Transform
4. Kloosterman Sums, Algebraic Curves

Recall

Many modern ciphers are (roughly speaking) a series of ROUNDS, where each round consists of an S-box and a P-box, and a subkey input.

$$x \longrightarrow \underbrace{S(x) \longrightarrow P(S(x)) \longrightarrow}_{one \quad round} S(P(S(x))) \longrightarrow \cdots$$

The S-box has to satisfy certain criteria to be secure against certain attacks. Some are

1. The PN or APN property provides resistance of the S-box to differential attack.
2. The permutation property (i.e. S being invertible) makes it easier to invert (to decrypt).
3. High algebraic degree (resistance to algebraic attack).
4. High nonlinearity provides resistance of the S-box to linear attack (today).

Linear Attack

Linear Cryptanalysis tries to approximate an S-Box $f : (\mathbb{F}_2)^n \longrightarrow (\mathbb{F}_2)^n$ with a linear (or affine) function.
Search for a, b, δ such that

$$\langle a, x \rangle = \langle b, f(x) \rangle + \delta$$

for "many" x's. ($a, b \in (\mathbb{F}_2)^n$, $\delta = 0$ or 1)
(Matsui, 1993. Known Plaintext attack. DES in 2^{43}.)
So if the following sum

$$\sum_{x \in (\mathbb{F}_2)^n} (-1)^{\langle b, f(x) \rangle + \langle a, x \rangle}$$

for <u>some</u> $a, b \in (\mathbb{F}_2)^n$ is "large" in absolute value, then our S-box is in big trouble.

Fourier Transform

If $(\mathbb{F}_2)^n = L = \mathbb{F}_{2^n}$, then we use $\langle x, y\rangle = \mathrm{tr}(xy)$ as the inner product, so we get

$$\sum_{x\in L}(-1)^{\mathrm{tr}(bf(x)+ax)}$$

Here tr is the trace from L to $\mathbb{F}_2$.

For the function

$$F_b(x) := (-1)^{\mathrm{tr}(bf(x))}$$

define its Fourier transform $\widehat{F}_b$ by

$$\widehat{F}_b(a) := \sum_{x\in L} F_b(x)(-1)^{\mathrm{tr}(ax)} = \sum_{x\in L}(-1)^{\mathrm{tr}(bf(x)+ax)}$$

So if $|\widehat{F}_b(a)|$ is “large” for some a, b then we are in trouble.
We want f such that all these Fourier coefficients are small.
Define $\mathbb{L}(f) = \max_{a,b\neq 0} |\widehat{F}_b(a)|$ (the linearity of f)
So we want functions with small linearity. (i.e. highly nonlinear)

Parseval's Equation says ($q = 2^n$)

$$\sum_a |\widehat{F}_b(a)|^2 = q^2$$

So the average of the squares is q, and it follows that $|\widehat{F}_b(a)|^2 \geq q$ for some a, so

Theorem

If $f : \mathbb{F}_q \longrightarrow \mathbb{F}_q$ then $\sqrt{q} \leq \mathbb{L}(f) \leq q$.

(We will not discuss funtions $\mathbb{F}_{2^n} \longrightarrow \mathbb{F}_2$ here.)
Sidelnikov, Chabaud-Vaudenay improved this for $f : \mathbb{F}_{2^n} \longrightarrow \mathbb{F}_{2^n}$

$$\mathbb{L}(f) \geq 2^{\frac{n+1}{2}} = \sqrt{2q}$$

Exercise: Show that a linear function f has $\mathbb{L}(f) = q$. (easy)
Exercise: Find the linearity of $f(x) = x^3$ (BCH code) (hard)

Inverse Function

What about $f(x) = x^{-1}$?
The Fourier coefficients become

$$\widehat{F}_b(a) = \sum_{x \in L} (-1)^{\mathrm{tr}(bx^{-1}+ax)}$$

With a change of variable (replace x by bx) this becomes

$$K(c) = \sum_{x \in L} (-1)^{\mathrm{tr}(x^{-1}+cx)}$$

This is known as a Kloosterman sum.
There is a lot of literature about Kloosterman sums.
In particular, from the Weil bound it is known that

$$-2^{n/2+1} \leq K(c) \leq 2^{n/2+1},$$

and every value which is congruent to 0 modulo 4 in that range is taken.
It follows that $\mathbb{L}(x^{-1}) \leq 2^{n/2+1} = 2\sqrt{q}$, and equality holds if n is even.

Note: equality can only hold in the Sidelnikov bound if n is odd.

Theorem

(Sidelnikov, Chabaud-Vaudenay) For functions $(\mathbb{F}_2)^n \longrightarrow (\mathbb{F}_2)^n$ we have $\mathbb{L}(f) \geq 2^{(n+1)/2}$. Further, equality holds if and only if the Fourier spectrum is $\{0, \pm 2^{(n+1)/2}\}$.

Functions for which equality holds are known as Almost Bent (AB) functions.
We saw that x^3 has this spectrum if n odd (BCH code weight distribution).

For n even, $\mathbb{L}(x^3) = 2\sqrt{q}$
(BCH spectrum is $\{0, \pm 2^{n/2}, \pm 2^{(n+2)/2}\}$)
For $n = 8$, the bound says $\mathbb{L}(f) \geq 2^{9/2}$ so at least 23.
The inverse function, and x^3, have $\mathbb{L}(f) = 32$. Best (smallest) known. Conjectured to be optimal. (In general, conjecture that $2^{(n+2)/2}$ is best possible linearity for n even.)

