Construction and Categories of Codes

G.R. Blakley, I. Borosh, and A. Klappenecker

Department of Mathematics, Texas A&M University, College Station, TX 77843-3368, USA, (blakley|borosh|andreask)@math.tamu.edu

Abstract. Blakley and Borosh introduced a general theory of codes, encompassing cryptographic and error control codes among others. They explored the properties of such general codes with methods from relational algebra and set theory. We provide a categorical point of view, which leads to new constructions of codes. We also exhibit a Jordan-Hölder type theorem and a Schreier refinement technique.

1 Introduction

In the late twentieth century a vast proliferation of codes occurred. Many new cardinalities became common, especially large finite or infinite. Many new arithmetics – infinite as well as finite – could be found in the newly introduced arithmetic-based codes. Hilbert spaces are as integral to the theory of quantum error control as Hamming spaces to classical error control. But many new codes arose without arithmetic, amounting to mere codebooks or databases.

Codes with no encode process, codes with no decode process, codes which encode every plaintext symbol into billions of different codetext expressions are now famous and widely used, as are codes which decode every codetext expression into every plaintext symbol.

Commerce has made ISSN, ISBN, UPC commonplace. Locks are codes, phonebooks are, genomes are, cash register receipts are, codes replace telephone wires, tollbooths, signatures.

This is not metaphorical talk. Every one of these objects is a code in a strict mathematical sense. And one realization that emerges from this mathematical view of codes is the profound importance of stuctural considerations. Codes have shapes, just as molecules have shapes. And the designer of codes has a larger repertory of kinds of structures to draw upon than an organic chemist. Moreover, these kinds of structures can be usefully described and combined by the methods of universal algebra, as adumbrated in [3]. But they also lend themselves to treatment by category theory, as will become clearer below.

After defining and visualizing precodes and codes, we introduce the corresponding categories in Section 4. We obtain products, limits, and colimits in the usual way in Sections 5 and 6, and then give some examples. The subprecodes of a precode form a lattice in a natural way. Section 8 discusses some of its properties and, in particular, establishes a Jordan-Hölder-Schreier theory for it. This result suggests a unified view of several cryptanalytic methodologies. We conclude with comments on this need for further results.

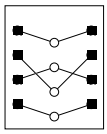
2 Visual Presentation of Precodes and Codes

The code definition given in the general theory [2] of codes goes as follows. A **precode** is a list (P, C, e, d) whose entries are a set P of plaintext symbols, a set C of codetext symbols, an encode relation $e \subseteq P \times C$, and a decode relation $d \subseteq C \times P$. A **code** is a precode for which the composite relation $de \subseteq P \times P$ is subdiagonal, i. e. contains only pairs of the form (p, p).

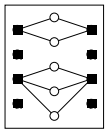
We follow [2] and introduce a graphical representation of precodes and codes called a **strip chart**. The items P, e, C, d, P are represented by five columns of marks. The marks in the three 'symbol' columns P, C, P are at various heights – different heights signifying different elements. The marks in the two 'relation' columns e, d are undirected line segments treated as if they were arrows going from left to right.

For example the first strip chart below represents the popular notion of a code, as a pair e, d of bijections between P and C. The bijection d going from the codetext symbols C (the set of open rings in the third column) to the plaintext symbols P (the set of blobs in the fifth column) is the inverse of the bijection e going from the blobs of P in the first column to the rings of C in the third column. Clearly each action of e moves a first-column blob b to a third-column ring r, and then d takes this ring to a fifth-column blob at the same height as the original blob. In other words, de takes each blob to itself.

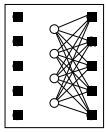
It is worth becoming acquainted with the strip charts below. They give a weak foreshadowing of the huge variety of codes already in use. And they set the stage for the purposeful use of abstract structure to produce novel codes of yet widely different structural types which the general theory of codes can supply for various information-related investigations or activities.



