SUM OF PERMUTATIONS

BAI LIN

Abstract. With growing need of secure authentication, data confidentiality, vybersecurity has now become one of our inevitable daily topic. People even get their private information leaked just because they connect to the public Wifi in a Cafe. Cryptographic hash functions have a large scale of applications in data security and are commonly used to verify data authenticity. My research focuses on the study of the algebraic properties in their underlying structure that dictate the security of a cryptographic hash functions. In particular, we investigate the algebraic design requirements of the Grøstl hash function and its generalizations. Grøstl is an iterated hash function with a compression function built from two distinct permutations, crucial for preserving its security. Grøstl is one of the five finalists in the recent NIST SHA-3 competition and is the hash function that probably received the most intense cryptanalysis during the competition. Its elegant design and simplicity inspires continued high interest in the security features of this hash function.

1. Introduction

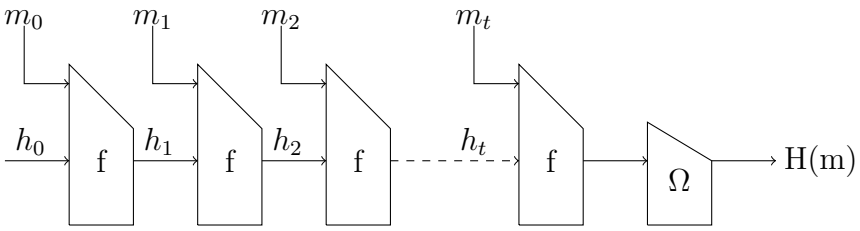
In the pursuit of efficient and provably secure constructions of cryptosystems, hash functions have emerged as important building blocks for many cryptographic purposes such as authentication protocols and digital signatures. Another important use of cryptographic hash functions is the generation of pseudo-random numbers. A hash function H maps an input message M of arbitrary length to a fixed-length hash value $h = H(M)$. Historically, the following three main security requirements have evolved (for a more formal treatment of these and related requirements we refer to [40]):

- *Collision resistance:* It is computationally infeasible to find two messages M_1 and M_2 with $M_1 \neq M_2$, which result in the same hash value $H(M_1) = H(M_2)$.

2010 *Mathematics Subject Classification.* 20B05, 20B30, 94A60, 11T71, 14G50.

- *Preimage resistance:* For a given hash value h, it is computationally infeasible to find any message M, which results in the given hash value $H(M) = h$.
- *Second preimage resistance:* For a given message M_1, it is computationally infeasible to find a second message M_2 with $M_1 \neq M_2$, which results in the same hash value $H(M_1) = H(M_2)$.

Most currently used hash functions are built form iterations of a compression function f using constructions such as Merkle-Damgard



$$h_0 = IV \; h_i := f(h_{i-1}, m_{i-1})$$

2. Preliminaries

2.1. **Group theoretical background.** In this section we present some background from the theory of permutation groups and finite fields which are used in this paper.

2.1.1. **Permutation groups.** For a finite set X, let $|X|$ denote the number of elements of X. For any nonempty finite set X with $|X| = n$, the set of all bijective mappings of X to itself is denoted by $\mathcal{S}_n$ and is called the *symmetric group* on X. A permutation $g \in \mathcal{S}_n$ is a *transposition* if g interchanges two elements $x, y \in X$ and fixes all the other elements of $X \setminus \{x, y\}$. A permutation $g \in \mathcal{S}_n$ is in *canonical form* if g maps the identity element to itself.

A permutation $g \in \mathcal{S}_n$ is in *canonical form* if g maps the identity element to itself.

A permutation $g \in \mathcal{S}_n$ is called an *odd* (*even*) permutation if g can be represented as a composition of an odd (even) number of transpositions[1].

[1]Note that in this terminology a cycle of even length is an odd permutation, while a cycle of odd length is an even permutation.

The set of all even permutations is a group under functional composition and is called the *alternating group* on X. The symbol $\mathcal{A}_n$ denotes the alternating group on a set X with $|X| = n$. The *degree* of a permutation group G over a finite set X is the number of elements in X that are moved by at least one permutation $g \in G$.

For any subgroup $G \leq \mathcal{S}_n$, for any $x \in X$, the set $orb_G(x) = \{\phi(x) : \phi \in G\}$ is called the *orbit* of x under G. The set $stab_G(x) = \{\phi \in G : \phi(x) = x\}$ is called the *stabilizer* of x in G. We will make use of the following well-known theorem, often called the Orbit-Stabilizer Theorem.

2.1.2. **Finite fields.** A structure $(\mathbb{F}, +, \cdot)$ is a *field* if and only if both $(\mathbb{F}, +)$ is an Abelian group with identity element 0_G and $(\mathbb{F} \setminus \{0_G\}, \cdot)$ is an Abelian groups and the law of distributivity of $\cdot$ over $+$ applies. If the number of elements in $\mathbb{F}$ is finite, $\mathbb{F}$ is called a *finite field*; otherwise it is called an *infinite field*.

It is known that every finite field has order p^n for some prime number p and some positive integer n. Such a field is called a *Galois field* of order p^n and is denoted by $\mathrm{GF}(p^n)$. The following classical fact from the theory of finite fields (see [23]) will be used.

Theorem 1. $\mathrm{GF}(p^{n_1}) \subseteq \mathrm{GF}(p^{n_2})$ *if and only if* n_1 *divides* n_2.

3. Cycle structure of the generalized Grøstl round functions

Let m, n, r be positive integers. The symbol $M_{m,n}(\mathrm{GF}(p^r))$ denotes the set of all $m \times n$ - matrices over $\mathrm{GF}(p^r)$. The elements of $\mathrm{GF}(p^r)^{mn}$ are defined as matrices $b \in M_{m,n}(\mathrm{GF}(p^r))$ with the mapping $t : \mathrm{GF}(p^r)^{mn} \to M_{m,n}(\mathrm{GF}(p^r))$, where $t(a) = b$ is defined by $b_{ij} = a_{ni+j}$, for $0 \leq i < m, 0 \leq j < n$. First we start with the analysis of the cycle structure of the component functions in the generalized Rijndael-like function.