Example: Back to Coding Theory

Consider the binary cyclic code C_f of length $2^n - 1$ with parity check matrix

$$\begin{bmatrix} \cdots & x & \cdots \\ \cdots & f(x) & \cdots \end{bmatrix}$$

It can be shown that the codewords in the dual code $C_f^\perp$ are the vectors $(\ldots\ldots, \mathrm{tr}(bf(x) + ax), \ldots\ldots\ldots)$ over all $a, b \in \mathbb{F}_{2^n}$.
Thus the weight $w_{a,b}$ of this codeword is given by

$$2^n - 2w_{a,b} = \sum_{x \in L} (-1)^{\mathrm{tr}(bf(x)+ax)}$$

which is a Fourier coefficient of f !

More generally, the weight distribution of the code is given by the Fourier spectrum of f.
Exercise: if the dual code has only three weights, show that the *distribution* is determined. (Hint: We know $d(C_f) \geq 3$, and use the MacWilliams identities to get three equations in three unknowns.)

Fourier Spectra and APN Functions

If n is odd,

$$\mathbb{L}(f) = 2^{(n+1)/2} \implies \text{APN}$$

Proof: We know the entire weight distribution of $C_f^{\perp}$ because equality holds in the bound. (using Sidelnikov, Chabaud-Vaudenay theorem and previous exercise)
Use the MacWilliams identities. Get the weight distribution of C_f, and it must be the same as the BCH code. In particular we must have $d = 5$ and f is APN.
Moral: Coding theory has useful tools!

Almost all known APN functions have the same Fourier spectrum as x^3.

Algebraic Curves

Consider

$$\sum_{x \in L} (-1)^{\mathrm{tr}(bf(x)+ax)}$$

To evaluate this, we want to know how often $\mathrm{tr}(bf(x) + ax)$ is 0.
Elements of trace 0 have the form $y^2 + y$.
So we want the number of solutions (x, y) to

$$y^2 + y = bf(x) + ax.$$

This is an algebraic curve defined over L.
We want the number of rational points on this curve.
Example: $f(x) = x^{-1}$, or x^3, the curve is an elliptic curve!
The number of rational points on elliptic curves was determined by Deuring, Waterhouse.
x^{-1} gives an *ordinary* elliptic curve.
x^3 gives a *supersingular* elliptic curve.
We recover the earlier results about the Fourier spectrum/weight distribution.

Fourier Transform in General

We note that all the Fourier transform theory can be done much more generally.
Let $f : A \longrightarrow B$ be a function between finite abelian groups.
We use isomorphisms $\alpha \mapsto \chi_\alpha$ from A to $\hat{A}$ (the group of characters of A) and $\beta \mapsto \psi_\beta$ from B to $\hat{B}$.

We define the value of the Fourier transform of f at $\alpha \in A$ and $\beta \in B$ by

$$\hat{f}(\alpha, \beta) = \sum_{a \in A} (\psi_\beta \circ f)(a)\ \chi_\alpha(a) \quad \text{for all } \alpha \in A. \tag{1}$$

We define the *linearity* of f by

$$\mathbb{L}(f) = \max_{\alpha \in A, \beta \in B^*} |\hat{f}(\alpha, \beta)|. \tag{2}$$

Let tr denote the absolute trace map from $L = \mathbb{F}_{p^n}$ to $\mathbb{F}_p$.
Let ζ be a primitive complex p-th root of unity.
Let $\widehat{L}$ denote the group of characters of the additive group of L.
The so-called canonical additive character on L is $\mu(x) = \zeta^{\mathrm{tr}(x)}$, and all elements of $\widehat{L}$ have the form $\mu_a(x) := \zeta^{\mathrm{tr}(ax)}$ for $a \in L$.
The Fourier transform of any complex-valued function F defined on L is the function $\widehat{F}$ defined on $\widehat{L}$ by

$$\widehat{F}(\mu_a) := \sum_{x \in L} F(x)\overline{\mu_a(x)} = \sum_{x \in L} F(x)\zeta^{-\mathrm{tr}(ax)}$$

for $a \in L$.
Usually we consider $\widehat{F}$ to be defined on L via the identification $a \leftrightarrow \mu_a$, and we write $\widehat{F}(a)$ instead of $\widehat{F}(\mu_a)$,

To a function $f : L \longrightarrow \mathbb{F}_p$ we associate the complex-valued function $F = \zeta^f$. Such a function f is called *bent* if F has $|\widehat{F}(a)|^2 = q$ for all a.
For a function $f : L \longrightarrow L$, the functions $f_b(x) = \mathrm{tr}(bf(x))$ are called the coordinate functions of f, for $b \in L$.
Continuing our notation, we let $F_b(x) = \zeta^{\mathrm{tr}(bf(x))}$.
The *Fourier spectrum* of f (or F) is the set of all values of the Fourier transform over all coordinate functions:

$$\Lambda_f := \{\widehat{F}_b(a) : a, b \in L, b \neq 0\}.$$

One can study the Fourier spectrum of PN functions, and so on.