Ultrametric integral cryptanalysis
Lecture notes

Tim Beyne

Preface

These are the notes for a series of lectures given at Ruhr University Bochum between July and September 2024. The goal of these lectures was to develop ultrametric integral cryptanalysis over products of finite fields with the same characteristic. A preliminary version of this theory appeared in [**1**, Chapter 5]. The special case of $\mathbb{F}_2^n$ was treated in [**3**].

The audience for these notes are researchers already familiar with linear and integral cryptanalysis, at least for primitives defined on $\mathbb{F}_2^n$. This background is not necessary to understand most of the material, but the main motivation for ultrametric integral cryptanalysis may be unclear without it. Most of the mathematical background (mainly p-adic numbers) is developed as necessary, with the exception of results from linear algebra.

The notes are organized in three chapters: (1) a general discussion of cryptanalysis on commutative inverse monoids, which includes linear and ultrametric integral cryptanalysis as special cases, (2) the main principles of ultrametric integral cryptanalysis on products of fields with the same characteristic and (3) a few examples of applications, primarily a 'cryptanalytic' proof of the Ax-Katz theorem and some analysis of generic Feistel ciphers.

Tim Beyne
April 2025

Contents

LECTURE 1

Cryptanalysis on commutative inverse monoids

1. Introduction

Ultrametric integral cryptanalysis is in some sense a generalization of the algebraic approach to integral cryptanalysis, but a proper motivation requires at least some familiarity with the geometric approach to cryptanalysis [**1**]. The goal of this lecture is to introduce the basic principles of this approach, and to identify the right mathematical setting to develop cryptanalytic theories that are 'as nice as' linear cryptanalysis. The prerequisites for this lecture are linear algebra and linear cryptanalysis, which will be used as a running example.

2. Geometric approach to cryptanalysis

We consider finite sets and functions $F : X \to Y$, such as cryptographic primitives, between them. The geometric approach replaces such functions by linear operators on vector spaces. In particular, it does so from two dual points of view:

2.1. Covariant picture. Let k be a field. The free vector space on X is the set of all formal k-linear combinations of elements of X:

$$k[X] = \mathrm{Span}\{\delta_x \mid x \in X\}\,,$$

where δ_x is the basis vector corresponding to the element x in X. In principle, $k[X]$ comes with an additional coalgebra structure, but this is outside of the scope of this lecture. Intuitively, a vector in $k[X]$ is an assignment of weights to each of the elements of X. In other words, it is a device to keep track of the state. A function $\mathsf{F} : X \to Y$ corresponds to a linear map $T^{\mathsf{F}} : k[X] \to k[Y]$ defined by

$$T^{\mathsf{F}}\,\delta_x = \delta_{\mathsf{F}(x)}\,.$$

The operator T^{F} describes how the state changes when a function F is applied. It will be called the *pushforward* operator of F.

2.2. Contravariant picture. The set of functions from X to k forms a vector space k^X under pointwise addition of functions. The vector space k^X has a basis consisting of the indicator functions δ^x of singleton sets $\{x\}$:

$$k^X = \mathrm{Span}\{\delta^x \mid x \in X\}\,.$$

Again, k^X in principle comes with a natural algebra structure, but this is outside of our current scope. The interpretation of functions in k^X is that they correspond to 'probes' of the state. This is related to the fact that functions in k^X correspond one-to-one to linear functions on $k[X]$, as we can set

$$v\left(\sum_{x\in X} u_x\,\delta_x\right) = \sum_{x\in X} u_x\,v(x)\,.$$

The fact that observations are assumed to be linear is a limitation, but there are compelling reasons to adopt this restriction and this point of view is powerful enough to describe all known general cryptanalytic methods. A more profound motivation can be given using category-theoretic language, but this will be avoided here. For a function $\mathsf{F} : X \to Y$, there is again a corresponding linear map $T^{\mathsf{F}^\vee} : k^Y \to k^X$, now defined by

$$T^{\mathsf{F}^\vee}\,\delta^x = \delta^x \circ \mathsf{F}\,.$$

Extending this linearly, we see that the *pullback* operator $T^{\mathsf{F}^\vee}$ corresponds to precomposition by F.

2.3. Duality. The dual of a vector space V is the space of linear functions from V to the base field k. Hence, k^X is naturally isomorphic to the dual vector space of $k[X]$. The basis vectors δ_x and δ^x are dual. Moreover, $T^{\mathsf{F}^\vee}$ is the adjoint of T^{F}:

$$\left(T^{\mathsf{F}^\vee} v\right)(u) = v\left(T^{\mathsf{F}} u\right).$$

Despite the similarity between $k[X]$ and k^X, it will be essential to keep them separate. For linear cryptanalysis, the distinction can mostly be ignored — but not for ultrametric integral cryptanalysis.

2.4. Cryptanalytic properties. A cryptanalytic property of F is a pair (U, V) with U a subspace of $k[X]$ and V a subspace of k^Y. The evaluation of (U, V) at u in U and v in V is defined as

$$v\left(T^{\mathsf{F}} u\right).$$

The combinatorial quantities of interest in cryptanalysis, such as correlations of linear approximations, can be expressed in this way. It is rarely feasible to evaluate properties exactly, so we need techniques to estimate them up to a 'small' error. For this to be meaningful, the field k should have a metric structure. This point is crucial to motivate ultrametric integral cryptanalysis, and we will come back to it at the end of this lecture.

Example 1. Let $\delta_A = \sum_{x \in A} \delta_x$ be the indicator of a subset $A \subset X$ and let δ^B be the indicator function of a subset $B \subset X$, then

$$\delta^B\left(T^{\mathsf{F}} \delta_A\right) = \left|\{x \in A \mid \mathsf{F}(x) \in B\}\right|.$$

The data of all evaluations of a cryptanalytic property (U, V) are equivalent to a linear map $U \otimes V \to k$ or equivalently

$$\begin{aligned} U &\to k[X]/V^0 \\ x &\mapsto T^{\mathsf{F}} x \pmod{V^0}, \end{aligned}$$

where $V^0 = \{x \in k[Y] \mid v(x) = 0 \text{ for all } v \in V\}$ is the annihilator of V. Equivalently, this is the solution set of the linear system described by V. This can also be expressed dually as the linear map

$$\begin{aligned} V &\to k^X/U^0 \\ x &\mapsto T^{\mathsf{F}^\vee} x \pmod{U^0}, \end{aligned}$$

where $U^0 = \{x \in k^X \mid x(u) = 0 \text{ for all } u \in U\}$. There is a more general definition of properties — not discussed here — as a pair of maps that simplifies these equivalent descriptions at the cost of more abstraction.

2.5. Propagation. We mention two basic but important results. Firstly,

$$T^{\mathsf{F}_r \circ \cdots \circ \mathsf{F}_2 \circ \mathsf{F}_1} = T^{\mathsf{F}_r} \cdots T^{\mathsf{F}_2} T^{\mathsf{F}_1}.$$

The dual version of this is that

$$T^{\mathsf{F}_r \circ \cdots \circ \mathsf{F}_2 \circ \mathsf{F}_1{}^\vee} = T^{\mathsf{F}_1^\vee} T^{\mathsf{F}_2^\vee} \cdots T^{\mathsf{F}_r^\vee}.$$

Secondly, we can define a tensor product of $k[X]$ and $k[Y]$ as $k[X \times Y]$, together with the bilinear map $\otimes : k[X] \times k[Y] \to k[X \times Y]$ defined by $\delta_x \otimes \delta_y = \delta_{(x,y)}$. Similarly, we define a tensor product of linear maps on $k[X]$ and $k[Y]$ that is compatible in the sense that if $A : k[X_1] \to k[X_2]$ and $B : k[Y_1] \to k[Y_2]$, then $A \otimes B$ is the

map $k[X_1 \times Y_1] \to k[X_2 \times Y_2]$ such that $\big(A \otimes B\big)(\delta_x \otimes \delta_y) = A(\delta_x) \otimes B(\delta_y)$. If $\mathsf{F}(x) = (\mathsf{F}_1(x_1), \dots, \mathsf{F}_n(x_n))$, then

$$T^{\mathsf{F}} = T^{\mathsf{F}_1} \otimes \cdots \otimes T^{\mathsf{F}_n} .$$

The prototypical example is a layer of parallel S-boxes. Similar definitions can be given for $k^X \otimes k^Y$.

2.6. Trails. Trails are the main tool to analyze compositions $\mathsf{F} = \mathsf{F}_r \circ \cdots \circ \mathsf{F}_1$ of functions $\mathsf{F}_i : X_i \to X_{i+1}$. The basic idea is that we want to lift properties of $\mathsf{F}_1, \dots, \mathsf{F}_r$ to properties of their composition. To do this in a sound way, we should have a decomposition of $k[X_i]$ as

$$k[X_i] = \bigoplus_{U \in \Omega_i} U .$$

Let U^{c} be the algebraic complement of U in Ω_i with respect to this decomposition. This is equivalent to the following decomposition for k^{X_i}:

$$k^{X_i} = \bigoplus_{U \in \Omega_i} (U^{\mathsf{c}})^0 .$$

In particular, for $(V^{\mathsf{c}})^0$ in Ω_i, we have $V \cong k[X_i]/(V^{\mathsf{c}})^0$ and the property $(U, (V^{\mathsf{c}})^0)$ corresponds a linear map $U \to V$ given by $\pi_V T^{\mathsf{F}} \iota_U$, where ι_U is inclusion and π_V is projection on V with complement V^{c}. The map $\langle V, U \rangle_{\mathsf{F}} = \pi_V T^{\mathsf{F}} \iota_U$ will be called an approximation (map) of F. The approximation map $\langle V, U \rangle_{\mathsf{F}} : U \to V$ is equivalent to the data of all evaluations of the property $(U, (V^{\mathsf{c}})^0)$

THEOREM 1. *For i in $\{1, \dots, r+1\}$, let Ω_i be a set of subspaces of $k[X_i]$ such that $k[X_i] = \bigoplus_{U \in \Omega_i} U$. Every approximation map $\langle U_{r+1}, U_1 \rangle_{\mathsf{F}}$ of $\mathsf{F} = \mathsf{F}_r \circ \cdots \circ \mathsf{F}_1$ with U_1 in Ω_1 and U_{r+1} in Ω_{r+1} satisfies*

$$\langle U_{r+1}, U_1 \rangle_{\mathsf{F}} = \sum_{U_2, \dots, U_r} \langle U_{r+1}, U_r \rangle_{\mathsf{F}_r} \cdots \langle U_3, U_2 \rangle_{\mathsf{F}_2} \langle U_2, U_1 \rangle_{\mathsf{F}_1} ,$$

where the sum is over all $(U_2, \dots, U_r)$ in $\prod_{i=2}^{r} \Omega_i$.

A sequence $(U_1, \dots, U_{r+1})$ of intermediate subspaces, including the endpoints, will be called a trail. This is equivalent to a sequence of compatible approximations. The map of a trail is the composition of approximations

$$\langle U_{r+1}, U_r \rangle_{\mathsf{F}_r} \cdots \langle U_3, U_2 \rangle_{\mathsf{F}_2} \langle U_2, U_1 \rangle_{\mathsf{F}_1} .$$

In these lectures, we will mostly consider the one-dimensional version of trails. In this case, the spaces $U_1, \dots, U_{r+1}$ are spanned by a single vector. Hence, the decomposition of $k[X_i]$ corresponds to a choice of basis:

$$k[X_i] = \bigoplus_{\beta \in B_i} \mathrm{Span}\{b_\beta\} ,$$

where B_i is a set of indices or labels for the basis functions. The corresponding decomposition of k^{X_i} then corresponds to the dual basis:

$$k^{X_i} = \bigoplus_{\beta \in B_i} \mathrm{Span}\{b^\beta\} ,$$

where $b^\beta(b_\beta) = 1$. The approximation map $\mathrm{Span}\{b_{\beta_i}\} \to \mathrm{Span}\{b_{\beta_{i+1}}\}$ is given by

$$x \mapsto b^{\beta_{i+1}}\big(T^{\mathsf{F}_i} b_{\beta_i}\big)\, x .$$

If we define B^{F_i} as the change-of-basis of T^{F} relative to the bases $\{b_\beta \mid \beta \in B_i\}$ and $\{b_\beta \mid \beta \in B_{i+1}\}$, then

$$B^{\mathsf{F}_i}_{\beta_{i+1},\beta_i} = \delta^{\beta_{i+1}}\left(B^{\mathsf{F}_i}\delta_{\beta_i}\right) = b^{\beta_{i+1}}\left(T^{\mathsf{F}_i}b_{\beta_i}\right),$$

where we index the coordinates of B^{F_i} by elements from the label sets B_i and B_{i+1}. The above theorem then translates to

$$B^{\mathsf{F}}_{\beta_{r+1},\beta_1} = \sum_{\beta_2,\ldots,\beta_r}\prod_{i=1}^{r} B^{\mathsf{F}_i}_{\beta_{i+1},\beta_i}.$$

Sequences $(\beta_1,\ldots,\beta_{r+1})$ can be reasonably called trails, as they are in one-to-one correspondence to trails as defined above. In this one-dimensional setting, a natural question is how to choose the bases for $k[X_1],\ldots,k[X_{r+1}]$. The numbers $B^{\mathsf{F}_i}_{\beta_{i+1},\beta_i}$ will be called correlations. It is worth mentioning that the decomposition into one-dimensional trails is equivalent to the simple equality

$$B^{\mathsf{F}} = B^{\mathsf{F}_r}\cdots B^{\mathsf{F}_2}B^{\mathsf{F}_1}.$$

The general 'sum of trails' principle can be interpreted as multiplication of matrices with a predetermined block structure.

Example 2. In linear cryptanalysis, B^{F} is the correlation matrix C^{F}. In the case of linear cryptanalysis over $\mathbb{F}_2^n$, $B_i = \mathbb{F}_2^n$ and

$$b_u = \frac{1}{2^n}\sum_{x\in\mathbb{F}_2^n}(-1)^{u\bullet x}\,\delta_x.$$

Furthermore, $b^u(x) = (-1)^{u\bullet x}$. Here, $u\bullet x = \sum_{i=1}^n u_i x_i$ denotes the dot product between $u = (u_1,\ldots,u_n)$ and $x = (x_1,\ldots,x_n)$.

3. Geometric approach with partial symmetries

The geometric approach may be applied to develop a theory of cryptanalysis for functions between sets with partial symmetries. This setting can be formalized by considering finite sets X with an action of a finite monoid M. This leads to the following setup, where k is a field:

- The free vector space $k[X]$ on X is a coalgebra *and* a $k[M]$-module.
- The function space k^X on X is an algebra *and* a k^M-comodule.

These two points of view are again dual. In general, as long as $k[X]$ is a semisimple $k[M]$-module, this setup leads to a way to choose trails (in the basis-free sense) that behave well with respect to the action of M and under morphisms of M-sets. If $k[X]$ is only completely decomposable, then the situation is more complicated. Here, we restrict this general setting in two important ways:

- It will be assumed that $X = M$, and M acts by multiplication: $x \mapsto m\,x$.
- The setup should lead to a unique theory of one-dimensional trails.

The prototypical example of such a theory is linear cryptanalysis. Hence, we essentially want to identify those cases that lead to something with the same properties as linear cryptanalysis.

3.1. Simultaneous diagonalization. The action of an element m of X corresponds to the function $x \mapsto m \cdot x$, the pushforward of which will be denoted by T^m. Equivalently, $k[X]$ acts on itself by

$$\delta_m \cdot u = T^m u \,,$$

and this can be extended to all of $k[X]$ by linearity. This turns $k[X]$ into an algebra, but we will not discuss this observation here. In 'nice' cryptanalytic theories such as linear cryptanalysis, the existence of an algebra or more generally module structure on $k[X]$ provides a unique decomposition of $k[X]$ into one-dimensional vector spaces that maximally 'simplifies' the description of the action of X on itself. That is, there exists a joint eigenvector basis of the linear maps T^m for all m in X. This is only possible if X is commutative, since if v is a joint eigenvector then

$$T^{m_1} T^{m_2} v = \lambda_{m_1} \lambda_{m_2} v = \lambda_{m_2} \lambda_{m_1} v = T^{m_2} T^{m_1} v \,.$$

For a finite group, commutativity is enough if k contains enough roots of unity. The situation for monoids is slightly more complicated. As shown below, it is necessary and sufficient that X is a so-called *inverse monoid*.

LEMMA 1. *Commuting linear endomorphisms are simultaneously diagonalizeable if and only if they are individually diagonalizeable.*

PROOF. Let $T_1, T_2, \ldots, T_n$ be commuting endomorphisms on a vector space V. We prove that if they are diagonalizeable, then they are simultaneously diagonalizeable. The other direction is clear. The eigenspaces V_λ of T_1 yield the following decomposition of V:

$$V = \bigoplus_\lambda V_\lambda \,.$$

Since $T_1 T_i x = T_i T_1 x = \lambda T_i x$ for all x in V_λ, each of the spaces V_λ is stable under T_i. Hence, it is sufficient to prove that the restrictions of $T_2, \ldots, T_n$ to V_λ are simultaneously diagonalizeable. Repeatedly applying this argument shows that $T_1, T_2, \ldots, T_n$ are simultaneously diagonalizeable. □

A commutative monoid X is called inverse if for all x there exists a y such that $x^2 y = x$ and $y^2 x = y$. The following result show that this is a necessary and sufficient condition for all the matrices T^m to be diagonalizeable.

THEOREM 2. *Let X be a finite monoid and k an algebraically closed field of characteristic zero. The endomorphisms T^m with m in X are simultaneously diagonalizeable if and only if X is a commutative inverse monoid.*

PROOF. Let us start by assuming that X is a commutative inverse monoid. By the results above, it is sufficient to prove that T^m is diagonalizeable for all m in X. Since k is algebraically closed, T^m is diagonalizeable if and only if its minimal polynomial has distinct roots. However, X is finite, so there exist a and $b \geq a$ such that $m^b = m^a$. This implies that the minimal polynomial of T^m divides $x^a(x^{b-a} - 1)$. In particular, if T^m is not diagonalizeable, then this must be because the geometric multiplicity of eigenvalue zero is less than its algebraic multiplicity. By the 'lemme des noyaux' or the existence of the Jordan normal form, there exists a subspace V such that $(T^m)^n V = 0$ for some $n \geq 1$ and T^m is diagonalizeable on a complement of V. However, X is inverse, so there exists an m^* such that

$$T^m = \left(T^m\right)^2 T^{m^*} \,.$$

However, this implies $\ker T^m = \ker \big(T^m\big)^2$. Iterating this argument for $m^2, m^3, \ldots$ shows that $\ker T^m = \ker (T^m)^n$ and hence $T^m V = 0$. This means that T^m is diagonalizeable on V. Hence, T^m is diagonalizeable.

For the converse claim, assume that the matrices T^m are simultaneously diagonalizeable. It was already shown above that X must be commutative. As already mentioned, the eigenvalues of T^m are necessarily zero or a root of unity. It follows that, by taking a sufficiently high power of T^m, we get a matrix $(T^m)^n$ with all eigenvalues equal to zero or one. We can assume $n \geq 1$. In diagonal form, one can see that

$$\big(T^m\big)^n T^m = T^m \,.$$

Equivalently, $m^2\, m^{n-1} = m$. Hence, X is a commutative inverse monoid. □

3.2. Characters. Finally, we show that the joint eigenvectors are uniquely determined by their *character*. This implies that the joint eigenvector basis is unique up to scale.

THEOREM 3. *Let X be a finite commutative inverse monoid and k an algebraically closed field of characteristic zero. There exist precisely $|X|$ distinct monoid homomorphisms $\chi : X \to k$. For every character χ, there exists a unique joint eigenvector b_χ of T^m for all m in X such that*

$$T^m\, b_\chi = \chi(m)\, b_\chi \,,$$

Furthermore, up to scale, the dual basis vectors b^χ are equal to the characters of X.