3.1. Analysis of the AddRoundConstant-like function ($\sigma[k]$-function).

Definition 2. Let $\sigma[k] : M_{m,n}(\mathrm{GF}(p^r)) \to M_{m,n}(\mathrm{GF}(p^r))$ denote the mapping defined by $\sigma[k](a) = b$ if and only if $b_{ij} = a_{ij} + k_{ij}$ and $k \in M_{m,n}(\mathrm{GF}(p^r))$ for all $0 \leq i < m$, $0 \leq j < n$.

3.2. **The SubBytes-like function (λ-function).**

Definition 3. Let $\lambda : M_{m,n}(\mathrm{GF}(p^r)) \to M_{m,n}(\mathrm{GF}(p^r))$ denotes the mapping defined as a parallel application of $m \cdot n$ bijective S-box-mappings $\lambda_{ij} : \mathrm{GF}(p^r) \to \mathrm{GF}(p^r)$ and defined by $\lambda(a) = b$ if and only if $b_{ij} = \lambda_{ij}(a_{ij})$ for all $0 \leq i < m, 0 \leq j < n$.

Each S-box mapping consists of an inversion, multiplication by a fixed $A \in \mathrm{GF}(p^r)$, and addition of a fixed element $B \in \mathrm{GF}(p^r)$ i.e. it is a mapping of the form $Ax^{-1} + B$ where $A, B \in \mathrm{GF}(p^r)$ are fixed. For convenience we define this map on all of $\mathrm{GF}(p^r)$ so that it maps 0 to B, and any nonzero x to $Ax^{-1} + B$.

3.3. **The ShiftBytes-like function (π-function).**

Definition 4. Let $\pi : M_{m,n}(GF(p^r)) \to M_{m,n}(GF(p^r))$ denote the mapping for which there is a mapping $c : \{0, \cdots, m-1\} \to \{0, \cdots, n-1\}$ such that $\pi(a) = b$ if and only if $b_{ij} = a_{i(j-c(i))} \bmod n$ for all $0 \leq i < m$, $0 \leq j < n$.

4. Group generated by the generalized Grøstl-like round function

5. Sum of Permutations

5.1. **Motivation and Preliminary Definitions.** The Grøstl hash function iterates a sum of two permutations, P and Q, of the message space over the Galois field $GF(2^8)$. In order for the overall function to be secure and collision-resistant, it is imperative that this sum $P \oplus Q$ be as close to a permutation as possible. This begs the mathematical question: under what conditions, and how often, is the sum of two permutations a permutation or close to a permutation? We begin by defining the sum of two permutations.

Definition 5. Let a $= (a_1, \cdots, a_n)$ and b $= (b_1, \cdots, b_n)$ be permutations in one-line notation over a group G under addition. The *sum* of those permutations, $a+b$, is $(a_1+b_1, \cdots, a_n+b_n)$.

Because the Galois fields used in cryptography are very large, finding 'near'-permutations is also useful for application. Below, we formalize the concept of a 'near'-permutation. In the following definitions, f is a function on a finite set of n elements.

Definition 6. The function f is said to be *k-near* if $|\text{Domain}(f)| - |\text{Range}(f)| = k$.[2]

Definition 7. The k^{th}*-step* of f is f^k where f^k is k compositions of f with itself and $f^1 = f$.

Definition 8. The *terminal size* of f is $min(|\text{Range}(f^k)|)$ for all $k \in \mathbb{N}$.

Lemma 9. *Let f be a function of size ω. The terminal size of f is $|Range(f^{\omega-1})|$.*

Proof. Let $f : S \to S$, let $A \subseteq S$ denote the largest subset of S on which f acts as a permutation, and denote the size of f by $\omega := |range(f)|$. Since S is finite f must contain a cycle making A nonempty. By definition $a \in A$ implies $f(a) \in A$, which shows no power of f will have a range smaller than $|A|$. Thus it suffices to show that $x \in S \backslash A$ implies $f^{\omega-1}(x) \in A$.

Suppose $x \in S \backslash A$. Consider the indexed set $T = \{t_i = f^i(x) | i \in \{1, 2...(\omega)\}\}$. Suppose T contains a pair of repeated elements $t_j = t_k$ corresponding to $f^j(x) = f^k(x)$. Then each of t_j and t_k is part of a cycle, making them elements of A and implying that all subsequent elements of T are also in A. We assume instead then that T is composed of ω distinct elements of $S \backslash A$. But then $|S \backslash A| = |range(f)|$, which would make A empty.

Thus the terminal size of a function f on a finite set S corresponds to the magnitude of the largest subset of S on which f acts as a permutation. □

The size of a function is our primary means of assessing how well it emulates a permutation. Over some groups, it is impossible for two permutations to sum to a permutation. Therefore, the best possible outcome may be a 1-near function. However, the size does not entirely capture the effectiveness of the function.

Consider the two permutations, written in one-line notation, $\pi_1 = (2345672)$ and $\pi_2 = (2345677)$. Despite that these two functions have the same size, π_1 is significantly stronger. Their graphical representation, shown in Figure 1, is more telling than their one-line notation.

[2]Stones and Wanless 'A Congruence Connecting Latin Rectangles and Partial Orthomorphisms' discuss a definition of partial orthomorphisms that generalizes our notion of k-near. Bedford (1991) 'Transversals in the Cayley Table of the non-Cyclic Groups of order 8 discusses a near-complete mapping and near-orthomorphism that closely mirror but slightly generalize our notion of a 1-near permutation.

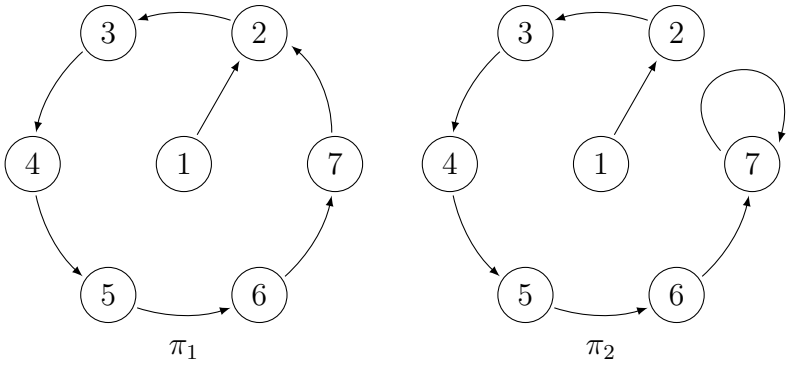


FIGURE 1. Graphical interpretation of π_1 and π_2.