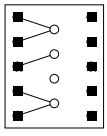
In the figure at left, the encode e is a bijection from P to C, the decode d is its inverse function, a bijection from C to P. Gray codes, key settings of Caesar ciphers, RSA, DES, AES, some commercial codebooks and Gödel numbering are among the many example of this matched-pair-of-bijections type of code.



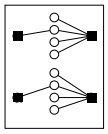
The code at left has a decode d which is a function (i. e., many-to-one relation). Its encode e is the converse of d (whence it is a one-to-many relation. There are many such codes, including some codebooks with homophones, the calculus (in which encode is antidifferention of a function and decode is differentiation, secret sharing schemes, and hash function codes (hashes are decodes, and their converses are encodes).



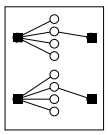
The code at left has an empty encode relation $e = \emptyset$, and a full decode (an all-to-all relation) d. Hence de is empty, so this strip chart does, indeed, present a code. The Diffie-Hellman key exchange [6] is often a pair of codes of this one-way 'encodeless' type. The genetic code is similar, but less extreme. Its encode is empty, and its decode is a function.



The code at left has an empty decode relation $d = \emptyset$, and an encode relation e which is a function. Clearly $de = \emptyset$. The Purdy high security login [9] is a code of this one-way 'decodeless' type. Other such examples are hash functions when viewed as one-way objects.



The code at left has an injective (i.e., one-to-one) encode relation e, and a decode relation d which is a function. Many error correcting codes are of this type. In particular, the code here is (isomorphic to) the triplication code with $e(0) = 000$, $e(1) = 111$, and with majority decode.



The code at left has an injective decode d, and a one-to-many encode relation e. It amounts to a variant of the magnetic-strip-card-key code, which large hotels use to give guests entry to their rooms. Each room is 'encoded' as a large set of bit strings, only one of which is valid today. A card with this string 'decodes' the room door to open today – thus revealing the encoded room. Cards with any other bit string don't open it. And tomorrow the decode d may be changed, but the encode e will remain the same. But no two rooms can ever be opened by the same card.

3 Basic Notions

Let P and C be sets. We will call P the plaintext symbols, and C the codetext symbols. Let e be a subset of $P \times C$, called the encoding relation, and d a subset of $C \times P$, called the decoding relation. Then $\mathcal{R} = (P, C, e, d)$ is called a **precode**. $\mathcal{R}$ is said to be a **code** if and only if $d \circ e$ is a subdiagonal relation on P.

A **precode homomorphism** from (P, C, e, d) to (P', C', e', d') is defined by a list of functions $(h, k, h \times k, k \times h)$, where $h\colon P \to P'$ and $k\colon C \to C'$ are required to satisfy $(h \times k)(e) \subseteq e'$ and $(k \times h)(d) \subseteq d'$. Sometimes we will write $\langle h, k \rangle$ to denote this homomorphism.

We obtain a **category $\mathfrak{P}$ of precodes** by taking precodes as **objects** and precode homomorphisms as **morphisms**.

The identity functions on the plaintext and the codetext symbols of a precode $\mathcal{R}$ induce the **identity morphism** $1_{\mathcal{R}}$ associated to the precode $\mathcal{R}$. The composition of morphisms is given by the composition of functions. The class of objects in the category $\mathfrak{P}$ is denoted by $\mathrm{Obj}(\mathfrak{P})$. The set of morphisms from $\mathcal{R}$ to $\mathcal{A}$ is denoted by $\mathrm{Hom}_{\mathfrak{P}}(\mathcal{R}, \mathcal{A})$ or sometimes simply by $\mathcal{R} \to \mathcal{A}$.

Let $\mathcal{R} = (P, C, e, d)$ and $\mathcal{A} = (P', C', e', d')$ be precodes. $\mathcal{R}$ is said to be a **subprecode** of $\mathcal{A}$ if and only if $P \subseteq P'$, $C \subseteq C'$, $e \subseteq e'$, and $d \subseteq d'$. Notice that a subprecode of a code is again a code.