PROOF. Let $v_1, \ldots, v_{|X|}$ be a joint eigenvector basis. One can see that for each i there exists a monoid homomorphism $\chi_i : X \to k$ such that

$$T^m\, v_i = \chi_i(m)\, v_i \,.$$

The fact that χ_i is a homomorphism follows from the multiplicativity of the matrices T^m. By linear algebra, there exists a dual basis $v^1, \ldots, v^{|X|}$ for k^X and the dual basis vectors are joint eigenvectors of ${T^m}^\vee$:

$${T^m}^\vee v^i = \sum_j v^i\big(T^m\, v_j\big)\, v^j = \chi_i(m)\, v^i \,.$$

Hence, $v^i\big(mx\big) = \chi_i(m) v^i(x)$ for all x and m in X. Since v^i is not the zero vector, this implies $v^i(1) \neq 0$. Hence, $x \mapsto v^i(x)/v^i(1)$ is a homomorphism of monoids. This shows that there are at least $|X|$ linearly independent monoid homomorphisms $X \to k$, and they are in one-to-one correspondence with the joint eigenvectors of T^m up to scaling.

To show that there are exactly $|X|$ characters, it is enough to prove that characters are linearly independent. Let $\chi_1, \ldots, \chi_n$ be distinct characters. We prove the result by induction on n, with $n = 1$ following from the fact that characters are nonzero. Let $c_1, \ldots, c_n$ be arbitrary scalars such that for all x in X,

$$\sum_{i=1}^n c_i\, \chi_i(x) = 0 \,.$$

Without loss of generality, let us assume that $c_1 \neq 0$. For j in $\{1, \ldots, n\}$, there exists an m in X such that $\chi_1(m) \neq \chi_j(m)$. Hence,

$$0 = \sum_{i=1}^n c_i \left(\chi_i(mx) - \chi_1(m)\chi_i(x)\right) = \sum_{i=2}^n c_i \left(\chi_i(m) - \chi_1(m)\right)\chi_i(x) \,.$$

By the induction hypothesis, any $n-1$ characters are linearly independent. Hence, $c_j\big(\chi_j(m) - \chi_1(m)\big) = 0$. However, $\chi_j(m) \neq \chi_1(m)$ for some m so $c_j = 0$. □

4. Structure of commutative inverse monoids

It is worth investigating the structure of commutative inverse monoids in more detail. In particular, it turns out that commutative inverse monoids can be understood as a kind of meet-semilattice of Abelian groups.

4.1. Idempotents. Above, we already implicitly used the following lemma.

LEMMA 2. *For every x in a finite commutative monoid X, there exists a positive n such that x^n is idempotent.*

PROOF. Since X is finite, there exist non-negative integers a and b such that $x^a = x^{a+b}$. It follows from this that $x^a = x^{a+ab} = x^{a(b+1)}$. Multiplication by $x^{a(b-1)}$ gives $x^{a+a(b-1)} = x^{a(b+1)+a(b-1)}$. It follows that $(x^{ab})^2 = x^{ab}$ so x^{ab} is idempotent. □

The idempotents of a commutative inverse monoid form a meet-semilattice. The partial order on the idempotents also extends to all of X by setting $x \leq y$ if and only if $e \leq f$ and $x = ey$, where e and f are the unique idempotent powers of x and y respectively.

LEMMA 3. *The submonoid $\mathcal{E}_X$ of idempotent elements of a commutative inverse monoid X is a bounded meet-semilattice, with partial order defined by $e \leq f$ if and only if $e = ef$.*

PROOF. The relation $\leq$ defines a partial order relation. Indeed, $e = e^2$, so $e \leq e$. If $e \leq f$ and $f \leq e$, then $e = ef = f$. Finally, if $e \leq f$ and $f \leq g$, then $e = hf = hg$ for some idempotent h. Hence, $e \leq g$. Furthermore, the product of e and f in the meet of e and f (their greatest lower bound). Indeed, suppose that there exists a g such that $g \leq e$, $g \leq f$ and $g \geq ef$, then $g = gf = g(ef) = ef$. Finally, $\mathcal{E}_X$ is bounded, since $e \leq 1$ for all e. □

For an idempotent e in X, we denote by $X_e = (eX)^\times$ the group of units of the monoid eX with unit e. The following result shows that these 'maximal subgroups' form a partition of X.

THEOREM 4. *Every commutative inverse monoid X can be partitioned as*

$$X = \bigsqcup_{e \in \mathcal{E}_X} X_e \,.$$

In particular, $x \in X_e$ if only if e is the unique idempotent power of x.

PROOF. Let $e = x^n$ be the unique idempotent power of x in X. Firstly, note that x is an element of eX. This follows by repeatedly applying the fact that, because X is an inverse monoid, there exists a y such that $x = y\,x^2$. Indeed, after n iterations, one gets $x = y^{n-1}\,e$. Since $x^{n-1}\,x = e$, the element x also has an inverse in eX. That is, $x \in X_e$. It follows that X is equal to the union of the groups X_e. Hence, it suffices to show that these groups are distinct. If x is an element of X_e, then there exists an n such that $x^n = e$. Indeed, X_e is a finite Abelian group so its exponent is finite. The element e is unique, because every group contains exactly one idempotent element. □

Theorem 4 can be taken a step further by noting that the groups corresponding to idempotent elements $e \leq f$ are connected by injective group homomorphisms $X_f \to X_e$ given by $x \mapsto ex$. The product of x and y in X can then be interpreted as follows. First, determine the idempotent powers e and f corresponding to x and y respectively. The product xy is equal to the product of $(ef)\,x = fx$ and $(ef)\,y = ey$ in the group attached to the meet ef of e and f. Every finite bounded meet-semilattice can be obtained as the set of idempotents of a commutative inverse monoid in this way, and conversely.

Example 3 (Groups). If X is a group, then its unit is the only idempotent element. This means that, for x and y in X, the inequality $x \leq y$ is equivalent to $x = y$.

Example 4 (Boolean lattice). Let $X = \mathbb{F}_q^n$ with $\mathbb{F}_q$ a finite field and coordinate-wise multiplication as the binary operation. The semilattice of idempotents $\mathcal{E}_X$ is isomorphic to $\mathbb{F}_2^n$ with bitwise-and or equivalently the semilattice of the subsets of a set on n elements, with set intersection as the meet operation. The groups attached to the elements of this lattice are products of $\mathbb{F}_q^\times$ and the trivial group $\{0\}$ with 0 as its unit. For example, the group corresponding to the idempotent element $(1, 1, 0, \ldots, 0)$ is $\mathbb{F}_q^\times \times \mathbb{F}_q^\times \times \{0\} \times \cdots \times \{0\}$.

4.2. Characters. Theorem 4 implies the Clifford-Munn-Ponizovskiĭ correspondence for the characters of X. This is a special case of a more general result in the representation theory of monoids.

Theorem 5. *Let X be a finite commutative inverse monoid. Every character $\chi : X \to k$ of X is an extension of a group character $\psi : X_e \to k$ with e an idempotent of X, as follows: for all x in X_f with f an idempotent of X, we have*

$$\chi(x) = \begin{cases} \psi(ex) & \text{if } e \leq f\,, \\ 0 & \text{else}\,, \end{cases}$$

Furthermore, different group characters lead to different characters of X.

Proof. The idea of the construction is that χ is a monoid homomorphism on every group X_f, that additionally preserves the meet operation. For x in X_f and y in X_g, note that the idempotent power of xy is the meet fg. Hence,

$$\chi(xy) = \begin{cases} \psi(exy) = \psi(ex)\psi(ey) & \text{if } e \leq fg\,, \\ 0 & \text{else}\,. \end{cases}$$

Since fg is the greatest lower bound on f and g, the condition $e \leq fg$ is equivalent to $e \leq f$ and $e \leq g$. This implies that $\chi(xy) = \chi(x)\chi(y)$. Finally, $\chi(1) = 1$ because $e \leq 1$ for all idempotents e and $\psi(e) = 1$ for all group characters ψ of X_e. □

The characters of X themselves form a commutative inverse monoid $\widehat{X}$ with binary operation defined by pointwise multiplication. For two idempotent characters χ and ψ, we again define $\chi \leq \psi$ if and only if $\chi = \chi\psi$. Suppose χ extends the trivial character of X_e and ψ the trivial character of X_f. The condition $\chi = \chi\psi$ is equivalent to equality between the following two sets (the corresponding supports):

$$\big\{g \in \mathcal{E}_X \mid e \leq g\big\} = \big\{g \in \mathcal{E}_X \mid e \leq g \text{ and } f \leq g\big\}\,.$$

This is the case if and only if $e \geq f$. In other words, $\mathcal{E}_{\widehat{X}}$ is isomorphic to the dual poset of $\mathcal{E}_X$.

4.3. Lattice structure. Due to the duality between X and $\widehat{X}$, the meet-semilattice $\mathcal{E}_X$ can be given the structure of a lattice. Indeed, the join $e \vee f$ of idempotent elements e and f can be defined as the idempotent corresponding to the meet of the trivial characters of the groups X_e and X_f. In fact, it is true in general that every finite bounded meet-semilattice is also a join-semilattice.

5. Fourier transformation

The Fourier transformation on $k[X]$ is the change-of-basis transformation from the standard basis to the joint eigenvector basis $\{b_\chi \mid \chi \in \widehat{X}\}$ from Section 3. To avoid the need to choose an arbitrary identification between X and $\widehat{X}$, we define the Fourier transformation as follows.

Definition 1 (Fourier transformation). Let X be a finite commutative inverse monoid with dual $\widehat{X}$. Let $\{b_\chi \mid \chi \in \widehat{X}\}$ be the dual of the character basis of k^X. The Fourier transformation on $k[X]$ is the linear map $\mathscr{F}_X : k[X] \to k[\widehat{X}]$ with

$$\mathscr{F}_X(b_\chi) = \delta_\chi \,.$$

The inverse-adjoint $\mathscr{F}_X^{-\vee} : k^X \to k^{\widehat{X}}$ of $\mathscr{F}_X$ will occasionally be referred to as the Fourier transformation on k^X.

Theorem 5 allows to give an explicit formula for the characters of any commutative inverse monoid. This in turn leads to an explicit formula for $\widehat{u} = \mathscr{F}_X\, u$:

$$\widehat{u}_\chi = b^\chi(u) = \sum_{x \in X} \chi(x)\, u_x \,.$$

The formula for $\mathscr{F}_X^{-\vee}$ is not as simple, unless X is a group. In that case, the basis vectors b_χ are given by $b_\chi = \sum_{x \in X} \overline{\chi(x)}\, \delta_x / |X|$ and $v(\chi) = v(b_\chi)$ then gives a similar formula. Any formula for b_χ is also equivalent to a 'Fourier inversion formula', since

$$u_x = \sum_{\chi \in \widehat{X}} \delta^x(b_\chi)\, \widehat{u}_\chi \,.$$

5.1. Fourier inversion formula. To obtain a formula for the inverse Fourier transformation, we will invert $\mathscr{F}_X$ on each of the groups X_e using the Fourier inversion formula for groups, and then combine these formulae using Möbius inversion on the poset X. The formula for the Fourier transformation is

$$\widehat{u}_\chi = \sum_{e \in \mathcal{E}_X} \sum_{x \in X_e} \chi(x)\, u_x \,.$$

For y in X, let $\widehat{X}_y$ denote the set of monoid characters extending the group characters of $X_y = X_f$ with f the unique idempotent power of y. Fourier inversion on the group X_y gives

$$\frac{1}{|X_y|} \sum_{\chi \in \widehat{X}_y} \widehat{u}_\chi / \chi(y) = \sum_{e \in \mathcal{E}_X} \sum_{x \in X_e} u_x \, \frac{1}{|X_y|} \sum_{\chi \in \widehat{X}_y} \chi(x)/\chi(y) \,.$$

It holds that $\chi(x) = 0$ unless $e \leq f$. Furthermore, if $e \leq f$, then

$$\frac{1}{|X_y|} \sum_{\chi \in \widehat{X}_y} \chi(x)/\chi(y) = \begin{cases} 1 & \text{if } y = ex \,, \\ 0 & \text{else} \,. \end{cases}$$

By definition $e \leq f$ and $y = ex$ is equivalent to $x \leq y$. Hence, we obtain

$$\frac{1}{|X_y|} \sum_{\chi \in \widehat{X}_y} \widehat{u}_\chi / \chi(y) = \sum_{y \geq x} u_x \,.$$

The remaining challenge is to invert this sum. This can achieved using the Möbius inversion formula. This a generalization of the inclusion-exclusion principle.

THEOREM 6 (Möbius inversion). *Let P be a finite partially ordered set and k a field. There exists a function $\mu : P \times P \to k$ such that if two functions $f : P \to k$ and $g : P \to k$ satisfy*

$$g(x) = \sum_{\substack{y \in P \\ y \geq x}} f(y) \,,$$

then they also satisfy

$$f(x) = \sum_{\substack{y \in P \\ y \geq x}} \mu(x,y) g(y) \,.$$

The function μ is called the Möbius function of P and satisfies the recurrence relation $\mu(x,y) = -\sum_{x \leq z < y} \mu(x,z)$ with $\mu(x,x) = 1$.

Let μ be the Möbius function of X. Applying Theorem 6 yields the following Fourier inversion formula, generalizing the usual formula for groups:

$$u_x = \sum_{y \geq x} \frac{\mu(x,y)}{|X_y|} \sum_{\chi \in \widehat{X}_y} \widehat{u}_\chi / \chi(y) \,.$$

As a special case of this formula, we obtain an expression for $b_\chi = \mathscr{F}_X^{-1} \, \delta_\chi$. If χ is the extension of a character of X_e, then

$$b_\chi = \frac{1}{|X_e|} \sum_{y \in X_e} \frac{1}{\chi(y)} \sum_{x \leq y} \mu(x,y) \, \delta_x \,.$$

5.2. Fourier transformation of linear maps. If $T : k[X] \to k[Y]$ is a linear map, then its Fourier transformation is the linear map

$$\mathscr{F}_Y \, T \, \mathscr{F}_X^{-1} : k[\widehat{X}] \to k[\widehat{Y}] \,.$$

By duality, the definition for linear maps between k^Y and k^X is analogous. For cryptanalysis, the Fourier transformation of the pushforward and pullback operators are of particular importance. Indeed, recall that the evaluation of one-dimensional properties defined by a basis vector and a dual basis vector is

$$b^\psi \big(T^{\mathsf{F}} b_\chi \big) = \delta^\psi \big(\mathscr{F}_Y \; T^{\mathsf{F}} \; \mathscr{F}_X^{-1} \; \delta_\chi \big) = B^{\mathsf{F}}_{\psi,\chi} \,.$$

The properties of B^{F} correspond precisely to the 'propagation rules' of our cryptanalytic theory. Some of these properties have nothing to do with the Fourier transformation. They are the functorial properties mentioned above. In particular, if $\mathsf{F} = \mathsf{F}_r \circ \cdots \circ \mathsf{F}_2 \circ \mathsf{F}_1$, then

$$B^{\mathsf{F}} = B^{\mathsf{F}_r} \cdots B^{\mathsf{F}_2} B^{\mathsf{F}_1} \,.$$

If $\mathsf{F}(x) = (\mathsf{F}_1(x_1), \mathsf{F}_2(x_2), \ldots, \mathsf{F}_n(x_n))$, then

$$B^{\mathsf{F}} = B^{\mathsf{F}_1} \otimes B^{\mathsf{F}_2} \otimes \cdots \otimes B^{\mathsf{F}_n} \,.$$

For the latter expression the tensor product needs to be chosen appropriately. More precisely, one needs to choose an isomorphism between the dual of $X_1 \oplus \cdots \oplus X_n$ and $\widehat{X}_1 \oplus \cdots \oplus \widehat{X}_n$. The following theorem summarizes some properties that are specific to the Fourier transformation. These correspond to 'propagation rules', in particular the standard propagation rules of linear cryptanalysis are a consequence.

THEOREM 7. *Let X and Y be commutative inverse monoids. The Fourier transformation of the pushforward operator has the following properties:*

(1) If $\mathsf{F}(x) = m \cdot x$ with m in X, then

$$B^{\mathsf{F}}_{\chi,\psi} = \begin{cases} \chi(m) & \text{if } \chi = \psi\,, \\ 0 & \text{else}\,. \end{cases}$$

(2) If $\mathsf{F} : X \to Y$ is a monoid homomorphism, then

$$B^{\mathsf{F}}_{\chi,\psi} = \begin{cases} 1 & \text{if } \psi = \chi \circ \mathsf{F}\,, \\ 0 & \text{else}\,. \end{cases}$$

(3) In particular, if $\mathsf{F} : X \oplus X \to X$ with $\mathsf{F}((x,y)) = x \cdot y$, then

$$B^{\mathsf{F}}_{\chi,(\psi_1,\psi_2)} = \begin{cases} 1 & \text{if } \chi = \psi_1 = \psi_2\,, \\ 0 & \text{else}\,. \end{cases}$$

(4) In particular, if $\mathsf{F} : X \to X \oplus X$ with $\mathsf{F}(x) = (x,x)$, then

$$B^{\mathsf{F}}_{(\chi_1,\chi_2),\psi} = \begin{cases} 1 & \text{if } \psi = \chi_1\chi_2\,, \\ 0 & \text{else}\,. \end{cases}$$

PROOF. The first property is by definition. Indeed $T^{\mathsf{F}} = T^m$ and the 'Fourier basis' was constructed precisely to jointly diagonalize these matrices. The eigenvalues were computed in Theorem 3. The last two properties follow from the second. For the second property, note that $\psi \circ \mathsf{F}$ is itself a character, so

$$B^{\mathsf{F}}_{\chi,\psi} = b^{\chi}\big(T^{\mathsf{F}} b_{\psi}\big) = \big(T^{\mathsf{F}^{\vee}}\, b^{\chi}\big)\, b_{\psi} = b^{\chi\circ\mathsf{F}}(b_{\psi})\,.$$

The result follows by duality of the basis vectors. □

For completeness, we mention an explicit formula for the coordinates of B^{F} in the standard basis, though this is rarely useful (one might advise to never use it):

$$B^{\mathsf{F}}_{\chi,\psi} = b^{\chi}\big(T^{\mathsf{F}} b_{\psi}\big) = \frac{1}{|X_e|} \sum_{y \in X_e} \frac{1}{\psi(y)} \sum_{x \le y} \mu(x,y)\chi(\mathsf{F}(x))\,,$$

assuming ψ extends a character of X_e.

6. Multidimensional properties

Let X be a finite commutative inverse monoid. Following the example of multiple linear cryptanalysis, we want to study properties (U, V) of the type

$$U = \operatorname{Span}\Big\{\delta_x \mid x \in M \subset X\Big\}$$
$$V = \operatorname{Span}\Big\{b^{\chi} \mid \chi \in N \subset \widehat{Y}\Big\}\,,$$

where M and N are sets of characters. If M and N are in fact submonoids, then U and V are related to certain partitions of $\widehat{X}$ and Y respectively. This can

be understood without calculations from general principles (coalgebra and algebra structure) that we will not discuss here. Instead, we give a more direct argument based on an isomorphism between M and a quotient of $\widehat{X}$. For the following theorem, we introduce the shorthand notation

$$b_S = \sum_{\chi \in S} b_\chi \,,$$

for a set of characters S.

THEOREM 8. *Let M be a submonoid of X, and $\sim$ an equivalence relation on $\widehat{X}$ defined by $\chi \sim \psi$ if and only if $\chi(x) = \psi(x)$ for all x in M. We have the following equality of subspaces:*

$$\operatorname{Span}\left\{\delta_x \;\middle|\; x \in M\right\}, = \operatorname{Span}\left\{b_{[\chi]} \;\middle|\; [\chi] \in \widehat{X}/\sim\right\}$$

where $[\chi]$ is an equivalence class in $\widehat{X}/\sim$ with representative element χ.