Notice that the $k^{th} - step$ of the permutations can be viewed as moving k steps forward from each vertex. Due to cycles that exist in these graphs, some elements are moved into those cycles and can never escape. This is why the size of the function decreases after repeated self-composition. Here we see that π_1, in its terminal state, (i.e. after n-1 compositions) will have size 6 but π_2 will have size 1. Note that the terminal size is equal to the number of vertices that are part of some cycle.

In examining sums of all pairs of permutations over a group, a few simplifications can be made. Below, we give a useful definition and a theorem that greatly reduces the set of permutation pairs necessary to consider.

Definition 10 (Canonical form)**.** A permutation π over a group G is said to be in *canonical form* if π maps the identity element to itself.

5.2. **An Equivalent Problem.** We have been able to reduce the problem to finding some π such that $\theta + \pi$ is a permutation. This task is equivalent to finding a transversal in a Latin square. In particular, the Latin square will be the Cayley table of the group that the permutations are over.

Definition 11. A *Latin square* of order n is an $n \times n$ array of n symbols such that each symbol appears exactly once in each column and row.

Definition 12. A *transversal* of a Latin square is a set of n entries such that no two entries are from the same row or column and share the same symbol.

Definition 13. A permutation θ of the elements of the quasigroup $(Q, \oplus)$ is a *complete mapping* if $\eta : Q \mapsto Q$ defined by $\eta(x) = x \oplus \theta(x)$ is also a permutation. The permutation η is known as an *orthomorphism* of $(Q, \oplus)$ [47].

Definition 14. A *near complete mapping* of a group $(G, \odot)$ is a bijection $\theta : g \mapsto \theta(g)$ from $G \setminus \{h\}$ to $G \setminus \{e\}$ such that the mapping $\phi : g \mapsto g\theta(g)$ is again a one-to-one mapping from $G \setminus \{h\}$ to $G \setminus \{k\}$. Here $h \neq e$ and k are fixed elements of G and e is the identity element.

In Ian Wanless's survey on transversals in latin squares [47] he provides the following theorem:

Theorem 15. *Let (Q,⊕) be a quasigroup and L_Q its Cayley table. Then $\theta : Q \mapsto Q$ is a complete mapping iff we can locate a transversal of L_Q by selecting, in each row x, the entry in column $\theta(x)$. Similarly, $\eta : Q \mapsto Q$ is an orthomorphism iff we can locate a transversal of L_Q by selecting, in each row x, the entry containing symbol $\eta(x)$.*

By theorem 15 we see that each transversal in the Cayley table of a group corresponds to a specific complete mapping. So to find the complete mappings of our groups we can instead find transversals in the Cayley table of our group. This Cayley table can be interpreted as a latin square because each row and column will contain every element of the group exactly once.

The notion of a k-near permutation is strongly related to that of a partial-transversal. Wanless, in [47], defines a *partial transversal* of *length* k of a latin square of order n to be a set of k entries each in different rows and columns and each containing a different symbol. Wanless notes that some papers define a partial transversal to be a set of n entries, each from different rows and columns, that contain k distinct elements. The latter definition adheres more closely to our notion of a k-near permutation and is the definition we will use.

Definition 16 (k-Transversals)**.** A k-transversal of a Latin square L of order n, where $1 \leq k \leq n$, is a list of n entries of L such that no two entries are in the same row, no two entries are in the same column, and there are k distinct symbols in the list.

5.3. **Results in $\mathbb{Z}_n$.**

Lemma 17. *If f is a permutation over $\mathbb{Z}_n$ that is the sum of the identity with another permutation then f has a fixed point.*

Proof. The identity maps every element to itself. Because the function we are adding to the identity is a permutation, it must be true that the additive identity will be added to some element in the identity permutation. For this element, the identity mapping will remain unchanged and yield a fixed point. □

Theorem 18. *Let f be a function defined over $\mathbb{Z}_n$ be the sum of the identity permutation with another permutation. Then $|range(f^{(n-1)})| \neq 2$.*

Proof. For a function to have a terminal range of 2 it is required that the total number of elements that are a part of some cycle is 2. This can manifest in two ways: two 1-cycles or one 2-cycle.

Case 1: Two 1-Cycles

Suppose f does contain two 1-cycles. A 1-cycle is a fixed point in our function. If we have two 1-cycles then there are two elements, i and j, such that $i \mapsto i$ and $j \mapsto j$. If f is the sum of the identity with some other permutation then when we subtract the identity we should be left with a permutation. However, if we subtract the identity from a function that maps i to itself and j to itself, the result is a function that maps two elements to 0. This is a contradiction.

Case 2: One 2-Cycle

Suppose the only cycle in f is a 2-cycle. This is a contradiction of lemma 17. □

Theorem 19. *Let $a = (a_1, \cdots, a_n)$ and $b = (b_1, \cdots, b_n)$ be permutations in one-line notation. Then*

(a) for even n, there exists a pair, a and b, such that $a+b$ is a 1-near permutation.

(b) for even n, there does not exist any a and b such that $a+b$ is a permutation.

(c) for odd n, there exists two a and b such that $a+b$ is a permutation.

(d) for odd n, there does not exist any pair, a and b, such that $a+b$ is 1-near.

Claim a). Let $a = (1, 2, \cdots, n)$ be the identity permutation. Let b be the permutation obtained by cyclically shifting the first $\frac{n}{2}$ elements of a by 1 place, so $b = (2, 3, \cdots, \frac{n}{2}, 1, \frac{n}{2}+1, \frac{n}{2}+2, \cdots, n)$. Then

$$a+b = (1+2, 2+3, 3+4, \cdots, \frac{n}{2}-1+\frac{n}{2}, \frac{n}{2}+1, \frac{n}{2}+1+\frac{n}{2}+1, \frac{n}{2}+2+\frac{n}{2}+2, \cdots, n+n).$$

$$= (3, 5, 7, \cdots, n-1, \frac{n}{2}+1, 2, 4, 6, \cdots, n).$$

Hence 1 does not appear in the one-line notation of the sum $a+b$, but $\frac{n}{2}+1$ appears twice (since $\frac{n}{2}+1$ is somewhere in one of the lists $3, 5, 7, \cdots, n-1$ or $2, 4, 6, \cdots n$). Thus, it is a 1-near permutation.

