On the measurement of the (non)linearity of Costas permutations

Konstantinos Drakakis, *University College Dublin, Ireland*

ABSTRACT. We study several criteria for the (non)linearity of Costas permutations, with or without the imposition of additional algebraic structure in the domain and the range of the permutation, aiming to find one that successfully identifies Costas permutations as more nonlinear than randomly chosen permutations of the same order.

KEYWORDS: *Algebraic constructions, APN permutations, Costas arrays, Costas permutations, (non)Linearity measures*

Costas arrays, namely, square arrangements of dots and blanks such that there lies exactly one dot per row and column, and such that no four dots form a parallelogram and no three dots lying on a straight line are equidistant, appeared for the first time in 1965 in the context of SONAR detection (Costas, 1965, 1984), when Costas, disappointed by the poor performance of SONAR systems, used them to describe a novel frequency hopping pattern for SONARs with optimal auto-correlation properties. About two decades later, Professor Solomon Golomb published two generation techniques for Costas permutations, both based on the theory of finite fields, known as the Welch and the Golomb method, respectively (Golomb, Taylor, 1984; Golomb, 1984; Drakakis, 2006). These are still the only general construction methods for Costas permutations available today. Despite the intensive mathematical research dedicated to Costas arrays in the last two decades, many key questions about them remain unresolved, and most notably the issue of their existence: do Costas arrays exist for all orders? There is currently no order known for which Costas arrays provably do not exist, while the two smallest orders for which no Costas arrays are known are 32 and 33 (Drakakis, 2006).

An interesting application of Costas arrays in cryptography was discovered when it was shown that Welch Costas arrays are Almost Perfect Nonlinear (APN) permutations (Drakakis et al., 2009). This prompted further an investigation of the nonlinearity

of Welch Costas permutations, in the sense defined in Carlet and Ding (2004) and Pott (2004), whereby Welch Costas permutations were interpreted as mappings on $\mathbb{Z}_n$, the group of integers modulo n, and were indeed shown to exhibit high nonlinearity among all such functions/permutations (Drakakis et al., 2010). Costas permutations, however, are not defined over $\mathbb{Z}_n$, but rather over $[n] \subset \mathbb{N}$, the set of the first n nonnegative integers, on which no group structure is imposed. The object of this work is to investigate the correct interpretation and calculation of the (non)linearity of a Costas permutation, and, by extension, of any discrete function, in this context. What does it mean for a discrete function to be linear? How can the concept of linearity be quantified? Can this quantification benefit, in the case of functions on $[n]$, from the fact that such functions can be extended to functions on $\mathbb{Z}_n$? Assuming that this latter extension exhibits indeed high nonlinearity, can we infer that the original function on $[n]$ is also highly nonlinear (according to some appropriate definition)?
In what follows we will study several (non)linearity criteria, and more specifically their performance on Costas permutations versus collections of random permutations. In order to maintain compatibility between them and to be able to compare them, we will use Costas permutations of order 15 (but also 16 and 27 on some occasions) as a test case and recurrent example. The conclusions drawn have, of course, been verified on a wider range of orders.

Costas permutations, APN functions, and linearity

In this section we provide some background information on Costas permutations and APN functions. We will note, in particular, that, though the definitions of Costas and APN permutations appear deceptively similar, there are nonetheless important differences one has to pay attention to.

The definitions

In what follows, let $[n]$ denote the set $\{0, 1, \ldots, n-1\}$, and $\mathbb{Z}_n$ the additive group of integers modulo n, $n \in \mathbb{N}^*$; in other words, $[n]$ and $\mathbb{Z}_n$ differ just by the imposition of an algebraic structure on the latter, which makes it a ring. We are now ready to define the Costas permutation.

Definition 2.1. Consider a bijection $f : [n] \to [n]$; is a Costas permutation if and only if:

$\forall i, j, k$ such that
$i, j, i + k, j + k \in [n], \quad f(i + k) - f(i) = f(j + k) - f(j) \Rightarrow i = j$ or $k = 0.$ (2.1)

An alternative yet fully equivalent way to state this condition is to stipulate that, for any $k \in [n]^*$ and any $l \in [n]$, the equation

$$f(i + k) - f(i) = l, \quad i \in [n - k], \tag{2.2}$$

has at most one root i.

A permutation f corresponds to a permutation array $A_f = [a^f_{i,j}]$ by setting the elements of the permutation to denote the positions of the (unique) 1 in the corresponding column of the array, counting from top to bottom: $a^f_{f(i),i} = 1$. It is customary to represent the 1*s* as "dots" and the 0*s* as "blanks". From now on the terms "array" and "permutation" will be used interchangeably, in view of this correspondence.
The Costas property is invariant under horizontal and vertical flips, as well as transposition (and therefore also under rotations of the array by multiples of 90°, which can be expressed as combinations of the previous two operations), hence a Costas array gives birth to an equivalence class that contains either eight Costas arrays, or four if the array happens to be symmetric: this Costas array is then considered to be the unique representative of the equivalence class, and normally the array within the equivalence class that comes first in lexicographical order is selected for this purpose.
We now give the definition of the APN function.