The **subcategory $\mathfrak{C}$ of codes** of the category $\mathfrak{P}$ of precodes is defined in the obvious way, taking codes as objects. Note that $\mathfrak{C}$ is a **full** subcategory of $\mathfrak{P}$. Indeed, consider a precode homomorphism between two codes $\mathcal{K}$ and $\mathcal{L}$. The image under this homomorphism is a subprecode of $\mathcal{L}$, and hence a subcode. Thus the set $\mathrm{Hom}_{\mathfrak{P}}(\mathcal{K}, \mathcal{L})$ of precode morphisms coincides with the set of code morphisms $\mathrm{Hom}_{\mathfrak{C}}(\mathcal{K}, \mathcal{L})$.

The code $\mathcal{I} = (\emptyset, \emptyset, \emptyset, \emptyset)$ is an **initial object** in the category of precodes, that is, there exists exactly one morphism from $\mathcal{I}$ to any another object $\mathcal{R}$ in $\mathfrak{P}$. Any code $\mathcal{T} = (\{p\}, \{c\}, \{(p, c)\}, \{(c, p)\})$ with singleton set symbols is a **terminal object** in $\mathfrak{P}$. The category $\mathfrak{P}$ does not have a zero object.

4 Morphisms

We take a closer look at the morphisms of precodes in this section. It turns out that the monomorphisms are just the injective functions respecting the encoding and decoding relations. We will see that epimorphisms need not be so well-behaved. Indeed, neither the category of precodes nor the category of codes is balanced, that is, bimorphisms need not be isomorphisms.

A morphism $f\colon \mathcal{A} \to \mathcal{B}$ of precodes $\mathcal{A}$ and $\mathcal{B}$ is said to be a **monomorphism**, or simply **monic**, if and only if $fg = fg'$ implies $g = g'$ for all $g, g' \in \mathrm{Hom}_{\mathfrak{P}}(\mathcal{R}, \mathcal{A})$ and all $\mathcal{R} \in \mathrm{Obj}(\mathfrak{P})$.

Lemma 1. *Let $f\colon \mathcal{R} \to \mathcal{A}$ be a morphism of precodes. Then $f = \langle f_1, f_2 \rangle$ is monic if and only if f_1 and f_2 are injective functions.*

Proof. Suppose that f_1 and f_2 are injective, hence monic, morphisms in the category of sets. This immediately implies that f is monic. Conversely, suppose that f is monic. Denote by $\mathcal{S} = (\{p\}, \{c\}, \emptyset, \emptyset)$ a precode with singleton symbol sets. Let x and y be (necessarily constant) morphisms from $\mathcal{S}$ to $\mathcal{R}$. Since $fx = fy$ implies $x = y$, it follows that f_1 and f_2 are injective. □

A morphism $f\colon \mathcal{R} \to \mathcal{A}$ of precodes $\mathcal{R}$ and $\mathcal{A}$ is said to be an **epimorphism** if and only if $gf = g'f$ implies $g = g'$ for all $g, g' \in \mathrm{Hom}_{\mathfrak{P}}(\mathcal{A}, \mathcal{B})$ and all $\mathcal{B} \in \mathrm{Obj}(\mathfrak{P})$.

Lemma 2. *Let $f\colon \mathcal{R} \to \mathcal{A}$ be a morphism of precodes. Then $f = \langle f_1, f_2 \rangle$ is epic if and only if f_1 and f_2 are surjective functions.*

Proof. Suppose that f_1 and f_2 are surjective functions, hence epimorphisms, in the category of sets. This implies that f is an epimorphism.