□

Claim b). See [20]. □

Claim c). Let $a = b = (1, \cdots, n)$, and consider $a+b = (1+1, 2+2, \cdots, n+n)$. The first $\lfloor \frac{n}{2} \rfloor$ elements of the sum will be $2, 4, 6, \cdots, n-1$. Now, $\lceil \frac{n}{2} \rceil + \lceil \frac{n}{2} \rceil$ is 1 mod n, since n is odd. Each subsequent entry is 2 more than the previous entry, so the remaining elements in the one-line notation are the rest of the odd elements from 1 to n: $3, 5, \cdots, n$. Thus $a+b = (2, 4, \cdots, n-1, 1, 3, 5, \cdots, n)$, a permutation.

□

Claim d). Let $\{1, 2, \cdots, n\}/\{t\}$ be the range of $a+b$, where $t \in \{1, 2, \cdots, n\}$. Let $a+b = (c_1, c_2, \cdots, c_n)$, so we have $a_i + b_i = c_i$ where $c_i \in \{1, 2, \cdots, n\}/\{t\}$ and $1 \leq i \leq n$. Now

by pigeon-hole principle, we know that there are exactly two of c_i's that are the same, say $c_k = c_l = h \quad (h \neq t)$, and the rest of c_i's are some arrangement of $\{1, 2, \cdots, n\}/\{t, h\}$. By summing up the quality $a_i + b_i = c_i$ over all i's, we have

$$\sum_{i=1}^{n}(a_i + b_i) \equiv \sum_{i=1}^{n} c_i \quad mod(n),$$

$$\sum_{i=1}^{n} a_i + \sum_{i=1}^{n} b_i \equiv \sum_{i=1}^{n} c_i \quad mod(n),$$

$$\frac{n(n-1)}{2} + \frac{n(n-1)}{2} \equiv \frac{n(n-1)}{2} - t + h \quad mod(n),$$

$$\frac{n(n-1)}{2} \equiv h - t \quad mod(n),$$

Since $h, t \in \{1, 2, \cdots, n\}$, we know that $n \nmid (h-t)$. This implies n must be even; otherwise, the left-hand side is divisible by n and the congruence would not hold. □

Theorem 20. *Let π be a permutation of the cyclic group of n elements, and let f denote $\pi \oplus \theta$. Then if $|Range(f)| = 2$, the one line representation of f is periodic.*

Proof. Let π be a permutation on $\mathbb{Z}_n$, θ the identity permutation, and f the sum of π and θ. Let $e := f(n)$, and the range of f be the set $\{2, e\}$ when $e \neq 2$ and $\{2, x\}$ otherwise.

Further assume that π is in canonical form, and let $\pi(n) := e$. Note that this implies $f(1) = 2$.

$$\begin{array}{lccccc} \theta: & (1 & 2 & 3 & \ldots & n-1 \quad n) \\ \pi: & (1 & & \ldots & & e) \\ \hline f: & (2 & & \ldots & & e) \end{array}$$

It can be seen above that $f(e-1) = \pi(e-1) + e - 1$. Assume $f(e-1) = e$; then $e = \pi(e-1) + e - 1$ which implies $\pi(e-1) = 1$, contradicting the assumption that π is a permutation in canonical form. Thus since f has a range of two elements, $f(e-1)$ must be 2. □

This observation is naturally generalized as follows.

Lemma 21. *If $f(i) = 2$, then $f(i + k(e - 2)) = 2$ for all integers k. Note that for large k this index wraps around f cyclically.*

Proof. The case for $k = 1$ is done above. From there, proceed inductively by applying the same reasoning but with $f(e - 1)$ substituted for $f(1)$, etc. □

Lemma 22. *If $gcd(e - 2, n) = 1$, then f has a range consisting solely of $\{2\}$, violating the assumption that f has size 2. Therefore $e-2$ must not be relatively prime to n for the choice of e to be valid.*

Proof. Let $\langle e-2\rangle$ denote the subgroup of $\mathbb{Z}_n$ generated by $(e-2)$. Lemma 21 can be restated concisely as: for $i \in \langle e - 2\rangle$, $f(i) = 2$. It is a theorem of abstract algebra that any element relatively prime to n generates the cyclic group of n elements. The proposition follows. □

Corollary 23. *For prime n, there is no permutation on Z_n such that $f = \pi \oplus \theta$ satisfies $|range(f)| = 2$.*

Theorem 24. *For every permutation π, The number that range decreases from step i to step $i+1$ is always bigger or equal that decreases from step $i+1$ to step $i+2$, i.e. $|ran((\pi + id)^i)| - |ran((\pi+id)^{(i+1)})| \geq |ran((\pi+id)^{(i+1)})| - |ran((\pi+id)^{(i+2)})|$, for every $0 \leq i \leq n-2$.*

Proof. Notice that the number of elements that are in some cycle doesn't change at all after composing arbitrary times. The change of the size is simply a result that only concerns the elements that are not in some cycle. In particular, every time the number of the size that decreases is equal to the number of the elements in $(\pi + id)^i$ such that no elements in the range of $(\pi + id)$ are mapping to them. Let's say in the i^{th} step (where the function is $(\pi + id)^i$), $a_1, a_2, \cdots, a_t$ are the ones that have no elements in the range of $(\pi + id)$ sending to them (as a consequence, they will vanish after being composed with $(\pi + id)$), and let $b_1, b_2, \cdots, b_l$ be the ones that $a_1, a_2, \cdots, a_t$ send to after being composed with $(\pi + id)$. It's obvious that $l \leq t \leq n$. Hence, next time when we reach the $(i + 2)^{th}$ step by composing $(\pi + id)$, only some of or none of $b_1, b_2, \cdots, b_l$ will in turn become the ones that have no

elements in the range of $(\pi + id)$ sending to them, since all the $a_1, a_2, \cdots, a_t$ have already disappeared from the i^{th} step to $(i+1)^{th}$ step and Thus $b_1, b_2, \cdots, b_l$ will vanish at this time. By the fact that $l \leq t$, we are confirmed that the inequality holds. □

Definition 25. Let $\#(k, s)$ denote the number of permutations π that has size s after composing $\pi + id$ with itself k times, where id is the identity permutation.