Definition 2.2. $f : \mathbb{Z}_n \to \mathbb{Z}_n$ is APN if and only if, for any $\alpha \in \mathbb{Z}^*_n$ and $\beta \in \mathbb{Z}_n$, the equation

$$f(x + a) - f(x) = \beta, \quad x \in \mathbb{Z}_n, \tag{2.3}$$

has at most two roots x.

The relation of the two definitions becomes clearer if we also look at the definition of the Perfect Nonlinear (PN) function:

Definition 2.3. $f : \mathbb{Z}_n \to \mathbb{Z}_n$ is PN if and only if, for any $\alpha \in \mathbb{Z}^*_n$ and any $\beta \in \mathbb{Z}_n$, the equation

$$f(x + a) - f(x) = \beta, \qquad x \in \mathbb{Z}_n, \tag{2.4}$$

has at most one root (hence exactly one root) x.

We now see how close Definitions 2.1 and 2.2 are. When we focus exclusively on permutations, thought, we see that a PN permutation is a contradiction in terms: by Definition 2.3, for any α, there has to be an x such that $f(x + a) - f(x) = 0$, hence f cannot possibly be a permutation! Consequently, when studying permutations, we can only hope for the next best thing, namely, an APN permutation. Note that the definitions of a Costas permutation and of an APN function show that these two types of functions are far from being "linear", namely, far from being similar to a "straight line", since the distance vectors between pairs of points in the function graph are not, in general, allowed to be collinear.

Construction methods for Costas permutations

We will denote the finite field of q elements by $\mathbb{F}(q)$, where q is, in general, a power of a prime. Recall that $\mathbb{Z}_p$, p is a prime, is the finite field $\mathbb{F}(p)$.

Algorithm 2.4 (Exponential Welch Construction $W_1(p, g, c)$). Let p be a prime, g a primitive root of the finite field $\mathbb{F}(p)$, and $c \in [p - 1]$; the *exponential Welch permutation* of order $p - 1$ corresponding to g and c is defined by

$$f(i) = g^{i+c} \bmod p - 1, \qquad i \in [p - 1]. \tag{2.5}$$

The inverse of an Exponential Welch permutation (corresponding to the transpose of the corresponding Costas array) is a Logarithmic Welch permutation, which is itself a Costas permutation. The two permutation sets are distinct for $p > 5$ (Drakakis et al., 2009), implying that there are $2(p - 1)\,\varphi\,(p - 1)$ distinct Welch Costas permutations of order $p - 1$. Here φ denotes Euler's totient function: $\varphi(x)$, $x \in \mathbb{N}$, is the number of positive integers less than and relatively prime to x. In particular, there are no self-inverse W_1-permutations (i.e.,

corresponding to symmetric Welch Costas arrays) for $p > 5$.

Algorithm 2.5 (Golomb Construction $G_2(q, a, b)$). Let $q = p^m$ where p is a prime and $m \in \mathbb{N}^*$, and let a, b be primitive roots of the finite field $\mathbb{F}(q)$; the Golomb permutation f of order $q - 2$ corresponding to a and b is defined through the equation

$$a^{i+1} + b^{f(i)+1} = 1, \quad i \in [q-2]. \tag{2.6}$$

There are $\varphi^2(q - 1) / m$ distinct G_2-permutations of order $q - 2$ (Drakasis, 2006).

A comprehensive example

Consider the W_1-permutation f resulting from $p = 11$, $g = 2$, and $c = 0$. The values corresponding to $0, 1, \ldots, 9$ are, in that order $0, 1, 3, 7, 4, 9, 8, 6, 2,$ and 5. As mentioned above, f is an APN permutation when construed as a function from $\mathbb{Z}_{10}$ to $\mathbb{Z}_{10}$: in this case, all additions take place in arithmetic modulo 10, and we write, for example, that $f(5 + 7) – f(5) = f(2) – f(5) = 1 – 4 = 7$. Note that, after generating f, we forget all about the prime number p used to generate it (in this case $p = 11$): henceforth, all modulo operations take place in arithmetic modulo $p – 1 = 10$, which is the size of the group $\mathbb{Z}_{10}$, in both the domain and the range.
Considering, however, f as a Costas permutation from $[10]$ to $[10]$; we see that $f(5 + 7) – f(5) = f(12) – f(5)$ is undefined, because $f(12)$ is undefined. On the other hand, $f(2) – f(5) = 1 – 4 = -3$, because addition takes place now in the usual integer arithmetic, in both the domain and the range.

Linearity

What does it mean for a function f to be linear? In general, we will assume that both the domain $D(f)$ and the range $R(f)$ of the function are subsets of a ring R, and we will call f linear if and only if there exist three constants $\alpha, \beta, \gamma \in R$ such that $\alpha f(x) + \beta x = \gamma$ for all $x \in D(f)$, where addition and multiplication are as defined in R. It is important to note that, occasionally, R can be chosen in more than one way: for example, in the example shown in the previous section (*A Comprehensive Example*) we may choose either $R = \mathbb{Z}$ or $R = \mathbb{Z}_{10}$, and this leads to different functions f, neither of which is linear, however.