PROOF. The first space is included in the second, because $b^\chi(\delta_x) = b^\psi(\delta_x)$ whenever $\chi \sim \psi$. That is, the first space only contains vectors whose Fourier transformation has coordinates that are constant on the equivalence classes defined by $\sim$. To complete the proof, we show that the spaces are equal by proving that the dimension of the first space is $|\widehat{X}/\sim|$. This follows from the isomorphism induced by the monoid homomorphism

$$\begin{aligned} \widehat{X} &\to \widehat{M} \\ \chi &\mapsto \chi|_M \,, \end{aligned}$$

where $\chi|_M$ is the restriction of M to χ. □

As a first supplement to the above theorem, we have the identity

$$b_{[\chi]} = \sum_{x \in M} \delta^x\big(b_{[\chi]}\big)\, \delta_x = \big|[\chi]\big| \sum_{x \in M} \delta^x\big(b_\chi\big)\, \delta_x \,,$$

where the value of $\delta^x(b_\chi)$ can be computed from the explicit formula for b_χ. There is a similar identity in the reverse direction:

$$\delta_x = \sum_{[\chi] \in \widehat{X}/\sim} b^{[\chi]}\big(\delta_x\big)/|[\chi]|\, b_{[\chi]} = \sum_{[\chi] \in \widehat{X}/\sim} \chi(x)\, b_{[\chi]} \,.$$

For the subspace V, we have the following dual theorem. The proof is based on the same ideas.

THEOREM 9. *Let N be a submonoid of $\widehat{X}$, and $\sim$ an equivalence relation on X defined by $x \sim y$ if and only if $\chi(x) = \chi(y)$ for all χ in N. We have the following equality of subspaces:*

$$\operatorname{Span}\left\{b^\chi \;\middle|\; \chi \in N\right\} = \operatorname{Span}\left\{\delta^{[x]} \;\middle|\; [x] \in X/\sim\right\},$$

where $[x]$ is an equivalence class in $X/\sim$ with representative element x.

PROOF. The first space is included in the second, because it only includes functions that are constant on equivalence classes. Indeed, $b^\chi(\delta_x) = b^\chi(\delta_y)$ whenever $x \sim y$. To show that the spaces are equal, it is now sufficient to prove that the

dimension of the first space is $|X/\sim|$. This follows from the isomorphism induced by the following monoid homomorphism:

$$\begin{aligned} X &\to \widehat{N} \\ x &\mapsto \mathrm{ev}_x\,, \end{aligned}$$

where $\mathrm{ev}_x : N \to k$ is the evaluation character defined by $\mathrm{ev}_x(\chi) = \chi(x)$. The kernel of this map is the congruence relation $\sim$, so $\widehat{N} \cong X/\sim$ by the first isomorphism theorem for monoids. It follows that $|N| = |X/\sim|$. □

Like for Theorem 8, we give two supplements to Theorem 9. Firstly,

$$\delta^{[x]} = \sum_{\chi\in N} \delta^{[x]}\big(b_\chi\big)/|[x]|\, b^\chi = \sum_{\chi\in N} \delta^x(b_\chi)\, b^\chi\,.$$

For the reverse direction, we have

$$b^\chi = \sum_{[x]\in X/\sim} b^\chi\big(\delta_x\big)\, \delta^{[x]} = \sum_{[x]\in X/\sim} |[x]|\, \chi(x)\, \delta^{[x]}\,.$$

There are results dual to Theorem 8 and Theorem 9 for quotient spaces. In this case one obtains an equality between two quotients of $k[X]$ (or k^X) by distinct subspaces. This too follows from general considerations about the Hopf algebra structures on $k[X]$ and k^X.

7. Places of fields

As mentioned in Section 2, cryptanalysis is generally concerned with *estimation* of evaluations of properties, as exact calculations are rarely feasible. This raises the question of what estimation means in this context. This requires the notion of absolute value function. A field with an absolute value is called a valued field.

Definition 2 (Absolute value). Let k be a field. An absolute value on k is a real-valued function $|\cdot| : k \to \mathbb{R}$ on k such that

(1) For all x in k, $|x| \geq 0$ with equality if and only if $x = 0$.
(2) The function $|\cdot|$ is multiplicative: for all x and y in k, $|xy| = |x|\,|y|$.
(3) The triangle-inequality holds: for all x and y in k, $|x + y| \leq |x| + |y|$.

Furthermore, if the strong triangle-inequality $|x+y| \leq \max\{|x|, |y|\}$ holds, then $|\cdot|$ is called a non-Archimidean or ultrametric absolute value. Otherwise, $|\cdot|$ is called Archimidean.

The most obvious example of a nontrivial absolute value function is perhaps the standard absolute value on $\mathbb{Q}$, namely $|x| = x$ if $x \geq 0$ and $|x| = -x$ otherwise. The following example gives infinitely many other important examples.

Example 5. Let p be a prime. Every nonzero rational number x can be written as $x = p^e\,(a/b)$ with a and b integers indivisible by p. Let $|x|_p = p^{-e}$ and $|0|_p = 0$. For example, $|10|_2 = 1/2$. The function $x \mapsto |x|_p$ is a non-Archimidean absolute value on $\mathbb{Q}$. It is called the p-adic absolute value.

If k is a valued field, the decomposition into one-dimensional trails can be turned into the principle of dominant trails, which is familiar from the case of

linear cryptanalysis (with $k = \mathbb{R}$ or $k = \mathbb{C}$):

$$B^{\mathsf{F}}_{\chi_{r+1},\chi_1} = \underbrace{\sum_{\chi\in\Lambda}\prod_{i=1}^{r} B^{\mathsf{F}_i}_{\chi_{i+1},\chi_i}}_{\text{estimate}} + \underbrace{\sum_{\chi\in\Omega\setminus\Lambda}\prod_{i=1}^{r} B^{\mathsf{F}_i}_{\chi_{i+1},\chi_i}}_{\text{error } \varepsilon},$$

where Ω is the set of all trails and Λ the set of 'dominant trails'. The idea is that $|\varepsilon|$ is small, so that the dominant trails provide a good estimate for the evaluation of the property.

If we want to develop a theory similar to linear cryptanalysis, we have to choose an absolute value function. It turns out that if k is an extension of $\mathbb{Q}$, then the number of options is limited up to equivalence. Two absolute values $|\cdot|$ and $|\cdot|'$ are called equivalent if there exists a positive real number t such that for all x in k,

$$|x|' = |x|^t .$$

Equivalently, the metrics defined by $|\cdot|$ and $|\cdot|'$ determine the same topology on k. A place is an equivalence class of absolute value functions. The places of $\mathbb{Q}$ are classified by Ostrowski's theorem, which we state without proof.

THEOREM 10 (Ostrowski). *Up to equivalence, the only nontrivial absolute value functions on $\mathbb{Q}$ are the ordinary absolute value function $|\cdot|$ and, for every prime number p, the p-adic absolute value functions $|\cdot|_p$.*

For extensions of k there are more choices, but when restricted to $\mathbb{Q}$ they must be equivalent to one of the options listed in Theorem 10. In general, for a number field, one can show that the ultrametric absolute value functions correspond to the prime ideals of its ring of algebraic integers. For this reason, the place corresponding to the ordinary absolute value is sometimes called the 'prime at ∞'.

Once an absolute value on k is fixed, it is often more convenient to work in the metric completion of k by this absolute value. In the case of $\mathbb{Q}$, the metric completion with respect to the ordinary absolute value is of course the field of real numbers $\mathbb{R}$. The completion with respect to the p-adic absolute value gives the field of p-adic numbers, which will be discussed in more detail in the next lecture. The fields $\mathbb{R}$ and $\mathbb{Q}_p$ are called *local fields*. A common approach in number theory is that it is often possible to understand the 'global' solutions of Diophantine equations by finding local obstructions. As the next section discusses in more detail, an important motivation for ultrametric integral cryptanalysis is that this philosophy is also of interest in the context of cryptanalysis.

8. Motivation for ultrametric integral cryptanalysis

Many symmetric-key primitives can be described in terms of a limited number of commutative ring operations. For example, $\mathbb{F}_2^n$ with exclusive-or and bitwise-and is a commutative ring and both of these operations are the natural building blocks of many cryptographic primitives. The theory of linear cryptanalysis maximally simplifies addition, making this a fundamental method to analyze such ciphers. If (and only if) the multiplicative structure of the ring is that of a commutative inverse monoid, then we can do the same for multiplication. Furthermore, note that the theory of linear cryptanalysis is traditionally developed over $\mathbb{R}$ or $\mathbb{C}$. That is, following the philosophy sketched in Section 7, it focuses on local aspects near the prime at infinity. By choosing another absolute value, we can similarly localize

at other primes. Ultrametric integral cryptanalysis is the unique cryptanalytic technique that combines these two aspects:

(1) It simplifies multiplication rather than addition. It turns out that, as a result of this, it generalizes the algebraic approach to integral cryptanalysis.
(2) It uses as the absolute value function the ultrametric p-adic absolute value. In particular, the theory is described over an extension of $\mathbb{Q}_p$.

Large parts of the theory of ultrametric integral cryptanalysis are immediate consequences of the general results for commutative inverse monoids that we have developed in this lecture. However, to be able to use it in applications, the more specific aspects of ultrametric integral cryptanalysis are essential. These will be worked out in the next lecture.

LECTURE 2

Ultrametric integral cryptanalysis

1. Introduction

In the first lecture, we developed the general principles of cryptanalysis over finite commutative inverse monoids. It was shown that in this case, there exists a basis that diagonalizes the pushforward operator of all multiplication maps. This led to a theory of one-dimensional trails following the example of linear cryptanalysis. The goal of this second lecture is to develop ultrametric integral cryptanalysis as a special case of this theory. Ultrametric integral cryptanalysis applies to rings that are also a commutative inverse monoid with respect to multiplication. Like linear cryptanalysis, it is a local theory — but one works over an extension of the field of p-adic numbers rather than over the complex numbers. This lecture builds on the contents of the first lecture, and additionally assumes some familiarity with integral cryptanalysis. To understand some proofs in this lecture, some familiarity with field extensions and commutative rings may be helpful.

2. Rings with commutative inverse monoid structure

Like for commutative inverse monoids, idempotent elements play a central role in the structure of rings. Two idempotent elements e and f are called orthogonal if $ef = 0$. In a ring, if e is an idempotent, then so is $1 - e$. Furthermore, e and $1 - e$ are orthogonal.

2.1. Products of fields. The following theorem shows that rings with a multiplicative commutative inverse monoid structure are isomorphic to products of finite fields. The statement of the theorem requires some terminology related to posets. A nonzero element a of a poset with zero is called an *atom* if there does not exist a nonzero element e such that $e < a$. The set of atoms of X will be denoted by $\mathcal{A}_X$.

Example 6. The atoms of $\mathbb{F}_3 \times \mathbb{F}_3$ are $(0, 1)$ and $(1, 0)$.

Theorem 11. *If a finite ring X is a commutative inverse monoid with multiplication, then there exist finite fields $\mathbb{F}_{q_1}, \ldots, \mathbb{F}_{q_n}$ such that*

$$X \cong \mathbb{F}_{q_1} \times \mathbb{F}_{q_2} \times \cdots \times \mathbb{F}_{q_n} .$$

More precisely, for every atom a of $\mathcal{E}_X$, the set aX is a finite field with unit a and the map $x \mapsto \big(a\,x \mid a \in \mathcal{A}_X\big)$ is a ring isomorphism between X and $\prod_{a \in \mathcal{A}_X} aX$.

Proof. For all elements of the poset $\mathcal{E}_X$ of idempotent elements, there exists an atom smaller than or equal to it. This is true in any finite poset that contains a zero element. We first show that every nonzero idempotent element e can be written as a sum of atoms. There exists an atom $a_1 \leq e$. Since the element $1 - a_1$ is idempotent, the element e is equal to the following sum of idempotent elements:

$$e = a_1\, e + (1 - a_1)\, e = a_1 + (e - a_1)\,,$$

where we use $a_1 \leq e$ and $e - a_1 \leq e$. The element $e - a_1$ is idempotent, and there again exists an atom smaller than it. This process can be iterated until the remaining idempotent element is zero, which must happen eventually since $e - a_1 < e$ and X is finite. Hence, there exist atoms $a_1, \ldots, a_n$ such that

$$e = a_1 + a_2 + \cdots + a_n\,.$$

In the first lecture, it was shown that every nonzero element x of X belongs to exactly one group X_e with unit equal to a nonzero idempotent element e. Hence, there exists an e such that $x = ex$. That is,

$$xe = x\,a_1 + x\,a_2 + \cdots + x\,a_n\,.$$

The values $x\,a_i$ are either zero or contained in X_{a_i} since $e\,a_i = a_i$. In general, for all x in X, there exist elements x_a in aX for all a in $\mathcal{A}_X$, such that

$$x = \sum_{a \in \mathcal{A}_X} x_a\,.$$

The values x_a are uniquely determined by x, because any two atoms $a \neq b$ satisfy $ab = 0$ (their greatest lower bound is zero). Hence, $x_a = a\,x$ so that x_a is uniquely determined.

As mentioned in the theorem statement, aX can be given a ring structure when a is an atom. Indeed, it is sufficient to show that if $x, y \in aX$, then $x - y \in aX$. This is true because $x - y = xa - ya = (x - y)a$. In fact, aX is a field because if $x \in aX$, then either $x = 0$ or $x \in X_a$. This is because there are no nonzero idempotent elements below a. This leads to the homomorphism

$$\begin{aligned} X &\to \prod_{a \in \mathcal{A}_X} aX \\ x &\mapsto \big(a\,x \mid a \in \mathcal{A}_X\big)\,. \end{aligned}$$

By the uniqueness of the decomposition of x above, this is an isomorphism. □

Theorem 11 implies that, without loss of generality, we can assume from now on that X is a product of finite fields. However, this does not mean that the monoid structure can be assumed to be a product of fields in all applications. The reason for this is that the cipher and its decomposition into rounds are generally considered to be input data, whereas the theory is only invariant up to post- and pre-compositions of all functions with the isomorphism and its inverse. Most of the results in this lecture will be stated for the general case, but the proofs will assume a product of fields. The final step of inserting ring isomorphisms in the right places will usually be implicit.

Remark 1. It is also possible to prove Theorem 11 using the Chinese remainder theorem for commutative rings. For every atom a, the set $(1-a)X$ is an ideal of X. Due to the orthogonality of atoms, these ideals are pairwise coprime[1]. The Chinese remainder theorem gives an isomorphism

$$X \cong \prod_{a \in \mathcal{A}_X} X/(1-a)X\,.$$

However, one can show that the ideals $(1-a)X$ are maximal, so that the quotients $X/(1-a)X$ are fields. The isomorphism that is obtained in this way is essentially the same as the one in Theorem 11, since $aX \cong X/(1-a)X$.

[1] If a and b are atoms, then $1 - a + a(1-b) = 1$.

2.2. Rings of characteristic p. For ultrametric integral cryptanalysis, we will for the most part assume that X is a ring of prime characteristic p. This assumption is not necessary from a technical point of view, but it leads to better results in applications and nicer theoretical properties. By Theorem 11, up to ring isomorphism,

$$X = \mathbb{F}_{p^{f_1}} \times \mathbb{F}_{p^{f_2}} \times \cdots \times \mathbb{F}_{p^{f_n}} .$$

Furthermore, up to ring isomorphism we can assume that all $\mathbb{F}_{p^{f_1}}, \ldots, \mathbb{F}_{p^{f_n}}$ are subfields of some field $\mathbb{F}_q$ of characteristic p. For a tuple $v = (v_1, \ldots, v_n)$ with v_i in $\{0, 1, \ldots, p^{f_i} - 1\}$, it is convenient to have the following notation:

$$x^v = \prod_{i=1}^{n} x_i^{v_i} ,$$

for $x = (x_1, \ldots, x_n)$. By composing with the relevant ring isomorphism, this notation extends to any X of prime characteristic p. It holds that $(xy)^v = x^v\, y^v$.

3. Field of p-adic numbers

Let $|\cdot|_p$ denote the p-adic absolute value on $\mathbb{Q}$. At the end of the first lecture, it was mentioned that the field of p-adic numbers $\mathbb{Q}_p$ is defined as the metric completion of $\mathbb{Q}$ with respect to $|\cdot|_p$. Below, we develop an alternative definition that is more algebraic. It will be shown at the end of this section that this definition coincides with the metric completion. The algebraic definition has the advantage that lends itself better to calculations, and in particular to the lifting lemma of Hensel.

3.1. Algebraic definition. The p-adic integers $\mathbb{Z}_p$ are defined as the following inverse limit (this concept is explained below):

$$\mathbb{Z}_p = \varprojlim \mathbb{Z}/p^n\mathbb{Z} .$$

Explicitly, $\mathbb{Z}_p$ consists of all sequences $x = (x_1, x_2, \ldots)$ with x_n in $\mathbb{Z}/p^n\mathbb{Z}$ and such that $x_n \equiv x_{n+1} \pmod{p^n}$. Addition and multiplication are defined coordinate-wise, and this makes $\mathbb{Z}_p$ into an integral domain. The p-adic absolute value is defined as $|x|_p = p^{-m}$, where $m \geq 1$ is the smallest integer such that $x_m \neq 0$. The integer m is called the p-adic valuation $\operatorname{ord}_p x$ of x.

The p-adic numbers are then defined as the field of fractions of $\mathbb{Z}_p$. Using Hensel lifting (see below), it can be shown that if $u \in \mathbb{Z}_p$ with $|u|_p = 1$, then also $u^{-1} \in \mathbb{Z}_p$. This implies that every element x of $\mathbb{Q}_p$ can be written as $p^m\, u$, where u is a unit of $\mathbb{Z}_p$ and m is an integer. The p-adic absolute value of x is equal to p^{-m}.

Since every element of $\mathbb{Z}_p$ is a unit times a power of p, the only ideals of $\mathbb{Z}_p$ are of the form $p^m\mathbb{Z}_p$. This means that $\mathbb{Z}_p$ has a unique maximal ideal $p\mathbb{Z}_p$ (such a ring is called a discrete valuation ring). From the inverse limit definition of $\mathbb{Z}_p$, it is clear that $\mathbb{Z}_p/p\mathbb{Z}_p$ is a finite field of order p.

3.2. Hensel lifting. Hensel's lifting lemma is an essential result for both the theory of p-adic numbers and its applications. In practice, the proof of this theorem is just as important as its statement, at it provides a way to calculate with p-adic integers (for example computing inverses, computing square roots, …).

THEOREM 12 (Hensel lemma). *Let $f(t)$ be a polynomial over $\mathbb{Z}_p$ with derivative $f'(t)$. If x_1 in $\mathbb{Z}/p\mathbb{Z}$ satisfies $f(x_1) \equiv 0 \pmod{p}$ and $f'(x_1) \not\equiv 0 \pmod{p}$, then there exists a unique p-adic integer x such that $f(x) = 0$ and $x \equiv x_1 \pmod{p}$.*

PROOF. The proof is a p-adic analogue of Newton's method. Let $f(t) = \sum_{i=0}^{n} c_i t^i$. Using the binomium, we obtain the following p-adic analogue of a first-order Taylor approximation:

$$f(t + hp) \equiv \sum_{i=0}^{n} c_i \, (t + hp)^i \equiv f(t) + hp \, f'(t) \pmod{p^2}.$$

Since $f(x_1) \equiv 0 \pmod{p}$, there exists an integer λ such that

$$f(x_1 + hp) \equiv p\big(\lambda + hf'(x_1)\big) \pmod{p^2}.$$

The choice $h \equiv -\lambda / f'(x_1) \pmod{p}$ leads to a unique x_2 in $\mathbb{Z}/p^2\mathbb{Z}$ such that we have $f(x_2) \equiv 0 \pmod{p^2}$. The same argument can now be applied to

$$f(t + hp^2) \equiv f(t) + hp^2 f'(t) \pmod{p^3}.$$

Repeating this process yields an infinite sequence $(x_1, x_2, \ldots)$ such that $x_n \in \mathbb{Z}/p^n\mathbb{Z}$ with $f(x_n) \equiv 0 \pmod{p^n}$. This infinite sequence is a p-adic integer. □

As a first application of Hensel lifting, we establish the following result.