Corollary 26. *If n is odd, then $\#(n-2, 2) = \#(n-1, 2) = 0$.*

Proof. It suffices to show that $\#(n-2, 2) = 0$. Suppose not, then the size decreases at least 1 from the step $n-2$ to step $n-1$. By Theorem 24, we know that the size decreases at least 1 from the i^{th} step to $(i+1)^{th}$ step, where $i = 1, 2, \cdots, n-3$. This will imply that the function f with $|range(f^{(n-2)})| = 2$ is initially from a function $(\pi + id)$ that has size at least $2 + 1 \times (n-3) = n-1$. But according to the Theorem 19 d) the range cannot be $n-1$ since n is odd, neither could it be n as in this case $(\pi + id)$ will be a permutation and arbitrary times of composition of permutation result in permutation rather than function of size 2. Thus it leads to a contradiction. □

Theorem 27. *For any n, if $(\pi + id)$ is 1-near permutation, then $|range((\pi + id)^2)| = n-2$ or $n-1$.*

Proof. Suppose x is the repeated element in the one-line notation of $(\pi + id)$.

Case 1: If one of the x's is in the x^{th} place, and the other one, let's say, is in the j^{th} place, then it's clear that in $(\pi + id)^2$, the x^{th} and j^{th} terms are x. Now if j appears in the rest $n-2$ place in $(\pi + id)$, then the j^{th} term of $(\pi + id)^2$ is x. In this case, there are 3 x's in $(\pi + id)^2$ and other terms are all distinct, so the size is $n-2$. On the other hand, if j doesn't appear in the rest of $n-2$ places, the only repeated terms in $(\pi + id)^2$ are just those 2 x's. Thus the size of $(\pi + id)^2$ is $n-1$.

Case 2: If neither of the x's is in the x^{th} place, then suppose they are in i^{th} and j^{th} places, respectively. Since $(\pi + id)$ is 1-near, in the rest of $n-2$ places i and j cannot be both

missing from the one-line notation, otherwise the size is $\leq n-2$. Hence, either i or j is missing, or i and j both exist in the one-line notation.

If both i and j exist, then the places where i and j stay in $(\pi+id)$ will be x's in $(\pi+id)^2$. And the i^{th} and j^{th} in $(\pi+id)^2$ are the same element since their preimages in $(\pi+id)$ are x. So $(\pi+id)^2$ has size $n-2$.

If only one of i and j remains, then the only repeated elements in $(\pi+id)^2$ is the x^{th} terms in $(\pi+id)$. So it's still 1-near. □

Corollary 28. *If* $|ran((\pi+id)^i)|=2$, *then* $|ran((\pi+id)^{(i+1)})|=1$.

Proof. If it's not, then it reaches the terminal size which is 2. By the Theorem 18 it's not possible. □

Theorem 29. $\#(k,1)+\#(k,2)=\#(k+1,1)$ *for* $k\geq\lfloor\frac{n-1}{2}\rfloor$.

Proof. By Corollary 28 we know that every function $(\pi+id)^i$ for certain i and permutation π that reaches size 2 must degenerate to size 1 after one more composition. And those have already reached size 1 will not change anymore in terms of size. Hence it suffices to show that for $k\geq\lfloor\frac{n-1}{2}\rfloor$, the k-step functions $(\pi+id)^k$ with size more than 2 cannot directly degenerate to size 1 after one more composition with $(\pi+id)$, i.e. $|range((\pi+id)^{(k+1)})|\geq 2$ if $|range((\pi+id)^k)|\geq 3$ for $k\geq\lfloor\frac{n-1}{2}\rfloor$.

Now starting with the step y where $y=\lfloor\frac{n-1}{2}\rfloor+1$. Suppose there is some function $(\pi+id)^{(y-1)}$ of size 3 such that $|range((\pi+id)^y)|=1$. The size decreases by 2 from step $y-1$ to step y, which means that during any consecutive two steps before the size decreases at least 2 by Theorem 24. So the y-step function $(\pi+id)^y$ must be initially degenerated from the function $(\pi+id)$ that has size at least

$$(1)\qquad 1+\underbrace{2+2+\cdots+2}_{(y-1)\ terms}=2y-1=2\lfloor\frac{n-1}{2}\rfloor+1=\begin{cases}n-1 & \text{if } n \text{ is even}\\ n & \text{if } n \text{ is odd}\end{cases}$$

And let's call the possible size of $(\pi + id)$ in this case the *initial size* for convenience.

Case 1: When n is even, we deduce from the previous statement that the y-step function $(\pi+id)^y$ might be initially degenerated from a 1-step function has at least size $n-1$. However, by Theorem 27 this is impossible since it only loses size by 1 at the first time of composition rather than 2. Also, the initial size cannot be n because if $(\pi + id)$ is of size n then it's a permutation, and so is $(\pi + id)^y$.

Case 2: When n is odd, the initial size is n, which means that $(\pi + id)$ is a permutation.

Consequently, either case yields contradiction. So there's no such function $(\pi + id)^y$ of size 1 that previously comes from $(\pi + id)^{(y-1)}$ which has size 3. Moreover, we claim that this implies as well the impossibility of degeneration from the function $(\pi + id)^{(y-1)}$ of size more than 3 to the function $(\pi + id)^y$ of size 1. The proof is analogous to (1) in the way we count the initial size. In particular, the resulting initial size $p(y - 1) + 1$, where p is the size that decreased from $(\pi + id)^{(y-1)}$ to $(\pi + id)^y$, will be greater than n hence an obvious contradiction.

So far we have proven that functions $(\pi+id)^y$ of size 1 are degenerated only from functions $(\pi + id)^{(y-1)}$ of size 1 or 2. Corollary 28 tells that all functions of size 2 will only result in size 1's. Thus $\#(k, 1) + \#(k, 2) = \#(k + 1, 1)$ holds for $k = \lfloor \frac{n-1}{2} \rfloor$.