As another example, consider f defined on Drakakis et al. (2009), where the values corresponding to $0, 1, \ldots, 10$ are, in this order, $0, 2, 4, 6, 8, 10, 1, 3, 5, 7,$ and 9. This function is not linear under $\mathbb{Z}$-arithmetic. Since Drakakis, Gow, O' Carroll (2009) is closed under arithmetic modulo 11, however, we may choose $R = \mathbb{Z}_{11}$, in which case $f(x) = 2x$ for all $x \in \mathbb{Z}_{11}$, and is, therefore, linear.

Linearity measures for discrete functions

How can we quantify the linearity of a discrete function, and especially of a Costas permutation, in a meaningful way? There are essentially two different ways to proceed, according to whether we are willing/able to introduce some sort of an algebraic structure to the problem or not. Note that we will follow the convention of labeling the criteria we study below by $\mathbb{L}$ or $\mathbb{NL}$, according to whether an increase in the value returned by the criterion implies increased linearity or nonlinearity for the tested function, respectively.

Linearity without algebraic structure

Least squares

In this version of the problem we are given a set of n points $(x_i, y_i = f(x_i)), i \in [n]$ on the plane as an input, and we are asked to determine how closely they correspond to the graph of a linear function. The obvious course of action is to fit a line of the form $c_1 x + c_2 y = c$, where $c, c_1, c_2 \in \mathbb{R}$, according to some fitting criterion, and determine the error of the approximation. The smaller the error, the more "linear" f is. Perhaps the most frequently used fitting method used in such cases is the familiar least squares approximation.

Nonmodular phases

Within the same context, an alternative, completely different concept of linearity can be defined based on the distance vectors between pairs of points $(x_i - x_j, f(x_i) - f(x_j)), i, j \in [n], i > j$, where, without loss of generality we may assume that $x_i \geq x_j$ whenever $i \geq j$: the function f is linear if and only if all such distance vectors have the same phase on the plane. A way to quantify this idea in a continuous way is to determine the unit vector with each such phase, sum the vectors, and find the length of the vector sum.

In other words, we consider

$$\sum_{i>j} \exp\,(i\varphi_f(x_i, x_j)), \qquad \varphi_f(x_i, x_j) := \angle\,(x_i - x_j, f(x_i) - f(x_i)). \tag{3.1}$$

As there are $n(n-1)/2$ such vectors in total, the length of the vector sum will be $n(n-1)/2$ when f is linear and less than that otherwise. The normalized

$$\mathbb{L}(f) = \frac{2}{n(n-1)} \left| \sum_{i>j} \exp\,(i \angle\,(x_i - x_j, f(x_i) - f(x_j))) \right| \tag{3.2}$$

is then a number between 0 and 1: the larger it is, the more linear f is. In particular, since each phase is confined to $(-\pi/2, \pi/2)$, we may substitute $\angle(u, v)$ by $\tan^{-1}(v/u)$. Given that

$$e^{iu} = \cos(u) + i\sin(u), \quad \cos(u) = \frac{1}{\sqrt{1 + \tan^2(u)}}, \quad \sin(u) = \frac{\tan(u)}{\sqrt{1 + \tan^2(u)}}, \tag{3.3}$$

we can write

$$\begin{aligned} \mathbb{L}(f) &= \frac{2}{n(n-1)} \left| \sum_{x>y} \frac{1 + i((f(x) - f(y))/(x - y))}{\sqrt{1 + ((f(x) - f(y))/(x - y))^2}} \right| \\ &= \frac{2}{n(n-1)} \left| \sum_{x>y} \frac{x - y + i(f(x) - f(y))}{\sqrt{(x - y)^2 + (f(x) - f(y))^2}} \right| \end{aligned} \tag{3.4}$$

The log-ratio

In order to obtain a more sensitive measure of linearity, we observe that if $f(x) = ax + \beta$, we would get

$$\begin{aligned} \sum_{x>y} \frac{x - y}{\sqrt{(x - y)^2 + (f(x) - f(y))^2}} &= \frac{n(n-1)}{2} \frac{1}{\sqrt{1 + \alpha^2}}, \\ \sum_{x>y} \frac{f(x) - f(y)}{\sqrt{(x - y)^2 + (f(x) - f(y))^2}} &= \frac{n(n-1)}{2} \frac{\alpha}{\sqrt{1 + \alpha^2}}. \end{aligned} \tag{3.5}$$

For a general function f, these two expressions yield two estimates for α, namely, α_1 and α_2, where we assume $\alpha_1 > \alpha_2$ without loss of generality. The log-ratio $\mathbb{NL}_c(f) := ln(\alpha_1/\alpha_2) \geq 0$ is a kind of condition number for f: the larger it is, the more nonlinear f is, so we can use

this log-ratio as a measure of the nonlinearity of f.