THEOREM 13. *For every x in $\mathbb{F}_p^\times$, the p-adic integers $\mathbb{Z}_p$ contain a unique $(p-1)^{st}$ root of unity $\tau(x)$ such that $\tau(x) \equiv x \pmod{p}$.*

PROOF. Let $f(t) = t^{p-1} - 1$. The derivative is $f'(t) = (p-1)\, t^{p-2}$, which is nonzero for nonzero t. The order of the multiplicative group $\mathbb{F}_p^\times$ is $p - 1$, so the result follows by applying Theorem 12. □

The function $\tau : \mathbb{F}_p \to \mathbb{Q}_p$ that maps zero to zero and x in $\mathbb{F}_p^\times$ to the corresponding root of unity is called the *Teichmüller character* of $\mathbb{F}_p$. This terminology is justified because τ is a monoid homomorphism.

3.3. Unramified extensions of $\mathbb{Q}_p$. To construct all the multiplicative characters of our ring X, it will in general be necessary to add some roots of unity to $\mathbb{Q}_p$. Let ζ be a primitive $(q-1)^{\text{st}}$ root of unity, with q a power of p. The p-adic absolute value can be extended to the extension field $\mathbb{Q}_p(\zeta)$ by defining

$$|x|_p = \sqrt[n]{|N_{\mathbb{Q}_p(\zeta)}(x)|_p}, \tag{1}$$

where $N_{\mathbb{Q}_p(\zeta)}$ is the absolute field norm and $n = [\mathbb{Q}_p(\zeta) : \mathbb{Q}_p]$. Intuitively, this is the right choice because the norm function is multiplicative and absolute values must be invariant under field automorphisms.

LEMMA 4. *Let ζ be a primitive $(q-1)^{st}$ root of unity, with q a power of a prime p. The only absolute value on $\mathbb{Q}_p(\zeta)$ that extends the p-adic absolute value of $\mathbb{Q}_p$ is given by* (1).

PROOF SKETCH. There are two claims to be verified: that (1) defines an absolute value, and that it is unique. For the first claim, we give an incomplete argument because the proof of the ultrametric triangle inequality is rather technical.

It is not difficult to check that $|\cdot|_p$ is multiplicative and $|x|_p = 0$ if and only if $x = 0$. The ultrametric triangle inequality can be proven by expressing the norm as the determinant of a matrix over $\mathbb{Q}_p$, but we will not work out the details.

To show that (1) is the unique absolute value on $\mathbb{Q}_p(\zeta)$ extending the p-adic absolute value, note that $|\cdot|_p$ is also a norm on the finite-dimensional $\mathbb{Q}_p$-vector space $\mathbb{Q}_p(\zeta)$. However, all such norms are equivalent. Hence, if $|\cdot|'_p$ is a second absolute value function on $\mathbb{Q}_p(\zeta)$, then there exist absolute positive constants c_1 and c_2 such that for all x in $\mathbb{Q}_p(\zeta)$

$$c_1 |x|_p \leq |x|'_p \leq c_2 |x|_p .$$

An application of the 'power trick' then shows that there is only one such norm. That is, for all $n \geq 1$, it holds that

$$c_1 |x^n|_p \leq |x^n|'_p \leq c_2 |x^n|_p .$$

Taking n^{th} roots and the limit $n \to \infty$ implies that $|x|'_p = |x|_p$ for all x in $\mathbb{Q}_p(\zeta)$. □

Hensel's lifting lemma can be generalized to $\mathbb{Q}_p(\zeta)$. To do this, we need some results about the algebraic structure of $\mathbb{Q}_p(\zeta)$.

LEMMA 5. *Let ζ be a primitive $(q-1)^{st}$ root of unity, with q a power of a prime p. The ring of integers of $\mathbb{Q}_p(\zeta)$ is equal to $\mathbb{Z}_p[\zeta]$. Furthermore, $p\mathbb{Z}_p[\zeta]$ is a maximal ideal of $\mathbb{Z}_p[\zeta]$.*

PROOF. Since the norm of an integral element of $\mathbb{Q}_p(\zeta)$ is always a p-adic integer, the integers of $\mathbb{Q}_p(\zeta)$ also satisfy $|x|_p \leq 1$. It is clear that $\mathbb{Z}_p[\zeta]$ is contained in the ring of integers of $\mathbb{Q}_p(\zeta)$. The converse is also true, because if $x = \sum_{i=0}^{n-1} a_i \zeta^i$ is an integer of $\mathbb{Q}_p(\zeta)$ with $|a_i|_p > 1$ for i in $I \subseteq \{0, 1, \ldots, n-1\}$, then $\sum_{i \in I} a_i \zeta^i$ would be an integral element with absolute value > 1. Finally, to see that the ideal generated by p is maximal, note that

$$\mathbb{Z}_p[\zeta]/p\mathbb{Z}_p[\zeta] \cong \mathbb{F}_p(\zeta) .$$

Since ζ is a primitive $(q-1)^{\text{st}}$ root of unity, $\mathbb{F}_p(\zeta)$ is a field with q elements. Hence, $p\mathbb{Z}_p[\zeta]$ is a maximal ideal. □

The extension $\mathbb{Q}_p(\zeta)/\mathbb{Q}_p$ is called unramified because the ideal $p\mathbb{Z}_p$ does not factor nontrivially as a product of prime ideals. Extensions in which $p\mathbb{Z}_p$ factors nontrivially are called ramified, but we will not discuss them in this lecture as they are only necessary to define the additive Fourier transformation over a p-adic field.

The field $\mathbb{F}_q = \mathbb{Z}_p[\zeta]/p\mathbb{Z}_p[\zeta]$ that came up in the proof of Lemma 5 is called the *residue field* of $\mathbb{Q}_p(\zeta)$. Hensel's lifting lemma carries over to polynomials over $\mathbb{Z}_p[\zeta]$. In this case we lift a solution in the residue field $\mathbb{F}_q$ to a solution in $\mathbb{Z}_p[\zeta]$. The proof is exactly the same as for Theorem 12.

THEOREM 14 (Hensel lemma). *Let ζ be a primitive $(q-1)^{st}$ root of unity, with q a power of a prime p, and let $f(t)$ be a polynomial over $\mathbb{Z}_p[\zeta]$ with derivative $f'(t)$. If x_1 in $\mathbb{Z}_p[\zeta]/p\mathbb{Z}_p[\zeta]$ satisfies $f(x_1) \equiv 0 \pmod{p}$ and $f'(x_1) \not\equiv 0 \pmod{p}$, then there exists a unique x in $\mathbb{Z}_p[\zeta]$ such that $f(x) = 0$ and $x \equiv x_1 \pmod{p}$.*

Theorem 14 implies that there exists a Teichmüller character $\tau : \mathbb{F}_q \to \mathbb{Q}_p(\zeta)$ such that $\tau(0) = 0$ and $\tau(x)$ is the unique $(q-1)^{\text{st}}$ root of unity such that $\tau(x) \equiv x \pmod{p}$. This character will be the starting point to construct the set of all characters of X in the next section.

3.4. Metric definition. The purpose of this section is to show that the above definition of $\mathbb{Q}_p$ coincides (up to isometry) with the metric completion of $\mathbb{Q}$ with respect to the p-adic absolute value. Since this is only necessary to motivate ultrametric integral cryptanalysis, this section may be skipped.

A sequence $x_1, x_2, \ldots$ is called a Cauchy sequence if for every real $\varepsilon \geq 0$, there exists an integer $N \geq 1$ such that for all integers $n, m \geq N$, it holds that $|x_n - x_m|_p < \varepsilon$. Two Cauchy sequences $x_1, x_2, \ldots$ and $y_1, y_2, \ldots$ are said to be equivalent if their distance is zero:

$$\lim_{n\to\infty} |x_n - y_n|_p = 0\,.$$

It follows from standard properties of limits that this indeed defines an equivalence relation on the set of Cauchy sequences. The completion of $\mathbb{Q}$ with respect to $|\cdot|_p$ can be defined as the set of these equivalence classes of Cauchy sequences. The absolute value of a Cauchy sequence $x_1, x_2, \ldots$ is defined as its limiting absolute value $\lim_{n\to\infty} |x_n|_p$.

The construction of the metric completion can be simplified because of the fact that the image of $|\cdot|_p$ is a discrete set. In particular, for any Cauchy sequence $x_1, \ldots, x_n$, we construct a particular subsequence $y_1, \ldots, y_n$ that is equivalent to the original sequence. Let $y_1 = x_{N_1}$ with N_1 the smallest N_1 such that $|x_n - x_m| < 1$ for all $n, m \geq N$. Then choose $y_2 = x_{N_2}$ with $N_2 \geq N_1$ in the same way with $\varepsilon = 1/p$. The subsequence constructed in this way with $\varepsilon = 1, 1/p, 1/p^2, \ldots$ has the property that $\lim_{n\to\infty} |x_n - y_n|_p = 0$. Furthermore, it satisfies $|y_{n+1} - y_n|_p < 1/p^n$. This implies that there exist rational numbers $c_1, \ldots, c_n$ such that

$$y_n = \sum_{i=0}^{n} c_i\, p^i\,.$$

The sequence can always be rewritten as a sequence

$$s_n = \sum_{i=-m}^{n} c_i\, p^i\,,$$

with $m \geq 0$ and $c_{-m}, \ldots, c_n \in \{0, 1, \ldots, p-1\}$. In other words, we can think of a p-adic number x as a convergent series

$$x = \sum_{i=-m}^{\infty} c_i\, p^i\,. \tag{2}$$

For $m = 0$, every sequence $(s_1, s_2, \ldots, s_n)$ is an element of the inverse limit $\mathbb{Z}_p$. Conversely, every element of the inverse limit gives rise to a unique convergent series with $m = 0$. Furthermore, every p-adic number is of the form $p^m\, u$ with u a unit of $\mathbb{Z}_p$. Hence, this gives an isometric isomorphism of fields.

4. Ultrametric integral change-of-basis

We now define ultrametric integral cryptanalysis for a ring X with commutative inverse monoid structure and prime characteristic. The local field of the theory will be chosen as an extension of $\mathbb{Q}_p$, where p is the characteristic of X. As far as generalities are concerned, this is not required — nor is it required that the characteristic of X is prime. For completeness, we briefly mention what happens if these assumptions do not hold:

- If the characteristic of X is not prime, then it makes sense to consider all of its prime divisors p. For the most part, one then ends up ignoring the factors of X with characteristic different from p (in other words, a localization of X). Depending on the prime numbers, it can be necessary to work in a ramified extension of $\mathbb{Q}_p$.
- If p does not divide the characteristic of X, then most of the specific results below cannot be rescued in a straightforward way.

Since there are currently no clear applications of the general case, we do not develop it any further here.

4.1. Characters of X. Let $X = \mathbb{F}_{p^{f_1}} \times \mathbb{F}_{p^{f_2}} \times \cdots \times \mathbb{F}_{p^{f_n}}$ so that $\mathbb{F}_{p^{f_1}}, \ldots, \mathbb{F}_{p^{f_n}}$ are subfields of a larger field $\mathbb{F}_q$. There is a primitive $(q-1)^{\text{st}}$ root of unity ζ_{q-1} so that $k = \mathbb{Q}_p(\zeta_{q-1})$ has residue field $\mathbb{F}_q$. For every tuple $v = (v_1, \ldots, v_n)$ with v_i in $\{0, 1, \ldots, p^{f_i} - 1\}$, we have a character of X defined by

$$\begin{aligned} \chi \colon X &\to k \\ x &\mapsto \tau(x^v)\,. \end{aligned}$$

Several important properties that will be presented in this lecture and the next do not depend on the precise value of v but only on its Hamming weight $\mathrm{wt}(v)$ or what we will call its p-weight $\mathrm{wt}_p(v)$. The latter is defined as follows:

$$\mathrm{wt}_p(v) = \sum_{i=1}^{n} s_p(v_i)\,,$$

where $s_p(v_i)$ is the sum of the base-p digits of v_i. It will be convenient to define $\mathrm{wt}(\chi) = \mathrm{wt}(v)$ and $\mathrm{wt}_p(\chi) = \mathrm{wt}_p(v)$. With this notation, the Hamming weight is the rank function of the lattice of idempotents of $\widehat{X}$.

The 'digit sum' function s_p can be tricky to work with. The following lemma lists two of its properties that we will use later.

LEMMA 6. *Let p be a positive integer and q a non-negative power of p. The digit sum function s_p satisfies the following properties:*

(1) For all non-negative integers x and y,

$$s_p(x+y) = s_p(x) + s_p(y) - (p-1)\,c(x,y)\,,$$

where $c(x,y)$ is the number of carries when x and y are added in base p.

(2) For all non-negative integers $x < q$ and k, $s_p(x) \leq s_p(x + k(q-1))$.

PROOF. The first property follows from the relation between the digits of the integers $z = x + y$, x and y:

$$z_i = x_i + y_i + c_{i-1} - p\,c_i\,,$$

where c_i is the i^{th} carry bit. In particular $c_i = 0$ and if z has a total of n digits, then $c_d = 0$. Summing these relations, we obtain

$$\sum_{i=1}^{d} z_i = \sum_{i=1}^{d} x_i + y_i + c_{i-1} - pc_i = s_p(x) + s_p(y) - (p-1)\sum_{i=1}^{d} c_i + \sum_{i=1}^{d} (c_i - c_{i-1})\,.$$

The last sum is telescoping and sums to $c_d - c_0 = 0$. Hence, the result follows. The second property follows from the inequality $s_p(x-y) \geq s_p(x) - s_p(y)$ for all

non-negative integers x and $y \leq x$. Indeed, since $x = (x-y)+y$, this inequality follows from the first property. Hence, $s_p(x+k(q-1))$ is lower bounded by

$$s_p(x+k(q-1)) \geq s_p(x+kq) - s_p(k) = s_p(qx) + s_p(qk) - s_p(k) = s_p(x)\,.$$

Here, the first equality follows from the assumption $x < q$. □

The following inequality is a useful consequence of Lemma 6:

$$\mathrm{wt}_p(\chi_1 \cdots \chi_l) \leq \mathrm{wt}_p(\chi_1) + \cdots + \mathrm{wt}_p(\chi_l)\,.$$

Note that this really relies on both parts of the lemma.

4.2. Eigenvector basis. As discussed in the first lecture, the functions $b^\chi = \chi$ define a suitable basis for the vector space k^X. To construct the dual basis vectors b_χ, the Fourier inversion formula can be used. It is sufficient to construct them for $\mathbb{F}_{p^e}$, since the dual basis vectors of X are obtained as the tensor products of a submonoid of these. By Fourier inversion, if $\chi \neq 1$, then

$$b_\chi = \frac{1}{p^e-1} \sum_{x \in \mathbb{F}_{p^e}^\times} \chi(1/x)\,\delta_x + \begin{cases} -\delta_0 & \text{if } \chi(x) = \tau(x^{p^e-1}) \text{ for all } x \text{ in } \mathbb{F}_{p^e}\,, \\ 0 & \text{else}\,. \end{cases}$$

For $\chi = 1$, we have $b_\chi = \delta_0$. The explicit formula for the basis vectors is rarely necessary. The basis vectors for $k[X]$ are then of the form

$$b_{\chi_1} \otimes \cdots \otimes b_{\chi_n}\,,$$

with χ_i a character of $\mathbb{F}_{p^{f_i}}$.

4.3. Fourier transformation. The change-of-basis transformation will be denoted by $\mathscr{U} : k[X] \to k[\widehat{X}]$. To avoid confusion with the additive Fourier transformation, it will be called the *ultrametric integral change-of-basis* map. The main properties of $\mathscr{U}$ and its dual $\mathscr{U}^{-\vee}$ have already been described in the first lecture. The following property is analogous to the unitarity of the Fourier transformation. In particular, the p-adic analogue of the Euclidean norm is the maximum norm

$$\left\| \sum_{x \in X} a_x\,\delta_x \right\| = \max_{x \in X} |a_x|_p\,.$$

The dual norm on k^X is given by essentially the same formula.

THEOREM 15. *The matrix representation of the ultrametric change-of-basis map $\mathscr{U} : k[X] \to k[\widehat{X}]$ with respect to the standard bases of $k[X]$ and $k[\widehat{X}]$ is unimodular. In particular, $\mathscr{U}$ is an isometry with respect to the p-adic maximum norm.*

PROOF. An invertible matrix over k is unimodular if its coordinates and the coordinates of its inverse are integer elements of k. The coordinates of the matrix representation of $\mathscr{U}$ are given by

$$\delta^\chi\big(\mathscr{U}\,\delta_x\big) = \chi(x)\,.$$

Since $|\chi(x)|_p \leq 1$, it is an integer element of k. Similarly, for the inverse we have

$$\delta^x\big(\mathscr{U}^{-1}\,\delta_\chi\big) = \delta^x\big(b_\chi\big)\,.$$

By the explicit formulas above, it is clear that $|\delta^x\big(b_\chi\big)|_p \leq 1$. Finally, note that any unimodular matrix defines an isometry with respect to the p-adic maximum norm.

Indeed, if the coordinates of $\mathscr{U}$ are integer elements of k, then $\|\mathscr{U}\, x\| \leq \|x\|$ for all x. Likewise, $\|\mathscr{U}^{-1} x\| \leq \|x\|$. It follows that $\|\mathscr{U}\| = 1$. □

It is useful to compute some examples of the matrix representation of $\mathscr{U}$ for small primes p and $n = 1$. For $p \leq 3$, no p-adic numbers apart from those also in $\mathbb{Q}$ appear.

EXAMPLE 7. Let $X = \mathbb{F}_2$. The rows of $\mathscr{U}$ correspond to the basis functions $x \mapsto 1$ and $x \mapsto \tau(x)$, where $\tau(0) = 0$ and $\tau(1) = 1$. The columns of $\mathscr{U}^{-1}$ correspond to the basis vectors δ_0 and $\delta_1 - \delta_0$. Hence,

$$\mathscr{U} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad \mathscr{U}^{-1} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}.$$

Note that all coordinates of $\mathscr{U}$ and $\mathscr{U}^{-1}$ are integers in $\mathbb{Q} \hookrightarrow \mathbb{Q}_2$.

EXAMPLE 8. Let $X = \mathbb{F}_5$. The rows of $\mathscr{U}$ correspond to the basis functions $x \mapsto \tau(x)^v$, where $\tau(0) = 0$, $\tau(1) = 1$, $\tau(2) = i$, $\tau(3) = -i$, $\tau(4) = -1$. Here, i the unique 5-adic square root of -1 so that $i \equiv 2 \pmod 2$. The first few digits are given by

$$i = 2 + 1 \cdot 5 + 2 \cdot 5^2 + 1 \cdot 5^3 + 3 \cdot 5^4 + \cdots$$

The columns of $\mathscr{U}^{-1}$ correspond to the basis vectors δ_0, $(\delta_1 - i\delta_2 + i\delta_3 - \delta_4)/4$, and so on. Hence,

$$\mathscr{U} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & i & -i & -1 \\ 0 & 1 & -1 & -1 & 1 \\ 0 & 1 & -i & i & -1 \\ 0 & 1 & 1 & 1 & 1 \end{bmatrix} \quad \text{and} \quad \mathscr{U}^{-1} = \frac{1}{4} \begin{bmatrix} 4 & 0 & 0 & 0 & -4 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & -i & -1 & i & 1 \\ 0 & i & -1 & -i & 1 \\ 0 & -1 & 1 & -1 & 1 \end{bmatrix}.$$

Note that all coordinates of $\mathscr{U}$ and $\mathscr{U}^{-1}$ are integers in $\mathbb{Q}_5$ (but not in $\mathbb{Q}$ or $\mathbb{Q}(i)$).