Finally, for any $k > \lfloor \frac{n-1}{2} \rfloor$ the same analysis on the initial size will apply. Hence we finish the proof. □

Theorem 30. *Let $f = \theta + \pi_1$ be a terminally 1-near permutation with $\pi_1(i) = n$ and $\pi_1(j) = \frac{n}{2}$. If $\pi_2(i) = \frac{n}{2}$ and $\pi_2(j) = n$, then $\theta + \pi_2$ is also a terminally 1-near permutation.*

Proof. We now know that the edge labeled n is in a cycle of length one (i.e. the fixed point). The edge labeled $\frac{n}{2}$ is not in a cycle. If we were to exchange the labels of those edges and redirect them accordingly, then their roles would exchange. The non-fixed point would become the fixed point and the fixed point would become the non-fixed point. This preserves the overall structure of the graph. Therefore, it is still a terminally 1-near permutation. □

Corollary 31. *The number of terminally 1-near permutations is even.*

5.4. **All the stuff that had to do with the full function spaces, classes of functions, etc.** Let n be even, $\quad F(n) := \{f | f \textit{ is a function on the cyclic group of order } n\}$

$F_1(n) := \{f | f \in F(n),\ f \textit{ is } 1-\textit{near}\}$

$F_{1t}(n) := \{f | f \in F(n),\ f \textit{ is terminally } 1-\textit{near}\}$

$F_{1t,\pi}(n) := \{f | f \in F(n), f = \pi_1 \oplus \pi_2, \pi_j \in S_n,\ f \textit{ is terminally } 1-\textit{near}\}$

$F_{1,c=0}(n) := \{f | f \in F(n),\ f \textit{ is } 1-\textit{near}, \sum_n f(i) \equiv_n 0\}$

$F_{1t,c=0}(n) := \{f | f \in F(n),\ f \textit{ is terminally } 1-\textit{near}, \sum_n f(i) \equiv_n 0\}$.

That is, $F(n)$ denotes the set of functions on n elements, $F_1(n)$ and $F_{1,t}(n)$ denote two subsets of $F(n)$ of natural interest in our investigation, $F_{1t,\pi}(n)$ those functions in $F_{1,t}(n)$ expressable as the sum of permutations, and the final two sets to be described shortly. In the following discussion the only restriction placed on n is that it is even, so the specification of n will generally be foregone for clarity; e.g. $F_{1t,c=0} \subset F_1$. Since all functions under consideration are 1-near, sometimes we will take x to be the excluded element in the one-line notation of a given f and r to be the repeated element.

The cardinality of F is well known to be n^n. The cardinality of each other set besides $F_{1t,\pi}(n)$ can be calculated directly.

Theorem 32. $|F_1| = \frac{n}{2}(n-1)n!$

Proof. Let $f \in F_1$. Consider the one-line notation of f. Since f is 1-near one must choose two distinguished elements, one to be repeated and one to be excluded. Since these elements are distinct, there are $n(n-1)$ ways to make this choice. The other $n-2$ elements simply comprise the remaining integers up to n. To form the one-line notation, then, the repeated elements may be placed in $\binom{n}{2}$ possible pairs of positions and the remaining elements arranged in $(n-2)!$ configurations. Taking the product yields $n(n-1)\binom{n}{2}(n-2)!$ possible functions f, which simplifies to the result. □

Theorem 33. $|F_{1t}| = (n-1)n!$

Proof. Let $f \in F_{1t}$. Consider the graph representing f. Since f is terminally one near, the graph necessarily contains a subset A of size $|A| = n-1$ on which f acts as a permutation, as well as an excluded element which is mapped into A. There are n choices for the excluded element, $n-1$ choices for its target (if it were a fixed point then it wouldn't be excluded), and $(n-1)!$ configurations for the permutation on A. Taking the product yields $n(n-1)(n-1)!$ possible functions f, which simplifies to the result. □

Theorem 34. *$F_1(n)$ has n-1 equivalence classes determined by $\left(\sum_{i=1}^{n} f(i)\right)$ mod n.*

Proof. $F_1(n)$ is the set of 1-near permutations on n elements. It is known that $\left(\sum_{i=1}^{n} i\right)$ *mod* $n = \frac{n}{2}$. To find the sum of an arbitary 1-near permutation we consider the following sum for $x, y \in \{1, 2, \cdots, n\}$ and $x \neq y$:

$$1 + 2 + 3 + \cdots + n - x + y.$$

The first n elements will sum to $\frac{n}{2}$ *mod* n. So we have

$$\frac{n}{2} - x + y.$$

From the constraints on x and y we have that $1 \leq |-x+y| \leq n-1$. Therefore, we have

$$\frac{n}{2} + 1 \leq \left(\frac{n}{2} - x + y\right) \text{ mod } n \leq \frac{n}{2} + n - 1.$$

This allows for every value on the range [1,n] with the exception of $\frac{n}{2}$. □

Theorem 35. *The n-1 equivalence classes of $F_1(n)$ are determined by $\left(\sum_{i=1}^{n} f(i)\right)$ mod n are the same size.*

Proof. We know that the sum of a 1-near permutation mod n is $\frac{n}{2} - x + y$ for some $x, y \in \{1, 2, \cdots, n\}$ with $x \neq y$. The $n-1$ equivalence classes are determined by the value of this sum. Let $c = \frac{n}{2} - x + y$. Then, rewriting, we see that for a given c, x is determined by y. So, in a particular class, c, there are n choices for an x, y pair that will satisfy the equation.

Each pair will result in a distinct function with a different repeated element. For each of those distinct functions there will be $\frac{n!}{2}$ ways to rearrange the one-line notation and create a new function. Thus each class contains at least $\frac{n}{2}n!$ distinct functions. Then across all classes there are at least $\frac{n}{2}n!(n-1)$ functions. However, this number was already proven to be the total number of functions. Theorem 36 follows. □

Theorem 36. $|F_{1,c=0}| = \frac{n}{2}n!$

Theorem 37. $|F_{1t,c=0}| = n!$

Proof. Let $f \in F_{1t,c=0}$. Consider the graph representing f. Since f is terminally one near, the graph necessarily contains a subset A of size $|A| = n-1$ on which f acts as a permutation, as well as an excluded element x which is mapped into A. Note that the target of x must be the repeated element, since that is the sole vertex of the graph with in-degree greater than one. There are n choices for x, and the constraint that the entries of f must sum to 0 mod n means there is only one valid choice for the repeated element: in particular, the repeated element must be $x + \frac{n}{2}$. Finally, there are $(n-1)!$ configurations for the permutation on A. Taking the product yields $n(n-1)!$ possible functions f, which simplifies to the result. □

Note that $\frac{|F_1|}{|F_{1t}|} = \frac{|F_{1,c=0}|}{|F_{1t,c=0}|} = \frac{n}{2}$, the same ratio we noticed in making the Harmo-Nick Conjecture.