Linearity with algebraic structure

Let us reformulate the ideas presented above regarding distance vectors and their phases in the special case of a function $f : [n] \to [n]$. An indication of the linearity of f is the degree to which a constant multiple of $x - y$ approximates (a constant multiple of) the difference $f(x) - f(y)$, the approximation holding for all pairs $(x, y), x, y \in [n]$: in other words, we consider the functions $F(\alpha, \beta; x, y) = \beta(f(x) - f(y)) - \alpha(x - y)$ and we determine whether any specific choice of the parameters α and β leads to values that lie uniformly "close to" 0 for all pairs (x, y), according to some proximity criterion.
A possible proximity criterion is again to apply "phase modulation", namely, to allow the values of $F(\alpha, \beta; x, y)$ to multiply the phase of complex exponential $\exp(i\varphi)$, which represents a vector of unit

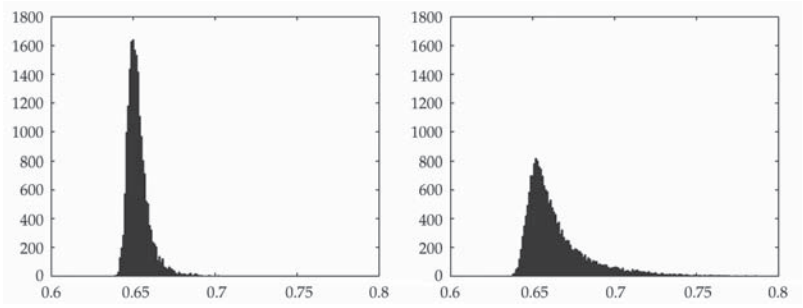

Figure 1. The histograms of all Costas permutations of order 15 (a) and an equinumerous collection of randomly chosen permutations of order 15 (b), according to $\mathbb{L}$: Costas permutations are shown to be more nonlinear

length, and then find the length of the aggregate vectors and choose the longest one. This we define as the (square of the) linearity of f:

$$\mathbb{L}^2_a(f) = \sup_{\alpha, \beta \in \mathbb{R}} \sum_{x, y \in [n]} e^{i[\beta(f(x) - f(y)) - \alpha(x - y)]\varphi} = \sup_{\alpha, \beta \in \mathbb{R}} \left| \sum_{x \in [n]} e^{i[\beta f(x) - \alpha x]\varphi} \right|^2 . \qquad (3.6)$$

Clearly, f is linear if and only if $\mathbb{L}(f) = n$.
Since f is an integer function, and remembering that our ultimate goal is to introduce algebraic structure in the problem at some point, it makes sense to confine α and β to integer values as well. Choosing further $\varphi = 2\pi / N, N \in \mathbb{N}^*$, we effectively impose a modulo N addition and consider F mod N instead of F:

$$\mathbb{L}_N(f) = \max_{\alpha,\beta} \left| \sum_{x\in[n]} e^{i(2\pi/N)[\beta f(x) - \alpha x]} \right|. \tag{3.7}$$

Sometimes (Drakakis et al. 2010) it even makes sense to generalize the previous expression slightly and use two different integer parameters M and N as follows:

$$\mathbb{L}_{M,N}(f) = \max_{\alpha,\beta} \left| \sum_{x\in[n]} e^{i2\pi[\beta(f(x)/N) - \alpha(x/M)]} \right|, \tag{3.8}$$

though we will mostly focus on the simple case $M = N$ from now on. So far we have not related N and n; how should we choose N for a given n? A first possibility is dictated by the extension of f to a function on $\mathbb{Z}_n$, that is, $f : \mathbb{Z}_n \to Z_n$, in which case the obvious choice would be $N = n$. Alternatively, considering still f as a function on $[n]$, both $x - y$ and $f(x) - f(y), x, y \in [n]$ range from $-(n-1)$ to $n-1$ included, so the range includes $2(n-1) + 1 = 2n - 1$ distinct values and hence it suffices to choose $N = 2n - 1$, if our goal is to avoid any fold-over of values within this range.

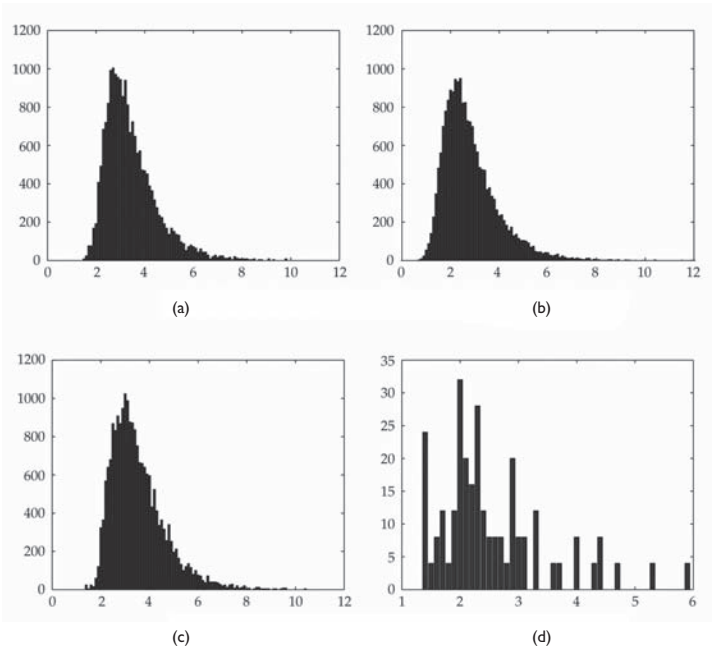

Figure 2. (a, b) the log-ratio histograms of all Costas permutations of order 15 (a) and an equinumerous collection of randomly chosen permutations of order 15 (b); Costas permutations are shown to be more nonlinear. (c, d) the log-ratio histograms of all Costas permutations of order 16 (c) and of all algebraically constructed Costas permutations of order 16 (d); algebraically constructed Costas permutations seem to be amongst the most linear ones