4.4. Integral cryptanalysis. An important property of ultrametric integral cryptanalysis is that if all quantities are reduced modulo p, then one obtains integral cryptanalysis over the residue field $\mathbb{F}_q$. In particular, if $q = 2$ and $X = \mathbb{F}_2^n$, then we obtain the description of integral cryptanalysis using the dual pair of 'precursor' and 'monomial' bases (leading to algebraic transition matrices [**2**]). Integral cryptanalysis with $q > 2$ has not been fully developed in the literature, but as shown below it is reasonable to define it as the modulo p reduction of ultrametric integral cryptanalysis.

EXAMPLE 9. If $X = \mathbb{F}_2^n$, then the characters are of the form $x \mapsto \tau(x^v)$ with v in $\{0,1\}^n$. The image of τ is just $\{0,1\}$ in this case. Modulo two, the basis functions b^χ are given by monomials $\chi(x) = x^v$. For $n = 1$, the basis vectors b_χ are equal to δ_0 and $\delta_0 + \delta_1$ modulo two. Hence, for general n,

$$b_\chi \equiv \sum_{x \preccurlyeq v} \delta_x \pmod 2.$$

This is the indicator function of the predecessor set of v. If S is a subset of $\mathbb{F}_2^n$, then up to identifying $\widehat{\mathbb{F}_2^n}$ and $\mathbb{F}_2^n$,

$$\mathscr{U}\, \delta_S \equiv \delta_{\mathcal{U}(S)} \pmod 2,$$

where $\mathcal{U}(S)$ is the parity set of S [**4**].

The concept of parity sets does not extend well beyond $\mathbb{F}_2$. The reason is that knowing the support of $\mathscr{U}\ \delta_S$ is in general not sufficient to recover S. It just happens to be the case that elements of the free vector space over $\mathbb{F}_2$ can be identified with sets. The division property [**5**] describes the weight of the elements in the parity set. This idea can be generalized to ultrametric integral cryptanalysis by saying that a set S has the $1/\varepsilon$-division property of order k if

$$\left|\delta^{\chi}\big(\mathscr{U}\ \delta_S\big)\right|_p \leq \varepsilon\,,$$

for all characters χ with $\mathrm{wt}_p(\chi) < k$. An equivalent definition is given below. A set then has the p-division property of order k if all monomials of p-degree strictly less than k sum to zero on the set.

Definition 3 (p^l division property). A multiset S with elements from X satisfies the p^l division property of order k if

$$\sum_{x\in S} \chi(x) \equiv 0 \pmod{p^l}\,,$$

for all characters χ of X with $\mathrm{wt}_p(\chi) < k$.

Although we will not use Definition 3 directly in applications, it will be useful for the analysis of generic constructions (functions of bounded degree, Feistel ciphers, ...) to characterize a set of trails by a sequence of upper bounds on the weights of the characters that they are defined on. In some cases, however, keeping track the p-weights is not sufficient. Furthermore, to analyze specific primitives, it is usually important to keep track of more precise information.

5. Ultrametric integral transition matrices

The ultrametric integral transition matrix of a function F is defined as

$$A^{\mathsf{F}} = \mathscr{U}\, T^{\mathsf{F}}\, \mathscr{U}^{-1}\,.$$

Since $\mathscr{U}$ is a Fourier transformation in the sense of the first lecture, all the properties that were mentioned there carry over. After mentioning a few examples of what these properties lead to, some results specific to ultrametric integral transition matrices will be given.

5.1. Basic properties. As discussed in the first lecture, multiplication by a constant m in X corresponds to a diagonal matrix with entries $\chi(m)$ in row and column χ. If F is a monoid homomorphism, then $A^{\mathsf{F}}_{\chi,\psi} = 1$ if $\psi = \chi \circ \mathsf{F}$ and zero otherwise. Applying this to multiplication and copy operations led to the propagation rules illustrated below.

(A) Copy. (B) Multiplication.

Figure 1. Propagation rules for copy and multiplication operations.

Other important classes of monoid homomorphisms are permutations of the factors of X (for $\mathbb{F}_2^n$ these are the bit-permutations) and power maps. If $\mathsf{F} : x \mapsto x^d$ is power map on a finite field, then

$$A^{\mathsf{F}}_{\chi,\psi} = \begin{cases} 1 & \text{if } \psi = \chi^d \\ 0 & \text{else}\,. \end{cases}$$

In the next lecture, we will use this propagation rule in the form $\mathrm{wt}_p(\psi) = \mathrm{wt}_p(\chi^d)$. There are two useful upper bounds related to this:

$$\begin{aligned} \mathrm{wt}_p(\chi^d) &\leq d\ \mathrm{wt}_p(\chi) \\ \mathrm{wt}_p(\chi^d) &\leq s_p(d)\ \mathrm{wt}_p(\chi)\,. \end{aligned}$$

The first upper bound follows from $\mathrm{wt}_p(\chi^d) = \mathrm{wt}_p(\chi \cdots \chi) \leq \mathrm{wt}_p(\chi) + \cdots \mathrm{wt}_p(\chi)$. For the second upper bound, note that there exist non-negative integers $u \leq q-1$ and k such that $u + k(q-1) = vd$. Let $d = \sum_{i=0}^{l-1} d_i\, p^i$. By Lemma 6, it holds that

$$s_p(u) \leq s_p(vd) = s_p\big(v\, d_0 + v\, d_1\, p + \cdots + v\, d_{l-1}\, p^{l-1}\big) \leq s_p(v) \sum_{i=0}^{l-1} s_p(d_i\, p^i)\,.$$

The result follows from $s_p(d_i\, p^i) = d_i$. The bound $\mathrm{wt}_p(\chi^d) \leq s_p(d)\ \mathrm{wt}_p(\chi)$ is better than $\mathrm{wt}_p(\chi^d) \leq d\ \mathrm{wt}_p(\chi)$, but both will be important for the discussion of the Ax-Katz theorem in the next lecture.

5.2. Reduction modulo p. It was shown in Theorem 15 that the coordinates of the matrices $\mathscr{U}$ and $\mathscr{U}^{-1}$ are p-adic integers. This means that the same is true for $A^{\mathsf{F}} = \mathscr{U}\ T^{\mathsf{F}}\ \mathscr{U}^{-1}$, and consequently it makes sense to reduce A^{F} modulo p. This yields a matrix that is closely related to the algebraic normal form of F.

The algebraic normal form of a function $f : X \to \mathbb{F}_{q_{f_n}}$ is a unique interpolating 'polynomial' in the ring $\mathbb{F}_q[x_1, \ldots, x_n]/(x_1^{p^{f_1}} - x_1, \ldots, x_n^{p^{f_n}} - x_n)$. In the following theorem, the notation F^v refers to the map $x \mapsto \mathsf{F}^v(x)$.

THEOREM 16. *Let $\mathsf{F} : X \to Y$ be a function with ultrametric integral transition matrix A^{F}. For all characters $\psi(x) = \tau(x^u)$ and $\chi(x) = \tau(x^v)$, the value $A^{\mathsf{F}}_{\chi,\psi}$ is congruent modulo p to the coefficient of x^u in the algebraic normal form of F^v.*

PROOF. By the definition of A^{F}, it holds that

$$A^{\mathsf{F}^\vee}\, \delta^\chi = \sum_\psi A^{\mathsf{F}^\vee}_{\psi,\chi}\, \delta^\psi = \sum_\psi A^{\mathsf{F}}_{\chi,\psi}\, \delta^\psi\,.$$

Evaluating at δ_x and reducing modulo p yields

$$\mathsf{F}^v(x) = \sum_{\substack{\psi \in \widehat{X} \\ \psi :\, x \mapsto \tau(x^u)}} A^{\mathsf{F}}_{\chi,\psi}\, x^u \pmod p\,.$$

Hence, we have obtained an interpolating polynomial for F^v. However, this polynomial is unique modulo the ideal generated by $x^{f_i} - x$, so its coefficients correspond to those of the algebraic normal form. □

5.3. Unimodularity. In linear cryptanalysis, correlation matrices are unitary. More generally, correlation matrices of injections are isometries with respect to the Euclidean norm. The following theorem is the ultrametric integral analogue of this result.

THEOREM 17. *If* F *is an injection, then* A^{F} *is an isometry with respect to the p-adic maximum norm. Furthermore, if* F *is invertible, then* A^{F} *is unimodular.*

PROOF. If F is an injection, then T^{F} is an isometry because $T^{\mathsf{F}}\,v$ and v have the same nonzero coordinates. Since $\mathscr{U}$ and $\mathscr{U}^{-1}$ are isometries, it follows that A^{F} is an isometry. The result about unimodularity follows by a similar argument, using the fact that T^{F} is unimodular if and only if it is invertible. □

6. Addition

In this section we analyze the ultrametric integral properties of addition in the ring X. As we will see in the next lecture, the analysis of most generic constructions essentially reduces to understanding addition. The absolute values of the ultrametric integral transition matrix for addition can be obtained from Stickelberger's theorem on Gauss sums. However, here we follow a more down-to-earth approach based on Kummer's theorem on binomial coefficients.

6.1. Lifting using Frobenius. In our analysis of addition we will reduce A^{F} modulo a power of p. To do this, we will rely on Lemma 7.

LEMMA 7. *Let* x *and* y *be integer elements of* k*, and* n *a positive integer. If* $x \equiv y \pmod{p^n}$*, then* $x^p \equiv y^p \pmod{p^{n+1}}$.

PROOF. If $x \equiv y \pmod{p^n}$, then $y = x + m\,p^n$ for some integer element m of k. By the binomial formula, we have

$$y^p = (x + m\,p^n)^p = \sum_{i=0}^{p} \binom{p}{i} x^{p-i}\,(mp^n)^i\,.$$

All terms on the right-hand side are divisible by p^{n+1}, except the term for $i = 0$. This term is equal to x^p, so the result follows. □

An interesting consequence of Lemma 7 is the formula $\tau(x) = \lim_{n\to\infty} x^{q^n}$.

6.2. Kummer's theorem. Kummer's theorem gives a formula for the p-adic valuation of binomial coefficients in terms of the s_p function. We will deduce it from Legendre's theorem, which does the same as Kummer's theorem but for factorials.

LEMMA 8 (Legendre). *For all non-negative integers* n*,*

$$\operatorname{ord}_p n! = \frac{n - s_p(n)}{p-1}\,.$$

PROOF. Generalizing the argument used in the proof of Lemma 6, we can see that

$$s_p(i-1) = s_p(i) - 1 + (p-1)\,\operatorname{ord}_p i\,.$$

Indeed, the number of times we borrow is equal to the number of trailing zeros in the base p expansion of i, which equals $\operatorname{ord}_p i$. Since $\operatorname{ord}_p n! = \operatorname{ord}_p n + \cdots +$

$\operatorname{ord}_p 2 + \operatorname{ord}_p 1$, summing up the above equation for i in $\{1, 2, \ldots, n\}$ yields

$$\sum_{i=1}^{n} s_p(i-1) = \sum_{i=1}^{n} s_p(i) - n + (p-1)\operatorname{ord}_p n!\,.$$

It follows from this that $s_p(n) - n + (p-1)\operatorname{ord}_p n! = 0$. □

THEOREM 18 (Kummer). *For all non-negative integers n and m, $\operatorname{ord}_p \binom{n}{m}$ is equal to the number of carries when adding m to $n-m$ in base p. That is,*

$$\operatorname{ord}_p \binom{n}{m} = \frac{s_p(m) + s_p(n-m) - s_p(n)}{p-1}\,.$$

PROOF. The result follows directly from Legendre's theorem:

$$\operatorname{ord}_p \binom{n}{m} = \operatorname{ord}_p n! - \operatorname{ord}_p m! - \operatorname{ord}_p (n-m)!\,.$$

The fact that this is the number of carries when adding m and $n-m$ in base p follows from the first part of Lemma 6. □

6.3. Addition with two inputs. We will determine only the divisibility of the coordinates of the ultrametric integral transition matrix for addition — although exact values can be obtained with a little more work.

THEOREM 19. *Let $\mathsf{F}\colon X \times X \to X$ with $\mathsf{F}(x,y) = x + y$. For all characters χ, ψ_1 and ψ_2 of X, we have $A^{\mathsf{F}}_{\chi,(\psi_1,\psi_2)} = 0$ unless $\chi = \psi_1\,\psi_2$. If $\psi_1\psi_2 \neq 1$ then*

$$\operatorname{ord}_p A^{\mathsf{F}}_{\psi_1\psi_2,(\psi_1,\psi_2)} \geq \frac{\operatorname{wt}_p(\psi_1) + \operatorname{wt}_p(\psi_2) - \operatorname{wt}_p(\psi_1\psi_2)}{p-1}\,.$$

PROOF. It is sufficient to prove the result for $X = \mathbb{F}_q$, since the conclusion then follows by taking tensor products. Since $\tau(x+y) \equiv \tau(x) + \tau(y) \pmod{p}$, Lemma 7 shows that

$$\big(\tau(x) + \tau(y)\big)^q \equiv \tau(x+y)^q \equiv \tau(x+y) \pmod{pq}\,.$$

To compute the coordinates of A^{F}, we first compute $\chi \circ \mathsf{F}$. If $\chi(x) = \tau(x^v)$, then

$$\chi(\mathsf{F}(x,y)) \equiv \tau(x+y)^v \equiv \big(\tau(x) + \tau(y)\big)^{qv} \pmod{pq}\,.$$

Expanding the right-hand side using the binomium formula yields the expression

$$\chi(\mathsf{F}(x,y)) \equiv \sum_{i=0}^{qv} \binom{qv}{i} \tau(x)^i \tau(y)^{qv-i} \pmod{pq}\,.$$

The functions $\omega_i\colon x \mapsto \tau(x)^i$ are characters of X. Hence, we have the following equality of functions:

$$T^{\mathsf{F}^\vee} b^\chi \equiv \sum_{i=0}^{qv} \binom{qv}{i} b^{\omega_i} \otimes b^{\omega_{qv-i}} \pmod{pq}\,.$$

Since $A^{\mathsf{F}}_{\chi,\psi} = \big(T^{\mathsf{F}^\vee} b^\chi\big)\big(b_\psi\big)$, we can evaluate the above expression in $b_\psi = b_{\psi_1} \otimes b_{\psi_2}$. The only nonzero terms in the resulting sum are those for which $\psi_1 = \omega_i$ and

$\psi_2 = \omega_{qv-i}$. This means that the sum is zero unless $\psi_1\psi_2 = \omega_i\,\omega_{qv-i} = \omega_{qv} = \chi$. For the second part of the result, we continue with the expression

$$A^{\mathsf{F}}_{\chi,\psi} \equiv \sum_{\substack{i=0 \\ \omega_i = \psi_1 \\ \omega_{qv-i} = \psi_2}}^{qv} \binom{qv}{i} \pmod{pq}.$$

If either $\psi_1 = 1$ or $\psi_2 = 1$, then the sum contains only one term ($i = 0$ or $i = qv$) and the result follows. Hence, we can assume that $\psi_1 = \omega_{u_1}$ and $\psi_2 = \omega_{u_2}$ with u_1 and u_2 in $\{1, \ldots, q-1\}$ with $v = u_1 + u_2$ or $v = u_1 + u_2 - (q-1)$. If $\omega_i = \psi_1$ and $\omega_{qv-i} = \psi_2$, then there exists a non-negative integers k and l such that $i = u_1 + k(q-1)$ and $qv - i = u_2 + l(q-1)$. Hence, by the second part of Lemma 6,

$$\begin{aligned} \operatorname{ord}_p \binom{qv}{i} &= \frac{s_p(u_1 + k(q-1)) + s_p(u_2 + l(q-1)) - s_p(v)}{p-1} \\ &\geq \frac{s_p(u_1) + s_p(u_2) - s_p(v)}{p-1}. \end{aligned}$$

To summarize, we have shown that $\operatorname{ord}_p A^{\mathsf{F}}_{\chi,\psi}$ is at least

$$\frac{s_p(u_1) + s_p(u_2) - s_p(v)}{p-1} = \frac{\operatorname{wt}_p(\psi_1) + \operatorname{wt}_p(\psi_2) - \operatorname{wt}_p(\chi)}{p-1},$$

assuming that $\psi_1\psi_2 = \chi \neq 1$. □

With some additional effort, it can be shown that the inequality in Theorem 19 can in fact be replaced by an equality. If X is a field of prime order, then Theorem 19 reduces to the following result:

$$\operatorname{ord}_p A^{\mathsf{F}}_{\psi_1\psi_2,(\psi_1,\psi_2)} = \begin{cases} 1 & \text{if } u_1 + u_2 \geq p \\ 0 & \text{else}, \end{cases}$$

for $\psi_1 : x \mapsto \tau(x^{u_1})$ and $\psi_1 : x \mapsto \tau(x^{u_2})$.

Example 10. For $X = \mathbb{F}_2$, the ultrametric integral transition matrix of addition is

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & -2 \end{bmatrix}.$$

In this case, divisibility is trivial except in the case $\operatorname{wt}(\psi_1) = \operatorname{wt}(\psi_2) = 1$ For $X = \mathbb{F}_2^n$ the result is the n-fold tensor product of the above, which gives

$$A^{\mathsf{F}}_{\psi_1\psi_2,(\psi_1,\psi_2)} = (-2)^{\operatorname{wt}(u_1 \wedge u_2)}$$

for $\psi_1 : x \mapsto \tau(x^{u_1})$ and $\psi_2 : x \mapsto \tau(x^{u_2})$.

Example 11. For $X = \mathbb{F}_4 = \mathbb{F}_2(\alpha)$ with $\alpha^2 + \alpha + 1 = 0$, the ultrametric integral transition matrix of addition is

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & -\frac{4}{3} & 0 & 0 & \frac{2}{3} & 0 & 0 & -\frac{4}{3} & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & \frac{2}{3} & 0 & 0 & 1 & 0 & 0 & -\frac{4}{3} & 0 & 0 & -\frac{4}{3} & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & -\frac{1}{3} & 0 & 0 & -\frac{1}{3} & 0 & 0 & 1 & 0 & 0 & -\frac{4}{3} \end{bmatrix}.$$

The columns correspond to exponents $(0,0), (0,1), (0,2), (0,3), (1,0), (1,1), \ldots$ Note the difference with addition in the ring $\mathbb{F}_2^2$.

6.4. Addition with a constant. The formula for addition with a constant is similar to the formula for additions with two inputs. The result is listed in Theorem 20. The proof is omitted because it is essentially identical to the proof of Theorem 19. The result is stated for constants t that are nonzero in every coordinate, because for all other coordinates the corresponding characters must match to obtain a nonzero correlation.

THEOREM 20. *Let $\mathsf{F}\colon X \to X$ with $\mathsf{F}(x) = x + t$ for a constant t in X such that $t_1 \neq 0, \ldots, t_n \neq 0$. For all characters $\chi \neq 1$ and ψ of X, we have*

$$\operatorname{ord}_p A^{\mathsf{F}}_{\chi,\psi} = \frac{\operatorname{wt}_p(\rho) + \operatorname{wt}_p(\psi) - \operatorname{wt}_p(\chi)}{p-1}\,.$$

For $X = \mathbb{F}_q$, the character ρ is defined by $\rho(x) = \tau(x^{v-u})$ if $u \leq v$ and $\rho(x) = \tau(x^{q-1+v-u})$ otherwise, where $\psi(x) = \tau(x^u)$ and $\chi(x) = \tau(x^v)$ and u and v are in $\{0, 1, \ldots, q-1\}$. This extends to products of fields by tensoring, and hence to arbitrary X isomorphic to a product of fields of characteristic p.

EXAMPLE 12. For $X = \mathbb{F}_2$, the ultrametric integral transition matrix of $x \mapsto x + t$ is

$$\begin{bmatrix} 1 & 0 \\ \tau(t) & (-1)^t \end{bmatrix},$$

where $\tau(t)$ is just the representation of t as a $\{0, 1\}$ integer.