Next we prove the conjecture from a few days ago that the number of terminally 1-near functions expressable as the sum of permutations, a.k.a. $F_{1t,\pi}(n)$, is $n!$.

5.5. **Results in $GF(p^r)$.** Knowing that the number of permutation sums that result in permutations is equivalent to counting the number of transversals in latin squares, we can make use of a result presented in [31].

Corollary 38. *(McKay et al. 2006) Let G be a group of order n. If G is abelian or n is even then the number of transversals in G is congruent to n modulo 2.*

By corollary 38 we know that $GF(p^r)$ for $p > 2$ has at least one transversal.

Consider $M_{n,n}(GF(p^r))$ $(n \geq 2)$ and the group formed by all the invertible matrices in it, namely $GL(n, GF(p^r))$.

The order of the group $GL(n, GF(p^r))$ is given by the formula $|GL| = (q^n - 1)(q^n - q) \cdots (q^n - q^{n-1})$ where $q = p^r$. If the entries of a Latin square are all from this group, we can investigate the the order of this group modulo 3 with the help of the theorem given below to acquire some information about the number of the transversals it can generate.

Theorem 39. *(McKay et al. 2006) If G is a group of order $n \not\equiv 1$ (mod 3) then the number of transversals in G is divisible by 3.*

Claim 1. *The number of the transversals of the Latin square over $GL(n, GF(p^r))$ is always divisible by 3.*

Proof. As we know that every the prime number $p \geq 5$ is of the form $6k \pm 1$ for some $k \in \mathbb{Z}$, $p^r = (6k \pm 1)^r$ is also of the form $6l \pm 1$ for some $l \in \mathbb{Z}$.

Case 1: $p \geq 5$

Write q as $6l \pm 1$. If $q = 6l + 1$, we see that $3 \mid (q^n - q)$ for all $n \in \mathbb{Z}$ thus the order of the group $|GL|$ is a multiple of 3 and apply the theorem we have that the number of transversals is divisible by 3. If $q = 6l - 1$ and n is even, then $3 \mid q^n - 1$; if n is odd, then $3 \mid q^n - q$. In either case above, the order is a multiple of 3 and thus the number of the transversals by theorem is also a multiple of 3.

Case 2 : $p = 3$

It's clear that the number of transversals is divisible by 3 in this case.

Case 3 : $p = 2$

Now $|GL| = (2^{rn} - 1)(2^{rn} - 2) \cdots (2^{rn} - 2^{rn-r})$

$$= 2^{rn(n-1)/2} \prod_{i=0}^{n-1} (2^{r(n-i)} - 1).$$

Since $n \geq 2$, there is some choice of i such that $r(n - i)$ is an even number. With the proper choice, we have $3 \mid 2^{r(n-i)} - 1$, i.e. $3 \mid |GL|$. Thus by the theorem we proved the claim.

In addition, an immediate consequence is that the number of the transversals over the group $SL(n, GF(p^r))$, consisting of all $n \times n$ invertible matrices whose $det = 1$, is also a multiple of 3. This is simply because that $|SL| = |GL|/(q-1) = (q^n-1)(q^n-q)\cdots(q^n-q^{n-2})q^{n-1}$ and a similar argument will apply. □

References

[1] L. Babai, *The probability of generating the symmetric group*, **Journal of Combinatorial Theory** 52 (1989), 148–153.

[2] E. Barkan, E. Biham, *In how many ways can you write Rijndael?*, **Lecture Notes in Computer Science**, Vol. 2501, Springer-Verlag (2002), 160–175.

[3] E. Biham and A. Shamir, *Differential Cryptanalysis of the Data Encryption Standard*, **Springer Verlag**, (1993).

[4] A. Biryukov and D. Khovratovich, *Related-key cryptanalysis of the full AES-192 and AES-256*, **Lecture Notes in Computer Science**, vol. 5912 (2009), 1–18.

[5] A. Biryukov, D. Khovratovich and I. Nikolic, *Distinguisher and related-key attack on the full AES-256*, **Lecture Notes in Computer Science**, Vol. 5677 (2009), 231–249.

[6] A. Biryukov, O. Dunkelman, N. Keller, D. Khovratovich and A. Shamir, *Key Recovery Attacks of Practical Complexity on AES-256 Variants with up to 10 Rounds*, **Lecture Notes in Computer Science**, Vol. 6110 (2010), 299–319.

[7] A. Bogdanov, D. Khovratovich and C. Rechberger, *Biclique cryptanalysis of the full AES*, **Lecture Notes in Computer Science**, Vol. 7073 (2011), 344–371.

[8] D.K. Branstead, J. Gait, S. Katzke, *Report of the Workshop on Cryptography in Support of Computer-Security*, **National Bureau of Standards**, (1977) NBSIR 77-1291.

[9] A. Caranti, F. Dalla Volta, M. Sala and F. Villani, *Imprimitive permutation groups generated by the round functions of key-alternating block ciphers and truncated differential cryptanalysis*, **Computing Research Repository - CoRR** , Vol. abs/math/0, (2006): see arXiv:math.GR/0606022 v2 12 Jun 2006.

[10] C. Cid, S. Murphy, and M.J.B. Robshaw, *Small scale variants of the AES*, **Proceedings of Fast Software Encryption**, Vol. 3557, (2005), 145–162.

[11] C. Cid, S. Murphy, and M.J.B. Robshaw, *Algebraic Aspects of the Advanced Encryption Standard*, **Springer**, New York, (2006).

[12] K. W. Campbell and M.J. Wiener, *DES is not a Group*, **Crypto 92**, 512–520.

[13] D. Coppersmith and E. Grossman, *Generators for Certain Alternating Groups with Applications to Cryptography*, **SIAM Journal on Applied Mathematics** Vol.29 (1975), 624–627.