Finally, note that if $f : \mathbb{Z}_{p-1} \to \mathbb{Z}_p^*$ is a W_1-permutation, it makes sense to choose either $M = N = p - 1$, as both the domain and the range contain $p - 1$ elements, or $M = p - 1$ and $N = p$, as these parameters reflect the natural modulo arithmetic in the domain and the range, respectively (both cases were studied in Drakakis et al., 2010).
Let us finish this discussion by mentioning that $1 - \mathbb{L}_n(f)/n$ has already been proposed as ameasure of the nonlinearity of $f : \mathbb{Z}_n \to \mathbb{Z}_n$ in the literature (Carlet, Ding, 2004; Pott, 2004), though, in our opinion, the presentation therein was much less straightforward and intuitive than the one given here.

Results

In this section we discuss the results obtained for each (non)linearity criterion through simulation. Simulation has been used extensively in recent times for the study of the properties of Costas arrays (see, e.g., Drakakis, 2007, 2008).

n	min $\mathbb{L}$	max $\mathbb{L}$	mean $\mathbb{L}$	std $\mathbb{L}$
3	2.4972	2.4972	2.4972	0
4	2.9750	3.7069	3.5083	0.3046
5	3.4886	4.3459	3.8902	0.2741
6	3.7244	5.1008	4.3918	0.3834
7	4.1534	5.1786	4.6496	0.2988
8	4.4234	5.9241	5.0275	0.2964
9	4.7015	6.8218	5.5497	0.3621
10	4.9301	7.4756	5.8875	0.3519
11	5.1913	7.6610	6.2660	0.3775
12	5.4815	8.5341	6.5952	0.3965
13	5.9015	8.7261	6.9115	0.4053
14	6.2014	9.4388	7.2270	0.4082
15	6.2186	9.8519	7.5333	0.4320
16	6.5159	11.0186	7.8509	0.4521
17	6.9582	11.4045	8.1315	0.4465
18	7.1846	12.9660	8.4167	0.4811
19	7.5165	12.4471	8.7136	0.4935
20	7.7579	11.8479	8.9461	0.4815
21	7.9894	13.9195	9.2645	0.5443
22	8.1603	15.2323	9.8122	0.9286
23	8.6376	14.7480	9.9175	0.8786
24	8.7028	15.9756	10.2508	1.0853
25	9.7256	16.5861	11.1684	1.5209
26	9.4019	14.8275	11.5788	1.6035
27	10.2502	16.8729	12.1096	1.4790

Table 1. Linearity results for all Costas permutations of orders $3 \leq n \leq 27$: the columns correspond from left to right to n, the minimal and maximal linearities observed, and the mean and standard deviations of the linearity

Least squares

When f is a Costas permutation of order n, linear least squares fitting fails to reveal any meaningful information, precisely because the points are very dispersed on the $n \times n$ square, owing to the Costas property. Computer simulations confirm our expectations in that the line fitted by least squares is invariably either horizontal or vertical, while the line fitted by orthogonal least squares, namely the variant of the method where the sum of the square distances of the points from the fitted line is minimized, yields invariably either $y = x$ or $y = n - 1 - x$ as the fitted line.

To conclude, Costas arrays are so far from being linear that it makes no sense to measure how far from linearity they are using this criterion.

Nonmodular phases

The real part of the vector sum (3.4) is in general much larger than the imaginary part, precisely because we always choose $x > y$, so the real parts of the summands add constructively.

This, in turn, implies that this criterion is not sensitive enough. For example, Figure 1 shows the histograms of $\mathbb{L}$ over all Costas arrays of order 15 and over an equinumerous collection of randomly chosen permutations of order 15: though the histograms look different, the range of the former lies entirely within the range of the latter, so this criterion is not sensitive enough to determine that Costas permutations are more nonlinear than random permutations.

(a)

p	min $\mathbb{L}$	max $\mathbb{L}$	mean $\mathbb{L}$	std $\mathbb{L}$
7	3.8344	3.8344	3.8344	0
11	4.7015	6.5598	5.9362	0.6950
13	5.1913	7.5414	6.4948	0.8282
17	6.2187	9.3125	7.8076	0.7503
19	7.6955	11.4045	8.9887	0.8291
23	8.2347	13.9195	9.9797	1.1294
29	10.2502	16.8729	12.0515	1.4318
31	10.9677	16.0584	13.0976	1.4772
37	11.2892	20.5871	14.8351	2.1180
41	12.7289	21.6292	16.1072	2.2266
43	12.6109	26.9280	17.1595	2.7143
47	13.6820	28.3311	18.3678	2.5944
53	14.7483	29.0807	20.7323	3.7493
59	16.5240	33.8105	23.1238	3.7493
61	18.0203	35.1987	23.8661	3.5871
67	17.9924	38.6901	26.2275	4.0515
71	19.2378	36.4151	27.6938	3.9807
73	18.1823	43.1096	28.6351	4.7167
79	19.4471	41.3463	30.9327	4.2014
83	21.2388	45.4478	32.5626	4.2763
89	21.3421	50.3580	35.0311	4.6115
97	24.6606	53.2888	38.2512	4.6876
101	27.1879	52.6007	39.8379	5.0796
103	25.4178	55.2598	40.6539	5.5482
107	27.3222	57.9382	42.2601	5.4518
109	24.9826	59.9810	43.1139	5.5113
113	29.6132	60.6273	44.7029	5.6808
127	32.5151	70.1353	50.4963	5.3487
131	31.2879	70.7198	51.9830	5.7581
137	37.2955	71.0958	54.4337	5.7566
139	39.2845	71.4561	55.2887	6.3653
149	40.5434	76.0981	59.2918	6.0764
151	37.4170	81.0333	60.1070	6.0674