6.5. Jacobi sums. There is an alternative approach to proving Theorems 19 and 20 based on Jacobi sums. For example, let t be a nonzero constant in $\mathbb{F}_q$ and $\mathsf{F} : x \mapsto x + t$. If $\psi \neq 1$ and ψ is not the character $x \mapsto \tau(x^{q-1})$, then

$$A^{\mathsf{F}}_{\chi,\psi} = \frac{1}{q-1} \sum_{x \in \mathbb{F}_q^\times} \chi(x+t)/\psi(x) = \frac{\chi(t)}{\psi(-t)} \frac{1}{q-1} \sum_{x \in \mathbb{F}_q^\times} \chi(1-x)\psi^*(x)$$

The sum on the right is called a Jacobi sum and the bound in Theorem 20 is a known result about such sums. More precisely, Jacobi sums can be expressed in terms of Gauss sums, whose factorization into prime ideals (in an algebraic number field) is understood by Stickelberger's theorem. From the point of view of cryptanalysis, Gauss sums show up when converting statements about correlations matrices to statements about ultrametric integral transition matrices and conversely. In other words, they express the coordinates of a change-of-basis matrix.

Apart from the factorization of Gauss sums into prime ideals, there is also a more detailed formula for their exact value in the p-adic numbers due to Gross and Koblitz. This result can also be applied to prove Theorems 19 and 20. Without giving the details, the Gross-Koblitz formula involves the p-adic gamma function

$$\Gamma_p(n) = (-1)^n \prod_{\substack{1 \leq i < n \\ p \nmid i}} i\,.$$

This is well-defined for integers n and is p-adically continuous, so admits a unique extension to the p-adic integers. The exact values of the coordinates of the ultrametric integral transition matrix of translation can be expressed in terms of the

p-adic gamma function as follows:

$$A^F_{\chi,\psi} = (-1)^e \, \frac{(-p)^\nu}{1-q} \, \prod_{i=0}^{e-1} \frac{\Gamma_p\Big(\frac{\mathrm{wt}_q\, \chi^{p^i}}{1-q} + 1\Big)}{\Gamma_p\Big(\frac{\mathrm{wt}_q\, \psi^{p^i}}{1-q} + 1\Big)\Gamma_p\Big(\frac{\mathrm{wt}_q\, \rho^{p^i}}{1-q} + 1\Big)} \,,$$

for $F(x) = x + 1$. This can be proven using the Gross-Koblitz formula.

7. Trails

The principle of trails was already discussed in the first lecture. For ultrametric integral cryptanalysis, it corresponds to the formula

$$A^{\mathsf{F}}_{\chi_{r+1},\chi_1} = \sum_{\chi_2,\ldots,\chi_r} \prod_{i=1}^{r} A^{\mathsf{F}_i}_{\chi_{i+1},\chi_i} \,,$$

when $\mathsf{F} = \mathsf{F}_r \circ \cdots \circ \mathsf{F}_2 \circ \mathsf{F}_1$. In the context of ultrametric integral cryptanalysis, trails are mainly used to identify properties with (p-adically) small correlation. This can be understood as a relaxation of zero-correlation properties, and in fact the techniques used to find them are conceptually similar.

7.1. Principle of dominant trails. A straightforward but essential consequence of the ultrametric triangle inequality is the following 'exact' variant of the principle of dominant trails. As mentioned above, we will mainly use this result with $\Lambda = \emptyset$.

Theorem 21 (Dominant trails). *Let $\mathsf{F} = \mathsf{F}_r \circ \cdots \circ \mathsf{F}_2 \circ \mathsf{F}_1$. For all subsets Λ of the set Ω of all trails from χ_1 to χ_{r+1},*

$$\left| A^{\mathsf{F}}_{\chi_{r+1},\chi_1} - \sum_{\chi\in\Lambda} \prod_{i=1}^{r} A^{\mathsf{F}_i}_{\chi_{i+1},\chi_i} \right|_p \le \max_{\chi\in\Omega\setminus\Lambda} \left| \prod_{i=1}^{r} A^{\mathsf{F}_i}_{\chi_{i+1},\chi_i} \right|_p ,$$

where $\chi = (\chi_1, \chi_2, \ldots, \chi_{r+1})$.

Theorem 21 leads to the 'approximate miss-in-the-middle' principle: one constructs a set of trails in the forward and a backward directions (with correlation not too small), and shows that they cannot be matched with high correlation. This approach is convenient for the manual analysis of constructions.

7.2. Application to linear functions. Let $\mathsf{F} : X \to Y$ be a homomorphism of the additive groups of X and Y. Such a map can be decomposed as a composition of three layers:

(1) A layer of copy operations.
(2) A layer of multiplications with various constants.
(3) One or more layers of additions, each between two inputs.

Let us consider a trail $(\chi_1, \chi_2, \chi_3, \chi_4)$ with nonzero correlation through this network of three layers. For every copy operation, the weight of the output character (χ, ψ) is greater than the weight of the input character $\chi\psi$ because $\mathrm{wt}_p(\chi) + \mathrm{wt}_p(\psi) \ge \mathrm{wt}_p(\chi\psi)$. Hence $\mathrm{wt}_p(\chi_2) \ge \mathrm{wt}_p(\chi_1)$. Multiplications by constants affect only the sign of trails, so $\mathrm{wt}_p(\chi_3) \ge \mathrm{wt}_p(\chi_2)$. For an addition, the weight of the input characters (χ, ψ) can be higher than the weight of the output character $\chi\psi$, but Theorem 19 shows that if the weight is lower by Δ, then the correlation is divisible

by $\Delta/(p-1)$. It follows that the correlation of every trail through the addition layers is divisible by

$$p^{\lceil (\mathrm{wt}_p(\chi_3)-\mathrm{wt}_p(\chi_4))/(p-1)\rceil}.$$

Using the fact that $\mathrm{wt}_p(\chi_3) \geq \mathrm{wt}_p(\chi_2) \geq \mathrm{wt}_p(\chi_1)$ shows that

$$\mathrm{ord}_p\, A^{\mathsf{F}}_{\chi_4,\chi_1} \geq \left\lceil \frac{\mathrm{wt}_p(\chi_1)-\mathrm{wt}_p(\chi_4)}{p-1} \right\rceil,$$

based on the principle of dominant trails. After reduction to the residue field, this result says that the image of a set with division property of order k under a linear map still has the division property of order k. In the next lecture, it will be shown that more complicated constructions can be analyzed in essentially the same way.

LECTURE 3

Applications

1. Introduction

The basics of ultrametric integral cryptanalysis were developed in the previous two lectures. This lecture turns to applications, starting with the analysis of arbitrary functions of a given degree. Cryptanalytic proofs of the Ax-Katz and Moreno-Moreno theorems will be given. These theorems show that the number of solutions to a system of polynomial equations over a finite field of characteristic p is divisible by some power of p depending on the degree of the equations. Further applications include the analysis of Feistel ciphers. This lecture assumes familiarity with the contents of the first two lectures, and with basic cryptographic constructions (Feistel ciphers, some concrete primitives) and their integral cryptanalysis in the $\mathbb{F}_2^n$-case.

2. Input sets and data-complexity

In the applications that follow, we will obtain 'almost zero-correlation' approximations for various constructions. That is, we will find pairs (ψ, χ) such that $|A^{\mathsf{F}}_{\chi,\psi}|_p \leq \varepsilon$ for some $0 < \varepsilon < 1$. More generally, we may suppose a one-dimensional property (u, v) with evaluation

$$\left|v\big(A^{\mathsf{F}}\, u\big)\right|_p \leq \varepsilon\,.$$

In practice, we then want to know the minimum amount of data required to exploit this property. For one-dimensional properties such as (u, v), at most $|\operatorname{supp} \mathscr{U}^{-1}\, u|$ data is required. This is tight when all nonzero coordinates of $\mathscr{U}^{-1}\, u$ have the same p-adic absolute value, as is the case for $u = \delta_\psi$. If ψ extends a character of the maximal subgroup X_e, then the explicit formula for b_ψ from Lecture 1 shows that $\operatorname{supp} b_\psi = \bigcup_{f \leq e} X_f$, where f ranges over all idempotents below e.

Example 13. For $X = \mathbb{F}_q^n$, the amount of data required to exploit a property (b_ψ, b^χ) with ψ extending a character of $e\,\mathbb{F}_q^n$ for some idempotent e, is equal to

$$|\operatorname{supp} b_\psi| = \sum_{f \leq e} (q-1)^{\mathrm{wt}(f)} = q^{\mathrm{wt}(e)} = q^{\mathrm{wt}(\psi)}\,.$$

Example 13 shows that, if we only use a single approximation, then the data-complexity is determined by the weight of ψ — although in general, differences in the size of the factors of X should be accounted for. However, when multiple approximations are available, the data complexity can sometimes be reduced significantly. This is analogous (in fact, as discussed below, in a precise way) to data-complexity reductions for multidimensional zero-correlation linear cryptanalysis.

Completely determining the minimum data-complexity for multiple approximations would lead us too far from the applications in this lecture, but it is useful to discuss how multiple approximations are related to divisibility properties for certain structured input sets. By analogy with multidimensional linear cryptanalysis, we will consider submonoids of X. As explained in the first lecture, if $M \subset X$ is a submonoid, then the ultrametric integral change-of-basis of the subspace of $k[X]$ spanned by the elements of M is equal to a subspace spanned by the indicators of certain equivalence classes of characters:

$$\operatorname{Span}\Big\{\delta_x \mid x \in M\Big\} = \operatorname{Span}\Big\{b_{[\chi]} \mid [\chi] \in \widehat{X}/\sim\Big\}\,,$$

where the equivalence relation $\sim$ is defined by $\chi \sim \psi$ if and only if $\chi(x) = \psi(x)$ for all x in M. More explicitly, we have

$$\delta_x = \sum_{[\chi] \in \widehat{X}/\sim} \chi(x)\, b_{[\chi]}\,.$$

This in particular yields the following formula.

LEMMA 9. *Let M be a submonoid of X and $\sim$ the equivalence relation defined above. The coordinates of $\mathscr{U}\,\delta_M$ are constant on the equivalence classes $\widehat{X}/\sim$. Furthermore, if χ is a character of X extending a character χ_e of the group X_e, then*

$$\frac{1}{|[\chi]|}\, b^{\chi}\big(\,\delta_M\,\big) = \begin{cases} \sum_{f \geq e} |M_f| & \text{if } \chi_e \in M_e^1\,, \\ 0 & \text{else}\,, \end{cases}$$

where the sum is over the idempotent elements f of M and $M_e = M \cap X_f$.

PROOF. A submonoid of M consists of a sublattice $\mathcal{E}_M$ of $\mathcal{E}_X$, with for each f in $\mathcal{E}_M$, a subgroup M_f of X_f. If $f \geq e$, then there is a connecting group homomorphism $M_f \to M_e$ defined by $x \mapsto ex$ with kernel of order $|M_f|/|M_e|$. Hence,

$$\sum_{x \in M} \chi(x) = \sum_{f \geq e} \sum_{x \in M_f} \chi_e(ex) = \sum_{f \geq e} \frac{|M_f|}{|M_e|} \sum_{x \in M_e} \chi_e(ex)\,.$$

By the orthogonality of group characters, the inner sum is zero unless $\chi_e \in M_e^1$. □

In the following example, the equivalence relation on the dual of X is trivial, but it helps to illustrate how Lemma 9 can be used.

EXAMPLE 14. Take $M = X = \mathbb{F}_q$. The equivalence relation on the characters of X is given by $\chi \sim \psi$ if and only if $\chi = \psi$. Lemma 9 shows that

$$b^{\chi}\big(\delta_M\big) = \begin{cases} q & \text{if } \chi = 1\,, \\ q-1 & \text{if } \chi(x) = \tau(x^{q-1})\,, \\ 0 & \text{else}\,. \end{cases}$$

In Section 3, we will use the following consequence of this expression:

$$\delta_{\mathbb{F}_q^n} = \sum_{\chi} q^{n-\mathrm{wt}(\chi)}\,(q-1)^{\mathrm{wt}(\chi)}\, b_{\chi}\,,$$

where the sum is over all idempotent characters. This follows by taking tensor products, or directly from Lemma 9 with $M = X = \mathbb{F}_q$.

In integral cryptanalysis over $\mathbb{F}_2$, the input sets are usually chosen to be of the form $u\mathbb{F}_2^n$, since these correspond to the precursor basis vectors. The following example shows that if such sets are used for ultrametric integral cryptanalysis, then one also needs to take into account characters with weight greater than $\mathrm{wt}(u)$. However, the higher the weight, the smaller the contribution of the character.

EXAMPLE 15. Let $X = \mathbb{F}_q^n$ and $M = u\mathbb{F}_q^n$ with u an idempotent element. Let $\chi = (\chi_1, \ldots, \chi_n)$ and $\psi = (\psi_1, \ldots, \psi_n)$ be characters of $\mathbb{F}_q^n$. The equivalence relation corresponds to $\chi \sim \psi$ if and only if $\chi_i = \psi_i$ when $u_i = 1$ and $\chi_i^{q-1} = \psi_i^{q-1}$

when $u_i = 0$. In other words, on the positions i with $u_i = 0$, only the idempotent power of the character χ_i matters. Lemma 9 yields

$$\delta_M = \sum_{\chi \geq \psi_u} q^{\mathrm{wt}(u)-\mathrm{wt}(\chi)}(q-1)^{\mathrm{wt}(\chi)}\, b_\chi\,,$$

where the sum is over idempotent characters greater than the idempotent character $\psi_u \colon x \mapsto \prod_{i=1}^n \tau(x^{u_i(q-1)})$. Indeed,

$$\sum_{f \geq e} |M_f| = \sum_{f \geq e} (q-1)^{\mathrm{wt}(f)} = q^{\mathrm{wt}(u)-\mathrm{wt}(e)}\,(q-1)^{\mathrm{wt}(e)}.$$

This result can also be deduced from the isomorphism $u\mathbb{F}_q^n \cong \mathbb{F}_q^{\mathrm{wt}(u)}$.

As a warning, note that the lattice of idempotents of a submonoid is not necessarily a Boolean lattice and submonoids are not necessarily subrings. For example, $\mathbb{F}_p$ does not contain any proper subfields but when $p-1$ is smooth it has many submonoids corresponding to subgroups of the cyclic group $\mathbb{F}_p^\times$.

Example 16. Let $X = \mathbb{F}_3^2$ and $M = \{0,1\}^2$ the submonoid of squares. The equivalence relation on characters of $\mathbb{F}_3$ is defined by the partition $\{\{x \mapsto \tau(x), x \mapsto \tau(x^2)\}, \{x \mapsto 1\}\}$. Lemma 9 gives

$$\begin{aligned}\delta_M &= \left(2\, b_{x\mapsto 1} + b_{x\mapsto\tau(x)} + b_{x\mapsto\tau(x^2)}\right)^{\otimes 2} \\ &= 4\, b_{x\mapsto 1} + 2\left(b_{x\mapsto\tau(x_2)} + b_{x\mapsto\tau(x_2^2)}\right) + 2\left(b_{x\mapsto\tau(x_1)} + b_{x\mapsto\tau(x_1^2)}\right) \\ &\quad + \left(b_{x\mapsto\tau(x_1x_2)} + b_{x\mapsto\tau(x_1^2x_2)} + b_{x\mapsto\tau(x_1x_2^2)} + b_{x\mapsto\tau(x_1^2x_2^2)}\right),\end{aligned}$$

where the terms between each pair of parentheses correspond to equivalent characters. In particular, the above expression can also be written as $\delta_M = 4b_{[x\mapsto 1]} + 2b_{[x\mapsto\tau(x_1)]} + 2b_{[x\mapsto\tau(x_2)]} + b_{[x\mapsto\tau(x_1x_2)]}$. The support is full in this case, since the annihilator of $\{1\} \subset \mathbb{F}_3^\times$ is the group of all characters of $\mathbb{F}_3^\times$.

The results about multidimensional properties from Lecture 1 also imply that if the output characters form a submonoid of the dual of X, then one obtains a divisibility property for a 'projection function' (in the sense of Wagner) to a set $X/\sim$. This is useful for other reasons but does not help reduce the data-complexity, except when divisibility is so high that statistical methods become available.

3. Functions with prescribed degree

In the following, $\mathsf{F} : X \to Y$ will be a function with prescribed degree. As we will see, there is more than one notion of 'degree' that is relevant to the analysis of such functions. For the Ax-Katz theorem discussed below, traditionally $X = \mathbb{F}_q^n$ and $Y = \mathbb{F}_q^m$, but the result generalizes to other products of fields of common characteristic p. By cryptanalytic standards, arbitrary functions of prescribed degree are not particularly interesting because they do not provide a specific enough model for most cryptographic primitives. However, there are several reasons why discussing this class of functions is particularly important in the context of ultrametric integral cryptanalysis:

- In many cases, it is reasonably accurate to model a component of a larger construction (such as a Feistel cipher or an AES-like cipher) by an arbitrary function with the same degree.

- It is often useful to compare properties of a concrete function with their counterparts for a generic function of the same degree.

Nevertheless, the primary motivation in this section is the proof of some theorems that are of interest to pure mathematics (the Ax-Katz theorem and the Moreno-Moreno theorem). The proof will be based on the analysis of ultrametric integral trails in a generic function of given degree.

3.1. Ax-Katz theorem. The Ax-Katz theorem has a rather long history, going back to a theorem of Chevalley and Warning that states that the number of solutions to a system of equations of degrees $d_1, \ldots, d_m$ in n variables over a finite field of characteristic p, is divisible by p if $d_1 + \ldots + d_m < n$. This result is not particularly deep or difficult[1], but it has an interesting consequence: a homogeneous system of equations in n variables has a nontrivial solution whenever $d_1 + \ldots + d_m < n$. Ax, and later Katz in a more general form, generalized the Chevalley-Warning theorem as follows.

THEOREM 22 (Ax-Katz). *The number of solutions to a system of equations $f_1, \ldots, f_m$ in n variables over $\mathbb{F}_q$ satisfies*

$$\operatorname{ord}_q N(f_1, \ldots, f_m) \geq \left\lceil \frac{n - \sum_{i=1}^m \deg f_i}{\max_{1 \leq i \leq n} \deg f_i} \right\rceil .$$

In particular, the number of solutions to a system of m equations of degree d in n variables is divisible by $q^{\lceil n/d \rceil - m}$.

The Ax-Katz theorem can be understood as a p-adic analogue of the Riemann hypothesis for local zeta functions, as it equivalent to an upper bound on the p-adic absolute value of the poles and zeros of the local zeta function of the variety cut out by the given equations.

A variant of the Ax-Katz theorem that is sometimes more precise was proven by Moreno and Moreno. Their bound depends on the p-degree of a polyonomial, which for a monomial is defined by

$$\deg_p x_1^{i_1} \cdots x_n^{i_n} = \sum_{k=1}^n s_p(i_k) .$$

The proof is based on rewriting the system of equations as a system of equations over $\mathbb{F}_p$, and $\deg_p f$ is precisely the algebraic degree of f as a polynomial over $\mathbb{F}_p$.

THEOREM 23 (Moreno-Moreno). *The number of solutions to a system of equations $f_1, \ldots, f_m$ in n variables over $\mathbb{F}_q$ with $q = p^e$ satisfies*

$$\operatorname{ord}_p N(f_1, \ldots, f_m) \geq \left\lceil e \left(\frac{n - \sum_{i=1}^m \deg_p f_i}{\max_{1 \leq i \leq n} \deg_p f_i} \right) \right\rceil .$$

In particular, the number of solutions to a system of m equations of p-degree d in n variables is divisible by $p^{\lceil en/d \rceil - em}$.

As discussed before, the class of functions with prescribed degree is too broad, so Theorems 22 and 23 are results with few immediate cryptanalytic applications.

[1]It follows from the fact that $\sum_{x \in \mathbb{F}_q} x^i = 0$ whenever $i < q - 1$.