Let f be an epimorphism. Seeking a contradiction, we assume that not both f_1 and f_2 are surjective. Denote by $\mathbf{2}$ the two element set $\{0, 1\}$. Define the precode $\mathcal{R} = (\mathbf{2}, \mathbf{2}, \mathbf{2} \times \mathbf{2}, \mathbf{2} \times \mathbf{2})$. Let g and h be two *distinct* morphisms in $\mathrm{Hom}_{\mathfrak{P}}(\mathcal{A}, \mathcal{B})$ that take the same values on the image of f. It follows from our assumption that such morphisms exist. However, since $gf = hf$ implies $g = h$, we get the desired contradiction. □

A morphism $f\colon \mathcal{R} \to \mathcal{A}$ of precodes $\mathcal{R}$ and $\mathcal{A}$ is called an **isomorphism** if and only if there exists a morphism $g\colon \mathcal{A} \to \mathcal{R}$ such that $fg = 1_{\mathcal{A}}$ and $gf = 1_{\mathcal{R}}$.

Lemma 3. *Let $f\colon \mathcal{R} \to \mathcal{A}$ be a morphism of precodes. If $f = \langle f_1, f_2 \rangle$ is an isomorphism then f_1 and f_2 are bijective functions.*

Proof. An isomorphism is monic and epic, implying that f_1 and f_2 are bijective functions. □

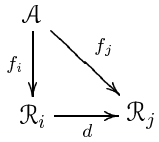
The evident asymmetry in the statement of this lemma reflects the fact that an epimorphism $f = \langle f_1, f_2 \rangle$ need not be surjective on the encoding relation or decoding relation, even though the functions f_1 and f_2 *are* surjective. This fact has some quizzical consequences. For instance, a monic and epic morphism in the category of precodes is not necessarily an isomorphism. To see this, let ι denote the identity function on $\mathbf{2}$. Then $\langle \iota, \iota \rangle$ is a monic and epic morphism from $\mathcal{R} = (\mathbf{2}, \mathbf{2}, \emptyset, \emptyset)$ to $\mathcal{A} = (\mathbf{2}, \mathbf{2}, \mathbf{2} \times \mathbf{2}, \mathbf{2} \times \mathbf{2})$. But it is obviously not an isomorphism.

5 Limits

In this section we derive some fairly general constructions of codes and precodes. The constructions are based on the categorical notion of a limit.

Recall that a **diagram** D in a category $\mathfrak{P}$ is a directed graph whose vertices $i \in I$ are labelled by objects $\mathcal{R}_i$ in $\mathfrak{P}$ and whose edges $i \to j$ are labelled by morphisms in $\mathrm{Hom}_{\mathfrak{P}}(\mathcal{R}_i, \mathcal{R}_j)$. The underlying graph is called the **scheme** of the diagram.

A family of morphisms $(f_i \colon \mathcal{A} \to \mathcal{R}_i)_{i \in I}$ with common domain $\mathcal{A}$ is said to be a **cone** for D, provided that for each arrow $d : \mathcal{R}_i \to \mathcal{R}_j$ in the diagram D, the triangle



commutes. A **limit** for D is a cone for D with the universal property that any other cone for D uniquely factors through it. In other words, if $(f_i \colon \mathcal{A} \to \mathcal{R}_i)_{i \in I}$ is the limit of a diagram D and $(g_i \colon \mathcal{B} \to \mathcal{R}_i)_{i \in I}$ is a cone for D, then there exists exists exactly one arrow $u \colon \mathcal{B} \to \mathcal{A}$ such that $g_i = f_i \circ u$ for all $i \in I$.

We want to show that the category $\mathfrak{P}$ of precodes and the category $\mathfrak{C}$ of codes are complete. In other words, we need to show that limits exist for all diagrams. Fortunately, it is sufficient to prove that products and equalizers exists [1, 7].