Table 2 (a). Linearity results for W_1(a) and G_2(b) -permutations generated in $\mathbb{F}(p)$, $7 \le p \le 151$: the columns correspond from left to right to p, the minimal and maximal linearity observed, and the mean and standard deviations of the linearities

(b)

Table 2 (b).

p	min $\mathbb{L}$	max $\mathbb{L}$	mean $\mathbb{L}$	std $\mathbb{L}$
7	3.7880	4.8439	4.4465	0.4119
11	5.5053	7.4756	6.1462	0.4734
13	6.5154	8.5341	7.3052	0.6860
17	7.0501	11.0186	8.5779	0.8158
19	7.5172	12.9660	9.2807	1.1130
23	8.9254	15.2323	10.6936	1.2187
29	10.3331	18.3192	12.9565	1.8357
31	10.9192	19.9770	13.5635	1.6404
37	11.9643	24.3642	15.7800	2.5823
41	12.7863	25.1720	17.1771	2.8822
43	12.7419	25.3322	17.8591	3.0673
47	14.0212	28.1627	19.1228	3.2689
53	15.0066	33.3970	21.5939	3.8445
59	16.4316	35.5213	23.5966	4.3340
61	16.2902	36.1331	24.3684	4.5547
67	17.4479	39.8553	26.7267	4.9172
71	18.4225	40.7985	28.2850	5.0988
73	18.3244	42.1267	29.0482	5.1430
79	18.5635	47.8331	31.4486	5.5526
83	20.0101	49.1780	33.0352	5.6957
89	21.2534	51.4717	35.4218	5.5741
97	22.1986	57.7464	38.6617	6.2288
101	22.9944	56.8656	40.2868	6.6020
103	25.5362	57.2796	41.0855	6.5606
107	24.0941	63.6045	42.7078	6.5930
109	24.7888	64.2623	43.5123	6.6427
113	26.5017	64.3054	45.1397	7.0107
127	35.1980	65.9180	50.7644	6.1753
131	30.9777	73.9441	52.4322	7.4101
137	32.7740	77.6968	54.8584	7.5626
139	31.6170	79.5067	55.6664	7.5732
149	35.0306	83.2909	59.7193	7.8052
151	36.3423	84.3540	60.5394	8.0897

The log-ratio

What if $\mathbb{L}_c$ is used instead of $\mathbb{L}$? The log-ratio histograms for all Costas permutations of order 15 and an equinumerous collection of random permutations of order 15, as well as the logratio histograms for all Costas permutations of order 16 and for all algebraically constructed Costas permutations of order 16 are shown in Figure 2. Costas permutations are indie found to be more nonlinear than random ones, even if only slightly so: though the random permutations histogram contains a few outliers at higher values, its main body lies clearly at smaller values compared to the Costas permutations histogram. Similarly, algebraically constructed Costas permutations are observed to be, on average, some of the most linear Costas permutations.

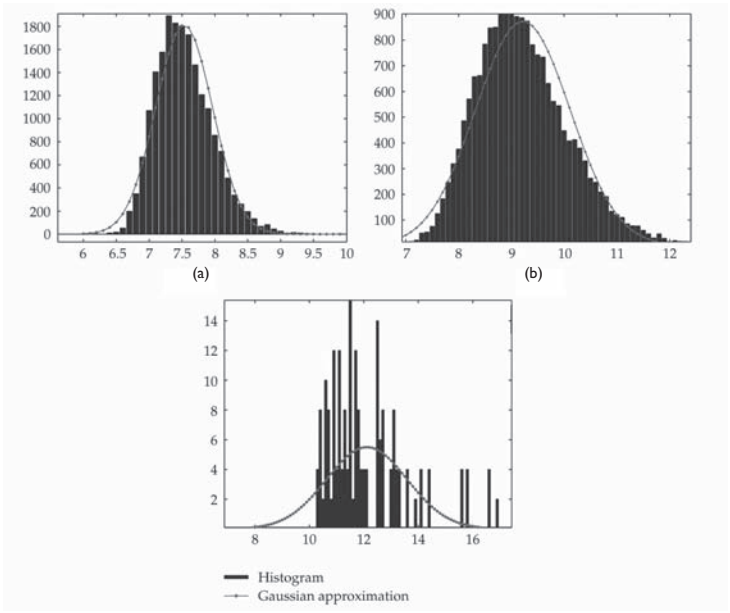

Figure 3. The linearity histogram of all Costas permutations of order 15 (a) is well approximated by the Gaussian of the same mean and variance, but the corresponding histogram of order 27 (c) is not, due to the small number of samples. Furthermore, the linearity criterion is efficient: histograms show that the linearity of Costas permutations of order 15 is clearly less than that of an equinumerous collection of randomly chosen permutations of order 15 (b)