3.2. Ultrametric integral cryptanalysis and systems of equations. To prove the Ax-Katz theorem using ultrametric integral cryptanalysis, we first express it in terms of the properties of a function. The solutions of the system of equations are the preimages of 0 for the function $\mathsf{F} : \mathbb{F}_q^n \to \mathbb{F}_q^m$ with $\mathsf{F}_i = f_i$. Since adding a constant does not affect the degree, this means the Ax-Katz theorem can be rephrased as a result about the number of preimages of a function with prescribed degree. This can be expressed as a cryptanalytic property, and hence in terms of ultrametric integral approximations, as made explicit by the following lemma.

LEMMA 10. *The number of preimages N of $(0, 0, \ldots, 0)$ under $\mathsf{F} : \mathbb{F}_q^n \to \mathbb{F}_q^m$ satisfies*

$$|N|_p \leq \max_{\psi,\chi} q^{\mathrm{wt}(\psi)-n} \left|A^{\mathsf{F}}_{\chi,\psi}\right|_p ,$$

where the maximum is over all idempotent multiplicative characters of $\mathbb{F}_q^n$.

PROOF. The number of preimages N is equal to

$$N = \delta^0\left(T^{\mathsf{F}}\, \delta_{\mathbb{F}_q^n}\right) = \sum_{\chi} (-1)^{\mathrm{wt}(\chi)}\, \delta^{\chi}\left(A^{\mathsf{F}}\ \mathscr{U}\ \delta_{\mathbb{F}_q^n}\right) ,$$

where the sum is over all idempotent characters χ. By the result from Example 14,

$$\mathscr{U}\ \delta_{\mathbb{F}_q^n} = \sum_{\psi} q^{n-\mathrm{wt}(\psi)}\, (q-1)^{\mathrm{wt}(\psi)}\, \delta_{\psi} ,$$

where the sum is over all *idempotent* characters ψ. Hence, $\mathrm{wt}(\psi)$ is the number of ψ_i equal to $x \mapsto \tau(x^{q-1})$ when $\psi = (\psi_1, \ldots, \psi_n)$. It follows that the number of preimages is

$$N = \sum_{\chi,\psi} (-1)^{\mathrm{wt}(\chi)}\, q^{n-\mathrm{wt}(\psi)}\, (q-1)^{\mathrm{wt}(\psi)}\, A^{\mathsf{F}}_{\chi,\psi} .$$

The result follows by taking the p-adic absolute value and applying the ultrametric triangle inequality. □

Based on Lemma 10, the proof of the Ax-Katz and Moreno-Moreno theorems should fall out of a lower bound on $\mathrm{ord}_p\, A^{\mathsf{F}}_{\chi,\psi}$. In fact, it is enough to have such a bound for idempotent characters χ and ψ — but for cryptanalytic applications, it is useful to have a bound for all coordinates anyway.

3.3. Analysis of trails using p-weights. The following theorem gives generically tight upper bounds on the p-adic absolute values of the coordinates of A^{F}. It will be shown below that Theorem 23 follows from this result in the case with a uniform bound on the p-degree of the functions $f_1, \ldots, f_m$. The more detailed case can be proven in exactly the same way, but this will be left to the reader. Theorem 22 follows from a different upper bound that requires a separate analysis and is discussed in the next section.

THEOREM 24. *Let $\mathsf{F} : \mathbb{F}_q^n \to \mathbb{F}_q^m$ be a function of p-degree d. For all multiplicative characters ψ and χ, it holds that*

$$\mathrm{ord}_p\, A^{\mathsf{F}}_{\chi,\psi} \geq \left\lceil \frac{\mathrm{wt}_p(\psi)/d - \mathrm{wt}_p(\chi)}{p-1} \right\rceil .$$

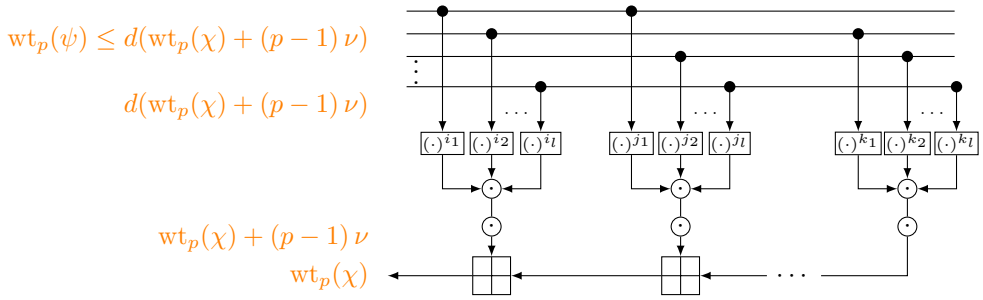

FIGURE 1. Ultrametric integral trails for a function with p-degree d.

PROOF. Each of the functions F_i can be written as a sum of one or more monomials with p-degree d, each scaled by some constant coefficient. This construction is depicted in Figure 1. Note that the constant term can be realized by adding a degree-zero monomial.

To estimate the correlation $A^{\mathsf{F}}_{\chi,\psi}$ of the approximation (ψ, χ), we will analyze trails through m copies of the construction in Figure 1. The basic strategy will be to upper bound the maximum p-weight of a character at the beginning of each trail with correlation $p^{-\nu}$. Since the upper bounds that we derive are linear in the p-weight of the output character, it is sufficient to analyze one of the functions F_i that is of highest degree.

As discussed in the second lecture, if the total p-weight of the input characters of an addition exceeds the p-weight of the output character by $\nu_i(p-1)$, then the corresponding approximation has absolute correlation $p^{-\nu_i}$. This holds for every addition, so the p-weight at the input of the addition layer is at most $\mathrm{wt}_p(\chi) + (p-1)\,\nu_1 + (p-1)\,\nu_2 + \cdots \le \mathrm{wt}_p(\chi) + (p-1)\,\nu$. Multiplication by a constant does not affect the p-weight, so the weight at the output of the layer of monomials is at most $\mathrm{wt}_p(\chi) + (p-1)\,\nu$. Every monomial function is a monoid homomorphism and, as we saw in the second lecture, increases the p-weight from output to input character by a factor of at most d. Hence, the p-weight of any character at the input of the layer of monomials in a trail is at most $d(\mathrm{wt}_p(\chi) + (p-1)\,\nu)$. Since the p-weight cannot increase further under copy operations, every trail with absolute correlation $p^{-\nu}$ or higher must satisfy

$$\mathrm{wt}_p(\psi) \le d\left(\mathrm{wt}_p(\chi) + (p-1)\,\nu\right).$$

Rearranging to isolate ν, which is an integer, yields

$$\nu \ge \left\lceil \frac{\mathrm{wt}_p(\psi)/d - \mathrm{wt}_p(\chi)}{p-1} \right\rceil.$$

This holds for every trail, so by the principle of dominant trails $\mathrm{ord}_p\, A^{\mathsf{F}}_{\chi,\psi} \ge \nu$. □

Proving the Moreno-Moreno theorem (Theorem 23) is now a matter of applying Lemma 10 to the result of Theorem 24. In fact, Theorem 24 itself is more important for applications to cryptanalysis. For example, it will be useful in Section 4.

PROOF OF THEOREM 23 WITH EQUAL DEGREES. Applying Lemma 10 to Theorem 24 shows that the number of solutions N satisfies (the minimum is over all

idempotent characters, and achieved for $\chi\colon x \mapsto \tau(\prod_{i=1}^m x_i^{q-1}))$

$$\begin{aligned}\operatorname{ord}_p N &\geq \min_{\psi} e\big(n - \operatorname{wt}(\psi)\big) + \left\lceil \frac{\operatorname{wt}_p(\psi)/d - me(p-1)}{p-1} \right\rceil \\ &\geq \min_{\psi} e\big(n - m - \operatorname{wt}(\psi)\big) + \left\lceil \frac{\operatorname{wt}_p(\psi)/d}{p-1} \right\rceil ,\end{aligned}$$

with $q = p^e$. The character ψ is idempotent, so $\operatorname{wt}_p(\psi) = e\,(p-1)\operatorname{wt}(\psi)$. The minimum is achieved for $\operatorname{wt}(\psi) = n$ and hence

$$\operatorname{ord}_p N \geq \left\lceil \frac{en}{d} \right\rceil - em\,.$$

This is precisely the Moreno-Moreno theorem with equal degrees. □

3.4. Analysis of trails using q-weights. The bound in Theorem 23 is best possible for functions of fixed p-degree, but not for all functions of fixed polynomial degree. The issue is that, in the proof of Theorem 24, we only kept track of the p-weight of characters in trails. However, this may be suboptimal for a function with given polynomial degree d. Below, we give another formula for $\operatorname{ord}_p A^{\mathsf{F}}_{\chi,\psi}$ in terms of the degrees of characters ψ and χ. The *degree of a character* $\chi : x \mapsto \tau(x^v)$ with $v_1, \ldots, v_n$ in $\{0, 1, \ldots, q-1\}$ is equal to $\sum_{i=1}^n v_i$. By analogy with the p-weight of characters, we will instead use the term 'q-weight':

$$\operatorname{wt}_q(\chi) = \sum_{i=1}^n s_q(v_i)\,,$$

where $s_q(v_i) = v_i$ since $0 \leq v_i < q$. A similar definition can be given for other powers of p, and would lead to further variants of the Ax-Katz theorem — but we will not go as far in these lecture notes.

By keeping track of q-weights of characters rather than p-weights, we will prove the following variant of Theorem 24. The proof follows the same strategy, but is slightly more technical because we will need to keep track of not only the q-weight of the characters, but also the q-weight of their p^{th} powers.

THEOREM 25. *Let $\mathsf{F} : \mathbb{F}_q^n \to \mathbb{F}_q^m$ be a function of degree d. For all multiplicative characters ψ and χ, it holds that*

$$\operatorname{ord}_p A^{\mathsf{F}}_{\chi,\psi} \geq \sum_{i=0}^{e-1} \left\lceil \frac{\operatorname{wt}_q(\psi^{p^i})/d - \operatorname{wt}_q(\chi^{p^i})}{q-1} \right\rceil ,$$

for $q = p^e$.

The reason why Theorem 25 involves p^{th} powers of χ and ψ is related to the ultrametric integral properties of addition. Let $\psi_1 : x \mapsto \tau(x^{u_1})$, $\psi_2 : x \mapsto \tau(x^{u_2})$ and $\chi = \psi_1\psi_2 : x \mapsto \tau(x^v)$ be characters of $\mathbb{F}_q$. As shown in the previous lecture,

$$A^{\boxplus}_{\chi,(\psi_1,\psi_2)} = \frac{\operatorname{wt}_p(\psi_1) + \operatorname{wt}_p(\psi_2) - \operatorname{wt}_p(\chi)}{p-1}\,,$$

and this is equal to the number of carries when adding u_1 and u_2 in base p. If there is overflow, then we also count the number of carries introduced by the overflow[2] after reducing modulo $q-1$. It is possible to use the q-weight of characters to detect one of these carries: if there is no carry in the addition of the highest digits, then

[2]Since u_1 and u_2 are both less than $q-1$, this introduces at most one additional carry.

$v = u_1 + u_2$. Otherwise, if there is a carry in this position, then $v = u_1 + u_2 - (q-1)$. Equivalently, $u_1 + u_2 - v \leq \nu_0(q-1)$ with ν_0 in $\{0, 1\}$ indicating the occurrence of a carry. Since taking a p^{th} power of a character rotates the digits in its exponent to the left, we have

$$\mathrm{wt}_q(\psi_1^{p^i}) + \mathrm{wt}_q(\psi_2^{p^i}) - \mathrm{wt}_q(\chi^{p^i}) \leq \nu_i\,(q-1)\,,$$

where $\nu_i = 1$ if there is a carry when adding the digits in position i and $\nu_i = 0$ otherwise. Hence, the total number of carries ν satisfies $\nu = \nu_0 + \nu_1 + \ldots + \nu_{e-1}$. This provides an alternative characterization of the properties of addition, this time in terms of q-weights.

Proof of Theorem 25. Like the proof of Theorem 24, we analyze the ultrametric integral trails through m copies of the construction shown in Figure 2. The addition layer may also include an addition by a constant. Like in the proof

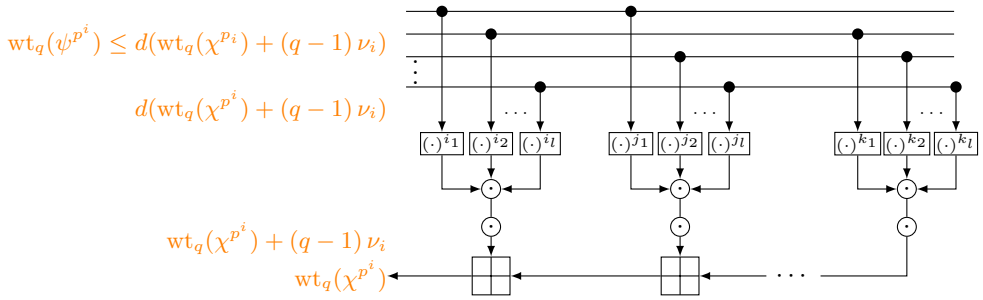

Figure 2. Ultrametric integral trails for a function with degree d.

of Theorem 23, it is sufficient to analyze one of the functions F_i that is of highest degree. The strategy is to keep track of the q-weight of the $p^{i\ \text{th}}$ power of the characters in a trail, for each i in $\{0, 1, \ldots, e-1\}$.

For the addition layer, the q-weight of the $p^{i\ \text{th}}$ power of the input character can be larger by $(q-1)\,\nu_i$, for some integer ν_i (for one addition, ν_i is either zero or one). That is, the q-weight at the output of the layer of monomials is at most

$$\mathrm{wt}_q(\chi^{p^i}) + (q-1)\,\nu_i\,.$$

Every degree d monomial is a morphism of monoids, and increases the q-weight by at most a factor of d (from output to input). The q-weight of the input character of a copy layer is always greater than the q-weight of the output character, so

$$\mathrm{wt}_q(\psi^{p^i}) \leq d\,(\mathrm{wt}_q(\chi^{p^i}) + (q-1)\,\nu_i)\,.$$

The absolute correlation of the trail is then at most $p^{-\nu}$, where $\nu = \nu_0 + \nu_1 + \cdots + \nu_{e-1}$. Rearranging the inequality above yields

$$\nu_i \geq \left\lceil \frac{\mathrm{wt}_q(\psi^{p^i})/d - \mathrm{wt}_q(\chi^{p^i})}{q-1} \right\rceil .$$

Summing this inequality for $i = 0, 1, \ldots, e-1$ yields the result. Note that it is quite important that we rearrange before computing the sum, because we want to account for the fact that each ν_i is an integer. □

To prove Theorem 22, we can apply Lemma 10 to Theorem 25 as follows.

Proof of Theorem 22 with equal degrees. By Lemma 10 and Theorem 25, the number of solutions N satisfies (the minimum is achieved for $\chi\colon x \mapsto \tau(\prod_{i=1}^{m} x_i^{q-1})$)

$$\begin{aligned}\operatorname{ord}_p N &\geq \min_{\psi} e\big(n - \operatorname{wt}(\psi)\big) + \sum_{i=0}^{e-1} \left\lceil \frac{\operatorname{wt}_q(\psi^{p^i})/d - m(q-1)}{q-1} \right\rceil \\ &\geq \min_{\psi} e\big(n - m - \operatorname{wt}(\psi)\big) + e \left\lceil \frac{\operatorname{wt}(\psi)}{d} \right\rceil ,\end{aligned}$$

where $\psi^{p^i} = \psi$ and $\operatorname{wt}_q(\psi) = (q-1)\operatorname{wt}(\psi)$ because the minimum is over all idempotent characters ψ. The minimum is achieved for $\operatorname{wt}(\psi) = n$. Hence,

$$\operatorname{ord}_q N \geq \left\lceil \frac{n}{d} \right\rceil - m\,.$$

This is precisely the Ax-Katz theorem with equal degrees. □

The proof of Theorem 25 (and Theorem 22 as a corollary) is a good example of the fact that it is often necessary to keep track of more detailed information than just the p-weights of the characters in a trail.

4. Feistel ciphers over $\mathbb{F}_q$

As a case study, we analyze Feistel ciphers on $\mathbb{F}_q^2$. The analysis will not be complete: for the most part, we restrict to generic Feistel constructions with a round function of prescribed p-degree and p a small prime number. For other cases, the general principles are the same but additional work is required. For tight bounds on concrete constructions, it is difficult to avoid computer assistance.

4.1. Analysis of trails. Figure 3 depicts a single round $\mathsf{F}\colon \mathbb{F}_q^2 \to \mathbb{F}_q^2$ of a Feistel cipher with round function $\mathsf{G}\colon \mathbb{F}_q \to \mathbb{F}_q$. As already mentioned, we focus here on the case that both branches are elements of the same finite field $\mathbb{F}_q$. The analysis of the more general case is similar, although the results can be quite different.

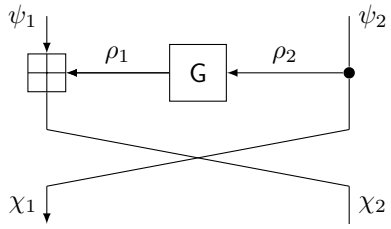

Figure 3. An ultrametric integral trail through one Feistel round.

Let (ψ, χ) be an approximation of F with $\psi = (\psi_1, \psi_2)$ and $\chi = (\chi_1, \chi_2)$ multiplicative characters. Based on the properties of addition and copy operations, the characters ρ_1 and ρ_2 shown in Figure 3 must satisfy

$$\begin{aligned}\chi_2 &= \psi_1 \rho_1\,, \\ \psi_2 &= \chi_1 \rho_2\,.\end{aligned}$$

In most cases, these relations uniquely determine ρ_1 and ρ_2. Suppose that $\chi_2 \neq 1$ and $\psi_2 \neq 1$. If in addition $\psi_1 \neq \chi_2$ and $\psi_2 \neq \chi_1$, then $\rho_1 = \psi_1^* \chi_2$ and $\rho_2 = \chi_1^* \psi_2$ with ψ_1^* the (pseudo)inverse of ψ_1 and likewise for χ_1^*. Hence,

$$\operatorname{ord}_p A^{\mathsf{F}}_{\chi,\psi} = \frac{\operatorname{wt}_p(\psi_1) + \operatorname{wt}_p(\psi_1^* \chi_2) - \operatorname{wt}_p(\chi_2)}{p-1} + \operatorname{ord}_p A^{\mathsf{G}}_{\psi_1^* \chi_2,\, \chi_1^* \psi_2} .$$

However, if $\psi_1 = \chi_2$ or $\psi_2 = \chi_1$, then a separate analysis is necessary because in this case ρ_1 or ρ_2 may be the trivial character. In particular, if $\psi_1 = \chi_2$, then there could be an additional trail with nonzero correlation for $\rho_1 = 1$. However, if $\rho_1 = 1$, then also $\rho_2 = 1$. This is possible only if $\psi_2 = \chi_1$. Hence, if both $\psi_1 = \chi_2$ and $\psi_2 = \chi_1$, then (since the absolute correlation of the trail with $\rho_1 = 1$ is greatest)

$$\operatorname{ord}_p A^{\mathsf{F}}_{\chi,\psi} = \frac{\operatorname{wt}_p(\psi_1) - \operatorname{wt}_p(\chi_2)}{p-1} = 0 .$$

If $\psi_2 = \chi_1$ but $\psi_1 \neq \chi_2$, then $\rho_2 = 1$ but $\rho_2 \neq 1$ so the bound becomes

$$\operatorname{ord}_p A^{\mathsf{F}}_{\chi,\psi} \geq \frac{\operatorname{wt}_p(\psi_1) + \operatorname{wt}_p(\psi_1^* \chi_2) - \operatorname{wt}_p(\chi_2)}{p-1} + \min\left\{ \operatorname{ord}_p A^{\mathsf{G}}_{\psi_1^* \chi_2,\, 1}, \operatorname{ord}_p A^{\mathsf{G}}_{\psi_1^* \chi_2,\, \chi_1^* \psi_2} \right\} .$$

Assuming the two trails do not cancel, the above inequality becomes an equality if

$$\operatorname{ord}_p A^{\mathsf{G}}_{\psi_1^* \chi_2,\, 1} \geq \operatorname{ord}_p A^{\mathsf{G}}_{\psi_1^* \chi_2,\, \chi_1^* \psi_2} ,$$

with $\chi_1^* \psi_2 \colon x \mapsto \tau(x^{q-1})$ in this case. These two conditions are generically satisfied.