[14] N. Courtois and J. Pieprzyk, *Cryptanalysis of Block Ciphers with Overdefined Systems of Equations*, **Lecture Notes in Computer Science**, Vol. 2501, Springer-Verlag (2001), 267–287.

[15] L. Cummings, M. Mays, *On the Parity of the Witt Formula*, **Congressus Numerantium Vol. 80** (1992), 49-56.

[16] J. Daemen, and V. Rijmen, *AES Proposal: Rijndael*, **NIST AES Proposal**, (1998).

[17] J. Daemen, and V. Rijmen, *The Design of Rijndael*, **Springer-Verlag**, Berlin, (2002).

[18] O. Dunkelman, N. Keller and A. Shamir, *Improved Single-Key Attacks on 8-Round AES-192 and AES-256)*, **Lecture Notes in Computer Science**, Vol. 6477, Springer-Verlag (2010), 158–176.

[19] J.D. Dixon, *The probability of generating the symmetric group*, **Mathematics Zeitschrift** Vol. 110 Issue 3 (1969), 199–205.

[20] L. Euler, Recherches sur une nouvelle espéce de quarrés magiques *Verh. Zeeuwsch. Gennot. Weten. Vliss.*, 9 (1782), pp. 85-239 Enestrm E530, **Opera Omnia OI7**, 291-392.

[21] N. Ferguson, J. Kelsey, S. Lucks, B. Schneier, M. Stay, D. Wagner, and D. Whiting, *Improved cryptanalysis of Rijndael*, **Lecture Notes in Computer Science**, Vol. 1978, Springer-Verlag (2000), 213–230.

[22] H. Gilbert and M. Minier, *A Collision Attack on 7 Rounds of Rijndael*, **In AES Candidate Conference** (2000), 230–241.

[23] J. A. Gallian, *Contemporary Abstract Algebra*, **Huston Mifflan Company**, (1992).

[24] H. Gilbert and T. Peyrin, *Super-Sbox Cryptanalysis: Improved Attacks for AES-Like Permutations*, **Lecture Notes in Computer Science**, Vol. 6147 (2010), 365–383.

[25] G. Hornauer, W. Stephan and R. Wernsdorf, *Markov ciphers and alternating groups*, **Lecture Notes in Computer Science**, Vol. 765 (1994), 453–460.

[26] K. Ireland and M. Rosen, *A classical introduction to modern Number Theory*, **Springer-Verlag Graduate Texts in Mathematics** 84 (Second Edition), 1990.

[27] B.S. Kaliski, R.L. Rivest, and A.T. Sherman, *Is the Data Encryption Standard a Group? (Results of Cycling Experiments on DES)*, **Journal of Cryptology**, Vol. 1 (1988), 3–36.

[28] T. Van Le, R. Sparr, R. Wernsdorf, and Y. Desmedt, *Complementation-like and cyclic properties of AES round functions*, **Proceedings of the 4th International Conference on the Advanced Encryption Standard**, Vol. 3373 (2005), 128-141.

[29] W. Mao, *Modern Cryptography: Theory and Practice*, **Prentice Hall**, (2003).

[30] S. Mattarei, *Inverse-closed additive subgroups of fields*, **Israel Journal of Mathematics** Vol. 159 (2007), 343–348.

[31] B.D. McKay, J.C. McLeod and I.M. Wanless, The number of transversals in a latin square, *Des. Codes Cryptogr.* **40** (2006), 269-284.

[32] L. Miller,*Generators of the Symmetric and Alternating Group*, **The American Mathematical Monthly**, Vol. 48, (1941), 43 – 44.

[33] S. Murphy, K.G. Paterson, P. Wild, *A weak cipher that generates the symmetric group*, **Journal of Cryptology** 7 (1994), 61–65.

[34] S. Murphy, M.J.B. Robshaw, *Essential algebraic structure within the AES*, **Proceedings of CRYPTO 2002** Vol. 2442 (2002), 1–16.

[35] National Institute of Standards and Technology (US), *Advanced Encryption Standard (AES)*, **FIPS Publication 197**, (2001).

[36] National Institute of Standards and Technology (US), *Recommendation for the Triple Data Encryption Algorithm (TDEA) Block Cipher*, **Special Publication 800-67** (2004).

[37] K.G. Paterson, *Imprimitive permutation groups and trapdoors in iterated block ciphers*, **Lecture Notes in Computer Science**, Vol. 1636 (1999), 201– 214.

[38] S. Patel, Z. Ramzan, G. S. Sundaram, *Luby-Rackof Ciphers: Why XOR Is Not So Exclusive*, **Lecture Notes in Computer Science**, Vol. 2595 (2003), 271–290.

[39] D. M. Rodgers, *Generating and Covering the Alternating or Symmetric group*, **Communications in Algebra**, 30 (2002), 425–435.

[40] P. Rogaway and T. Shrimpton, *Cryptographic Hash-Function Basics: Definitions, Implications, and Separations for Preimage Resistance, Second-Preimage Resistance, and Collision Resistance*, Fast Software Encryption, **Lecture Notes in Computer Science** , Vol. 3017 (2004), 371–388.

[41] Martin Schlaffer, 2011. *Cryptanalysis of AES-Based Hash Functions*. PhD Thesis, Graz University of Technology, Austria.

[42] C. E. Shannon, *A Mathematical Theory of Communication*, **Bell System Technical Journal**, 27 (1948), 379–423.

[43] R. Sparr and R. Wernsdorf, *Group theoretic properties of Rijndael-like ciphers*, **Discrete Applied Mathematics**, Vol. 156 (2008), 3139–3149.

[44] W. Trappe and L. C. Washington, *Introduction to Cryptography with Coding Theory*, **Pearson Education**, (2006).

[45] R. Wernsdorf, *The round functions of Rijndael generate the alternating group*, **Lecture Notes in Computer Science**, Vol. 2365, Springer-Verlag (2002), 143–148.

[46] A. Williamson, *On Primitive Permutation Groups Containing a Cycle*, **Mathematische Zeitschrift**, 130 (1973), 159–162.

[47] I. M. Wanless, *Transversals in Latin squares: A survey*, Surveys in Combinatorics 2011, **London Math. Soc. Lecture Note Series** 392, Cambridge University Press, (2011) 403–437.