Linearity with algebraic structure

We computed the linearity of several families of Costas permutations, using $\mathbb{L}_{2n-1}$ as the measure of linearity, n being the order of the Costas permutation. More specifically, we focused on the families of all Costas permutations of order 27 and below (Table 1), and on the families of W_1- and G_2-permutations generated in $\mathbb{F}(p)$, $3 \leq p \leq 151$

(Table 2). For each family we recorded the minimal and maximal linearities found, the mean linearity and the standard deviation.
As a general observation, the linearity histograms for all families are well approximated by Gaussian distributions (see, e.g., Figure 3), provided the families contain enough Costas permutations (at least a few hundred). Furthermore, the mean linearities $\overline{\mathbb{L}}_W(n)$ and $\overline{\mathbb{L}}_G(n)$ for W_1- and G_2-permutations of order n, respectively, seem to increase

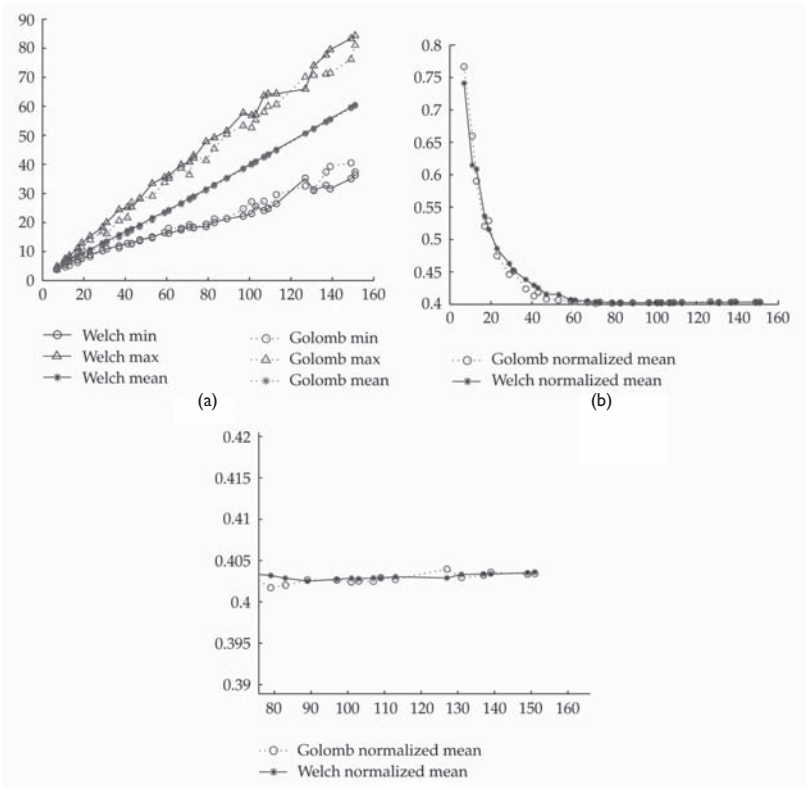

Figure 4. (a) plot of the minimal, maximal, and mean linearities for the Welch and Golomb family generated in $\mathbb{F}(p)$, $3 \le p \le 151$. (b) plot of the mean linearity divided by the order, indicating convergence near 0.4. (c) a detail of the tail of the previous plot

asymptotically linearly with n (see Figure 4): $\overline{\mathbb{L}}_G(n) \approx \overline{\mathbb{L}}_W(n) \approx 0.4035n$. Furthermore, it is clear from Figure 3 that $\overline{\mathbb{L}}_{2n-1}$ successfully distinguishes Costas permutations from random permutations, assigning on average smaller linearity to the former.

Conclusion

We proposed various (non)linearity measures for Costas permutations, divided in two broad categories, according to whether we are willing to impose some algebraic structure on the domain and the range or not. Amongst the measures that do not take advantage of any algebraic structure, the linear least squares fit was found inappropriate, as it was completely insensitive to the input, the nonmodular phases criterion was found not to be sensitive enough, while the log-ratio performed adequately in terms of distinguishing Costas permutations from randomly chosen permutations of the same order and correctly deciding that the former are more nonlinear than the latter; it also suggested that algebraically constructed Costas permutations are amongst the most linear Costas permutations. On the other hand, when the difference vectors are combined with an underlying modulo structure, the resulting criterion is sensitive enough to recognize that Costas permutations are less linear than randomly chosen permutations of the same order.

References

All URLs checked December 2010

Carlet Claude, Ding Cunseheng (2004), *Highly nonlinear mappings*, "Journal of Complexity", V. 20, n. 2-3, pp. 205-244

Costas John P. (1984), *A study of detection waveforms having nearly ideal range - Doppler ambiguity properties*. Proceedings of the IEEE, V. 72, n. 8, pp. 996-1009

Costas John P. (1965), *Medium constraints on sonar design and performance*, Techical Report R65EMH33, GE Co

Drakakis Konstantinos (2008), *Three challenges in Costas arrays*, "Ars Combinatoria", V. 89, pp. 167-182

Drakakis Konstantinos (2007), *Data mining and costas arrays*, "Turkish Journal of Electrical Engineering and Computer Sciences", V. 15, n. 1, pp. 67-76

Drakakis Konstantinos (2006), *A review of Costas arrays*, "Journal of Applied Mathematics", V. 2006, Article ID26385.