We need to introduce some more notation. Let $(r_i)_{i \in I}$ be a family of relations indexed by a set I, where $r_i \subseteq P_i \times C_i$. We can define a product of these relations by

$$\prod r_i = \left\{ (p, c) \in \prod P_i \times \prod C_j \mid \forall i \in I \ (p(i), c(i)) \in r_i \right\},$$

where all products range over the index set I. Sometimes we will denote the product of two relations r_i and r_j by $r_i \otimes r_j$. For example, if $r_1 = \{(2,1),(1,2)\}$, $r_2 = \{(a,b),(a,c)\}$, then the product relation $r_1 \otimes r_2$ is given by

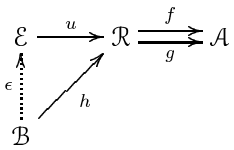
$$r_1 \otimes r_2 = \{((2,a),(1,b)),((2,a),(1,c)),((1,a),(2,b)),((1,a),(2,c))\}.$$

Theorem 1. *The category $\mathfrak{P}$ of precodes has products. The product of a family of codes is again a code.*

Proof. Let $\mathcal{R}_i = (P_i, C_i, e_i, d_i), i \in I$, be a family of precodes indexed by a set I. The product of this family is obtained by taking cartesian products of the symbol sets, and the product of the encoding and decoding relations. In other words, the product of the family $\mathcal{R}_i$ is given by $(\mathcal{R}, (\pi_i \colon \mathcal{R} \to \mathcal{R}_i)_{i \in I})$, where the precode $\mathcal{R}$ is given by the object $(\prod_{i \in I} P_i, \prod_{i \in I} C_i, \prod_{i \in I} e_i, \prod_{i \in I} d_i)$, and the projection map π_i is the obvious map onto the ith component. It is clear that $\mathcal{R}$ is a code if and only if all $\mathcal{R}_i$ are codes. □

The equalizer $(\mathcal{E}, u)$ of two morphisms $f, g \colon \mathcal{R} \to \mathcal{A}$ is an object $\mathcal{E}$ together with a morphism $u \colon \mathcal{E} \to \mathcal{R}$ such that $fu = gu$, with the additional property

that every morphism h satisfying $fh = gh$ factors uniquely through u. In other words, the triangle in the following diagram commutes:



Recall that in the category of sets, the equalizer of two functions $f, g\colon R \to A$ is given by the coincidence set $\{x \in R \mid f(x) = g(x)\}$ with the inclusion mapping.

Theorem 2. *The category $\mathfrak{P}$ has equalizers. If $(\mathcal{E}, u)$ is the equalizer of two morphisms between codes, then $\mathcal{E}$ is also a code.*

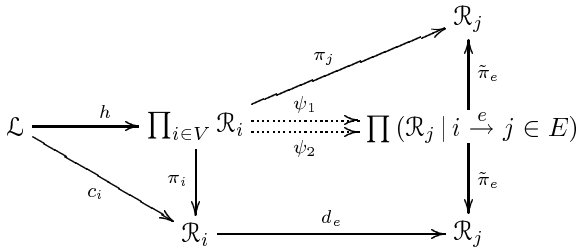
Proof. Let $\mathcal{R} = (P, C, e, d)$ and $\mathcal{A}$ be precodes. Let $f = \langle f_1, f_2 \rangle$ and $g = \langle g_1, g_2 \rangle$ be a pair of morphisms between $\mathcal{R}$ and $\mathcal{A}$. We give an explicit construction of the equalizer.

The equalizer $(\mathcal{E}, u)$ of f and g is given by the precode $\mathcal{E} = (P^*, C^*, e^*, d^*)$, where the plaintext symbols $P^* = \{\, a \in P \mid f_1(a) = g_1(a) \,\}$ and codetext symbols $C^* = \{\, a \in C \mid f_2(a) = g_2(a) \,\}$ are just coincidence sets, and the encoding and decoding relations are obtained from $\mathcal{R}$ by restriction, that is, $e^* = e\,|_{P^* \times C^*}$, $d^* = d\,|_{C^* \times P^*}$, and the morphism $u = \langle \iota_1, \iota_2 \rangle$ is induced by the set inclusion maps $\iota_1\colon P^* \to P$, $\iota_2\colon C^* \to C$.