Finally, we consider the case that $\chi_2 = 1$ or $\psi_2 = 1$. The formulae from above are still valid, except that some zero-correlation cases must be ruled out. In particular, if $\psi_2 = 1$, then we must also have $\chi_1 = 1$. Similarly, if $\chi_2 = 1$, then we must have $\psi_1 = 1$ and $\rho_1 = 1$. The latter condition implies that $\psi_2 = \chi_1$. To summarize, we have the following result.

THEOREM 26. *Let $\mathsf{F} \colon \mathbb{F}_q^2 \to \mathbb{F}_q^2$ be one round of a Feistel cipher with round function $\mathsf{G} \colon \mathbb{F}_q \to \mathbb{F}_q$ and let (ψ, χ) be a pair of multiplicative characters of $\mathbb{F}_q^2$ such that $\psi = (\psi_1, \psi_2)$ and $\chi = (\chi_1, \chi_2)$. If $\psi_2 = 1 \wedge \chi_1 \neq 1$ or $\chi_2 = 1 \wedge (\psi_1 \neq 1 \vee \psi_2 \neq \chi_1)$, then $A^{\mathsf{F}}_{\chi,\psi} = 0$. Otherwise, it holds that*

$$\operatorname{ord}_p A^{\mathsf{F}}_{\chi,\psi} \geq \begin{cases} 0 & \text{if } \psi_2 = \chi_1 \text{ and } \psi_1 = \chi_2 , \\ \nu + \min\left\{ \operatorname{ord}_p A^{\mathsf{G}}_{\psi_1^* \chi_2,\, 1}, \operatorname{ord}_p A^{\mathsf{G}}_{\psi_1^* \chi_2,\, \chi_1^* \psi_2} \right\} & \text{if } \psi_2 = \chi_1 \text{ and } \psi_1 \neq \chi_2 , \\ \nu + \operatorname{ord}_p A^{\mathsf{G}}_{\psi_1^* \chi_2,\, \chi_1^* \psi_2} & \text{else} , \end{cases}$$

where ν is given by

$$\nu = \frac{\operatorname{wt}_p(\psi_1) + \operatorname{wt}_p(\psi_1^* \chi_2) - \operatorname{wt}_p(\chi_2)}{p-1} .$$

Theorem 26 can be used to automate the ultrametric integral analysis of Feistel ciphers, provided that one has a model for the round function G. Possible strategies for representing this as an MILP, SAT or SMT problem will not be discussed here, and we focus instead on the generic case. An example of a concrete construction is discussed in Section 4.5.

4.2. Generic case using p-weights. Theorem 26 provides a model for one round of a Feistel cipher, but except using significant computer-assistance, it is not so easy to use. If $\mathsf{G}\colon \mathbb{F}_q \to \mathbb{F}_q$ is any function with p-degree d, then we might hope to upper bound $|A^{\mathsf{F}}_{\chi,\psi}|_p$ entirely in terms of the p-weights of the characters χ_1, χ_2, ψ_1 and ψ_2. By Theorem 24,

$$\operatorname{ord}_p A^{\mathsf{G}}_{\rho_1,\rho_2} \geq \max\left\{0, \left\lceil \frac{\operatorname{wt}_p(\rho_2)/d - \operatorname{wt}_p(\rho_1)}{p-1} \right\rceil \right\}. \tag{3}$$

To express everything in terms of p-weights, it is sufficient to determine the minimum of $\operatorname{wt}_p(\lambda_1^*\lambda_2)$ over all characters λ_1 and λ_2 with prescribed p-weight. Indeed, the valuation bound is minimized when $\operatorname{wt}_p(\chi_1^*\psi_2)$ and $\operatorname{wt}_p(\psi_1^*\chi_2)$ are minimal. In the former case, this follows from the fact that the bound in Theorem 26 is increasing in $\operatorname{wt}_p(\chi_1^*\psi_2)$. In the latter case, ν is increasing in $\operatorname{wt}_p(\psi_1^*\chi_2)$, but $\operatorname{ord}_p A^{\mathsf{G}}_{\psi_1^*\chi_2,\chi_1^*\psi_2}$ is not. However, an increase in $\operatorname{wt}_p(\psi_1^*\chi_2)$ is guaranteed to increase the integer ν but due to the ceiling function $\operatorname{ord}_p A^{\mathsf{G}}_{\psi_1^*\chi_2,\chi_1^*\psi_2}$ can increase at most by the same amount. Hence, the overall lower bound is increasing in $\operatorname{wt}_p(\psi_1^*\chi_2)$.

LEMMA 11. *The minimum of* $\operatorname{wt}_p(\lambda_1^*\lambda_2)$ *over all multiplicative characters* λ_1 *and* λ_2 *of* $\mathbb{F}_q$ *with prescribed* p*-weights* $w_1 = \operatorname{wt}_p(\lambda_1)$ *and* $w_2 = \operatorname{wt}_p(\lambda_2)$ *is given by*

$$\min_{\lambda_1,\lambda_2} \operatorname{wt}_p(\lambda_1^*\lambda_2) = \begin{cases} 0 & \text{if } w_2 = 0 \text{ and } w_1 = 0\,, \\ e(p-1) & \text{if } w_2 = 0 \text{ and } w_1 = e(p-1)\,, \\ e(p-1) - w_1 & \text{if } w_2 = 0 \text{ and } w_1 \notin \{0, e(p-1)\}\,, \\ w_2 & \text{if } w_2 \neq 0 \text{ and } w_1 = e(p-1)\,, \\ w_2 - w_1 & \text{if } w_2 > w_1\,, \\ r & \text{if } w_2 \leq w_1 \neq e(p-1)\,, \end{cases},$$

with the unique integer in $\{1, \ldots, p-1\}$ *such that* $r \equiv w_2 - w_1 \pmod{p-1}$.

PROOF SKETCH. The first four cases can be verified by plugging in the appropriate values. For the fourth case, note that we have $\operatorname{wt}_p(\lambda_2) = \operatorname{wt}_p(\lambda_1\lambda_1^*\lambda_2) \leq \operatorname{wt}_p(\lambda_1^*\lambda_2) + \operatorname{wt}_p(\lambda_2)$ which implies $\operatorname{wt}_p(\lambda_1^*\lambda_2) \geq \operatorname{wt}_p(\lambda_2) - \operatorname{wt}_p(\lambda_1)$. To see that this bound is achieved, it is sufficient to choose two exponents so that no carry occurs when they are subtracted. For the final case, note that

$$\operatorname{wt}_p(\lambda_1^*\lambda_2) \equiv \operatorname{wt}_p(\lambda_2) - \operatorname{wt}_p(\lambda_1) \pmod{p-1}\,.$$

Intuitively, r should be the smallest possible value satisfying this congruence. However, it is not possible that $r = 0$ since that would mean $\lambda_1^*\lambda_2 = 1$. To complete the proof, one should construct characters so that equality holds but for brevity we omit this final step. □

Plugging in Lemma 11 into Theorem 26 and Equation (3) provides a description of the ultrametric integral transition matrix for one round of a Feistel cipher with an arbitrary round function of p-degree d, in terms of the p-weights of the input and output characters. The resulting formula contains rather many case distinctions, so we instead give an example.

EXAMPLE 17. Let V be a matrix of size $(e(p-1)+1)^2 \times (e(p-1)+1)^2$ containing lower bounds on the valuations $\operatorname{ord}_p A^{\mathsf{F}}_{\chi,\psi}$ for different values of $(\operatorname{wt}_p(\psi_1), \operatorname{wt}_p(\psi_2))$

(columns) and $(\mathrm{wt}_p(\chi_1), \mathrm{wt}_p(\chi_2))$ (rows). For $q = 2^3$ and $d = 2$, we get

$$V = \begin{array}{c|cccc|cccc|cccc|cccc}
 & (\cdot,0) & & & & (\cdot,1) & & & & (\cdot,2) & & & & (\cdot,3) & & & \\
(\cdot,0) & 0 & \infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty \\
 & \infty & \infty & \infty & \infty & 0 & \infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty \\
 & \infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty & 0 & \infty & \infty & \infty & \infty & \infty & \infty & \infty \\
 & \infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty & 0 & \infty & \infty & \infty \\
\hline
(\cdot,1) & 0 & 0 & 2 & 3 & 0 & 1 & 2 & 3 & 0 & 1 & 2 & 3 & 1 & 2 & 3 & 4 \\
 & \infty & \infty & \infty & \infty & 0 & 0 & 2 & 3 & 0 & 1 & 2 & 3 & 0 & 1 & 2 & 3 \\
 & \infty & \infty & \infty & \infty & 0 & 1 & 2 & 3 & 0 & 0 & 2 & 3 & 0 & 1 & 2 & 3 \\
 & \infty & \infty & \infty & \infty & 0 & 1 & 2 & 3 & 0 & 1 & 2 & 3 & 0 & 0 & 2 & 3 \\
\hline
(\cdot,2) & 0 & 0 & 0 & 3 & 0 & 0 & 1 & 3 & 0 & 0 & 1 & 3 & 0 & 1 & 2 & 3 \\
 & \infty & \infty & \infty & \infty & 0 & 0 & 0 & 3 & 0 & 0 & 1 & 3 & 0 & 0 & 1 & 3 \\
 & \infty & \infty & \infty & \infty & 0 & 0 & 1 & 3 & 0 & 0 & 0 & 3 & 0 & 0 & 1 & 3 \\
 & \infty & \infty & \infty & \infty & 0 & 0 & 1 & 3 & 0 & 0 & 1 & 3 & 0 & 0 & 0 & 3 \\
\hline
(\cdot,3) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 & 0 & 3 & 0 & 0 & 1 & 3 \\
 & \infty & \infty & \infty & \infty & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 & 0 & 3 \\
 & \infty & \infty & \infty & \infty & 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 \\
 & \infty & \infty & \infty & \infty & 0 & 0 & 0 & 3 & 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0
\end{array}.$$

4.3. Multiple rounds. If p is small (such as $p = 2$), then the dimensions of the valuation matrix grow only quadratically with e. For many such cases, this makes it feasible to compute V directly. For two rounds of a Feistel cipher, the matrix U of valuation lower bounds can be computed as follows[3]:

$$U_{w_2,w_1} = \min_w \left\{ V_{w_2,w} + V_{w,w_1} \right\}.$$

This can of course be iterated. Instead of calculating and storing complete matrices, it is usually easier to store only a sparse vector encoding the current state clipped at a certain maximum valuation. As a further improvement, this can be combined with the miss-in-the-middle approach. Note that these vectors are nothing but compact representations of subspaces of $k[\mathbb{F}_q^2]$ with increasingly larger dimension. These subspaces form a trail in the general sense defined in the first lecture.

Before giving an example of a property for multiple rounds, it is worth mentioning that the calculation method sketched above is analogous to (but somewhat more sophisticated than) the 'matrix method' to find zero-correlation linear approximations and impossible differentials, and to the original description of the word-based division property [**5**]. In fact, the latter method is essentially the residue-field version of the above for Feistel ciphers on $\mathbb{F}_2^{2n}$. Note that for $\mathbb{F}_2^{2n}$, one keeps track of the Hamming weight of both branches rather than the p-weight[4].

[3]The operations are those of the 'min-plus tropical semiring'.

[4]The two-weight and the Hamming weight are not quite the same because they are defined on different monoids, but this does not matter in the generic case if only the 2-degree of G is used.

EXAMPLE 18. For the same parameters as in Example 17, the following matrix of valuation lower bounds is obtained for three Feistel rounds:

$$V = \begin{array}{c} \\ (\cdot,0) \\ \\ \\ \\ (\cdot,1) \\ \\ \\ \\ (\cdot,2) \\ \\ \\ \\ (\cdot,3) \\ \\ \\ \end{array} \begin{array}{c} \begin{array}{cccc} (\cdot,0) & & & \end{array} \\ \left[\begin{array}{cccc|cccc|cccc|cccc} 0 & \infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 2 & 0 & 0 & 1 & 3 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 2 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 3 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 2 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \\ \begin{array}{cccc} (\cdot,1) & & & \end{array} \\ \\ \\ \\ \begin{array}{cccc} (\cdot,2) & & & \end{array} \\ \\ \\ \\ \begin{array}{cccc} (\cdot,3) & & & \end{array} \\ \\ \\ \end{array} .$$

The last column does not contain any information, given that the Feistel construction is always a permutation. Not surprisingly, the lowest absolute correlation is obtained when the left output character has p-weight equal to one and the weights of the input characters are $(e(p-1), e(p-1)-1)$. As discussed below, this property can be evaluated using $p^{e-1} = 4$ chosen plaintexts.

4.4. Lowest-correlation approximation. As an example, and because it is the approximation that covers the greatest number of rounds, let us take a closer look at the approximation with input weights $(e(p-1), e(p-1)-1)$ and output weights $(1, 0)$ that was identified in Example 18. Without going into details, we mention that this property can be tested using p^{2e-1} encryptions, though this assumes one has additional (weaker) divisibility properties for other input characters with lower p-weights. The corresponding input set is a submonoid of $\mathbb{F}_p^{2e}$ but not of $\mathbb{F}_{p^{2e}}$.

Figure 4 shows how the valuation decreases with an increasing number of rounds for $q \in \{2^{32}, 3^{20}\}$ and $d = 2$. One result implied by this is that for $p = 2$ there is a zero-sum property on eleven rounds with 2^{63} data. This is precisely the property obtained by using the word-based division property. When using the word-based division property, one works over the monoid $\mathbb{F}_2^{2n}$ rather than $\mathbb{F}_{2^n}^2$ but for the generic case both monoids yield the same result. It must be emphasized that this is only the case when we allow arbitrary round functions of given p-degree.

If p is large, then directly computing (products with) the entire matrix of valuation lower bounds is not feasible. One can instead rely on off-the-shelf solvers, or follow a more analytical approach. For this lecture, we limit ourselves to the cases analyzed above.

4.5. Concrete cases. In concrete cases, absolute correlations may be much smaller than the generic upper bound. A dedicated analysis of the trails is then

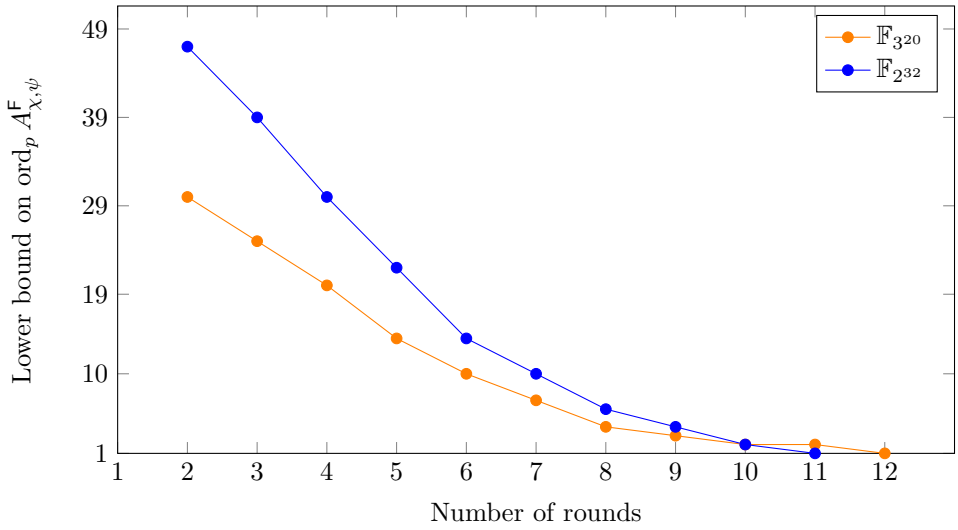

FIGURE 4. Maximum valuation as a function of the number of rounds for a generic Feistel cipher on $\mathbb{F}_q^2$ with round function of 2-degree $d = 2$ and $q \approx 2^{32}$ (specifically, $q \in \{2^{32}, 3^{20}\}$).

necessary. To illustrate this, we take $q = 2^{32}$ as in Figure 4 and assume the round function $\mathsf{G}\colon \mathbb{F}_q \to \mathbb{F}_q$ is given by

$$\mathsf{G}(x) = (x + k_1)^3 + k_2\,,$$

with k_1 and k_2 arbitrary nonzero constants. The constants may be different in different rounds. Note that $s_p(3) = 2$ so the generic analysis from Figure 4 is applicable. We construct a simplified model for this function (likely not tight), by keeping track only of p-weights but taking into account the specific properties of A^{G}. The result (for the same property as before) is shown in Figure 5. The valuation in round 11 is three (as opposed to one), but the valuation still drops to zero after 12 rounds.

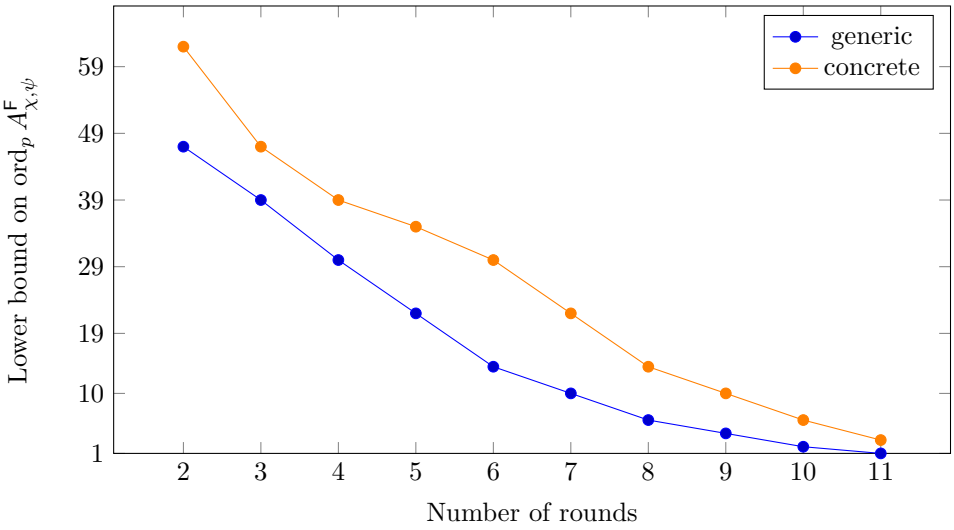

FIGURE 5. Valuation as a function of the number of rounds for a Feistel cipher on $\mathbb{F}_q^2$ with round function $x \mapsto (x + k_1)^3 + k_2$ and $q = 2^{32}$. The property is the same as in Figure 4.

Bibliography

[1] Tim Beyne. A geometric approach to symmetric-key cryptanalysis, June 2023.

[2] Tim Beyne and Michiel Verbauwhede. Integral cryptanalysis using algebraic transition matrices. *IACR Trans. Symm. Cryptol.*, 2023(4):244–269, 2023.

[3] Tim Beyne and Michiel Verbauwhede. Ultrametric integral cryptanalysis. In Kai-Min Chung and Yu Sasaki, editors, *ASIACRYPT 2024, Part VII*, volume 15490 of *LNCS*, pages 392–423. Springer, Singapore, December 2024.

[4] Christina Boura and Anne Canteaut. Another view of the division property. In Matthew Robshaw and Jonathan Katz, editors, *CRYPTO 2016, Part I*, volume 9814 of *LNCS*, pages 654–682. Springer, Berlin, Heidelberg, August 2016.

[5] Yosuke Todo. Structural evaluation by generalized integral property. In Elisabeth Oswald and Marc Fischlin, editors, *EUROCRYPT 2015, Part I*, volume 9056 of *LNCS*, pages 287–314. Springer, Berlin, Heidelberg, April 2015.