Drakakis Konstantinos, Requena Veronica, McGuire Gary (2010), *On the nonlinearity of exponential Welch costas functions*, "IEEE Transactions on Information Theory", V. 56, n. 3, pp. 1230-1238

Drakakis Konstantinos, Gow Rod, McGuire Gary (2009), *APN permutations on $\mathbb{Z}_n$ and Costas arrays*, "Discrete Applied Mathematics", V. 157, n. 15, pp. 3320-3326

Drakakis Konstantinos, Gow Rod, O'Carroll Liam (2009), *On the symmetry of Welch – and Golomb – constructed Costas arrays*, "Discrete Mathematics", V. 309, n. 8, pp. 2559-2563

Drakakis Konstantinos, Rickard Scott, Beard James K., et al. (2008), *Results of the enumeration of Costas arrays of order 27*, "IEEE Transactions on Information Theory", V. 54, n. 10, pp. 4684-4687

Golomb Solomon W. (1984), *Algebraic constructions for Costas arrays*, "Journal of Combinatorial Theory Series A", V. 37, n. 1, pp. 13-21

Golomb Solomon W., Taylor Herbert (1984), *Constructions and properties of Costas arrays*, "Proceedings of the IEEE", V. 72, n. 9, pp. 1143-1163

Pott Alexander (2004), *Nonlinear functions in abelian groups and relative difference sets*, "Discrete Applied Mathematics", V. 138, n.1-2, pp. 177-193

Sintesi

Negli anni '60, per trovare una soluzione alle basse prestazioni dei SONAR utilizzati dalla flotta navale statunitense, l'ingegnere John Costas introdusse un nuovo schema per la scelta delle frequenze da utilizzare che comportava la variazione a salti, distaccandosi così dallo schema a variazione continua utilizzato fin a quel momento. Il nuovo sistema di permutazione delle frequenze di Costas (frequency hopping pattern) può essere illustrato facendo uso di una scacchiera nella quale le case nere devono sottostare a dei vincoli: primo, in ogni riga e ogni colonna può esserne presente una sola; secondo, le case devono esser poste in modo che con quattro qualsiasi di esse, non si possa ottenere un parallelogramma (pensando le case come vertici dello stesso) e che, prese tre allineate, quella centrale non si trovi perfettamente nel mezzo delle altre due. Utilizzando questi schemi e strutture, Costas era riuscito a

ottenere dei SONAR con ottime prestazioni. Tali strutture, affermatesi come vettori di Costas, altro non sono che casi particolari delle più note matrici di permutazioni e, pur essendo studiate oramai da quasi un lustro per la loro evidente utilità in ambito di indagine marina, nascondono ancora alcuni lati oscuri. Uno su tutti: è possibile costruirle per un qualsiasi ordine? In altre parole, se l'ordine, ovvero il numero di righe e di colonne della nostra scacchiera, indica le frequenze a disposizione, è possibile determinare vettori di Costas per un qualsiasi numero di queste, permettendo così sia un utilizzo ottimale da parte di un unico strumento e, insieme, un utilizzo multiplo di più macchinari. Attualmente, per esempio, pur essendo costruibili per un'infinità di ordini, mancano all'appello quelle per ordini relativamente bassi quali il 32 e il 33. Neanche le attuali prestazioni degli elaboratori sono riuscite a risolvere il problema essendo questo, per sua natura, esponenziale.

D'altra parte, l'interesse per questo vuoto nelle nostre conoscenze è aumentato da quando è stato messo in evidenza come una particolare tipologia di vettori di Costas (Welch-Costas) può essere efficacemente utilizzato nella crittografia, ovvero nella cifratura delle informazioni. La caratteristica che ne esalta la loro utilità in questo ramo della matematica (ma anche del quotidiano, vista l'importanza della crittografia negli acquisti on line, nella telefonia cellulare e nella protezione di documenti riservati, per citare alcuni esempi) è quella di essere permutazioni quasi perfettamente non lineari (APN, Almost Perfect Nonlinear). Considerando che la linearità può essere interpretata come un indice di regolarità e quindi della prevedibilità di un andamento., va da sé che un elevato discostamento dalla linearità è una delle caratteristiche che un sistema crittografico deve possedere.

È in questo contesto che si inseriscono i recenti risultati di Drakakis, i quali, pur necessitando di uno studio su un più ampio numero di ordini, consolidano la migliore efficacia nella cifratura dei messaggi quando si utilizzano strutture quali i vettori di Costas. Forse è un'ironia della sorte che tali strutture siano nate allo scopo di perfezionare la scelta e l'interpretazione dei messaggi da parte dei SONAR. O forse, più semplicemente, è un'ulteriore dimostrazione della "irragionevole efficacia della matematica nelle scienze naturali".