The construction ensures that $u(\mathcal{E})$ is the largest subprecode of $\mathcal{R}$ such that the restrictions of the functions f and g on $u(\mathcal{E})$ coincide, $f|_{u(\mathcal{E})} = g|_{u(\mathcal{E})}$. We can express h by a composition of a morphism $\epsilon\colon \mathcal{B} \to \mathcal{E}$ with u, since $h(\mathcal{B})$ is a subprecode of $u(\mathcal{E})$. The morphism ϵ is uniquely determined, since u is a monomorphism. □

Theorem 3. *The category $\mathfrak{P}$ of precodes and the category $\mathfrak{C}$ of codes are complete.*

Proof. The categories $\mathfrak{P}$ and $\mathfrak{C}$ have products and equalizers and are therefore complete [1, 7]. The main idea of this standard construction goes as follows. Suppose that we are given a diagram D in $\mathfrak{P}$ with sets V of vertices and E of edges. We build two products: the product of all objects in D, and the product indexed by E of all codomains of arrows in D. The universal property of the E-indexed product induces unique maps ψ_1 and ψ_2 as is shown in the following diagram:



The map h is given by the equalizer of ψ_1 and ψ_2, and the maps c_i are given by composition of h with the projection maps π_i, that is, $c_i = \pi_i\, h$. It is not difficult to see that $(\mathcal{L}, (c_i)_{i\in I})$ is a cone of D. It follows from the universality of the equalizer and of the V-indexed product that this cone is the limit of D. □

6 Colimits

Reversing arrows, we obtain the concept of **cocones** and **colimits** of diagrams. We will derive the dual results for precodes.

Theorem 4. *The category $\mathfrak{P}$ of precodes has coproducts. The coproduct of a family of codes is again a code.*

Proof. The coproduct $(\mathcal{K}, (\iota_i\colon \mathcal{R}_i \to \mathcal{K})_{i\in I})$ of the family $\mathcal{R}_i$ is given by the disjoint union of the symbol sets and the induced disjoint union of the encoding and decoding relations together with the obvious inclusion maps. In other words,

$$\mathcal{K} = \left(\bigcup_{i\in I} P_i \times \{i\}, \bigcup_{i\in I} C_i \times \{i\}, \bigcup_{i\in I} e_i \otimes \Delta_i, \bigcup_{i\in I} d_i \otimes \Delta_i \right),$$

where Δ_i denotes the relation $\Delta_i = \{(i,i)\}$.

It is clear that $\mathcal{K}$ is a code if and only if all $\mathcal{R}_i$ are codes. □

Theorem 5. *The category $\mathfrak{P}$ has coequalizers.*

Proof. Let $\mathcal{R}$ and $\mathcal{A} = (P, C, e, d)$ be precodes, and let $f = \langle f_1, f_2\rangle$ and $g = \langle g_1, g_2\rangle$ be a pair of morphisms between $\mathcal{R}$ and $\mathcal{A}$. Let E_1 be the smallest equivalence relation on P such that $f_1(a)$ and $g_1(a)$ are equivalent. Similarly, let E_2 be the smallest equivalence relation on P such that $f_2(a)$ and $g_2(a)$ are equivalent. The coequalizer of f and g is given by the precode $(P/E_1, C/E_2, e/E_1 \otimes E_2, d/E_2 \otimes E_1)$ and the morphism $\langle c_1, c_2\rangle$ induced by the canonical quotient maps $c_1\colon P \to P/E_1$ and $c_2\colon C \to C/E_2$. □

Remark 1. The coequalizer of two codes in the category $\mathfrak{P}$ is not necessarily a code. For example, let $\mathcal{K} = (\{1,2\}, \{1,2\}, id, id)$ the code with identity encoding and decoding relations. Denote by i and s the two bijective functions from $\{1,2\}$ into itself. Then the coequalizer of the morphisms $\langle i, i\,\rangle$ and $\langle i, s\rangle$ is the precode $\mathcal{E}$ given by

$$\mathcal{E} = (\{1,2\}, \{[1]\}, \{(1,[1]),(2,[1])\}, \{([1],1),([1],2)\}).$$

Theorem 6. *The category $\mathfrak{P}$ of precodes is cocomplete.*

Proof. The category $\mathfrak{P}$ has coproducts and coequalizers and is therefore cocomplete. □

7 Examples

Example 1 (RSA). Denote by p and q two distinct odd primes. A key setting of an RSA public key cryptosystem [10] can be seen as a code over the symbol set $\mathbf{Z}/pq\mathbf{Z}$, where the encoding relation e is given by the function $x \mapsto x^{\varepsilon} \bmod pq$ and the decoding relation d is given by $x \mapsto x^{\delta} \bmod pq$. The exponents are assumed to satisfy the congruence $\varepsilon\delta \equiv 1 \bmod \varphi(pq)$, where φ is Euler's totient function. We denote this code by $\mathcal{RSA} = (\mathbf{Z}/pq\mathbf{Z}, \mathbf{Z}/pq\mathbf{Z}, e, d)$.

Reducing the symbol sets modulo p and q respectively, one obtains two key-settings of Pohlig-Hellman cryptosystems [8], denoted by

$$\mathcal{PH}_1 = (\mathbf{Z}/q\mathbf{Z}, \mathbf{Z}/q\mathbf{Z}, e_1, d_1) \quad \text{and} \quad \mathcal{PH}_2 = (\mathbf{Z}/p\mathbf{Z}, \mathbf{Z}/p\mathbf{Z}, e_2, d_2).$$

The encoding and decoding relations are obtained from e and d by reducing modulo p and q respectively. For instance, the relation e_1 is given by the function $x \mapsto x^{\varepsilon} \bmod q$.

The $\mathcal{RSA}$ code is, in the terminology introduced in the next section, an example of a **product** of the codes $\mathcal{PH}_1$ and $\mathcal{PH}_2$.

Example 2 (RSA, cont'd). Conversely, given two Pohlig-Hellman codes

$$\begin{aligned}\mathcal{PH}_1 &= (\mathbf{Z}/q\mathbf{Z},\ \mathbf{Z}/q\mathbf{Z},\ x \mapsto x^{\varepsilon_1} \bmod q, x \mapsto x^{\delta_1} \bmod q)\\ \mathcal{PH}_2 &= (\mathbf{Z}/p\mathbf{Z},\ \mathbf{Z}/p\mathbf{Z},\ x \mapsto x^{\varepsilon_2} \bmod p, x \mapsto x^{\delta_2} \bmod p)\end{aligned}$$

and assuming that $\gcd(p-1, q-1) | (\varepsilon_1 - \varepsilon_2)$, then it is easy to see that the greatest common divisor of $p-1$ and $q-1$ divides $\delta_1 - \delta_2$. The Chinese remainder theorem yields the integers ε, δ satisfying

$$\begin{aligned}\varepsilon &\equiv \varepsilon_1 \bmod q-1, \qquad & \delta &\equiv \delta_1 \bmod q-1,\\ \varepsilon &\equiv \varepsilon_2 \bmod p-1, \qquad & \delta &\equiv \delta_2 \bmod p-1,\end{aligned}$$

respectively. The $\mathcal{RSA}$ code

$$(\mathbf{Z}/pq\mathbf{Z}, \mathbf{Z}/pq\mathbf{Z}, x \mapsto x^{\varepsilon} \bmod pq, x \mapsto x^{\delta} \bmod pq)$$

is then isomorphic to the product of $\mathcal{PH}_1$ and $\mathcal{PH}_2$.

Example 3 (Unequal Error Protection). We construct a simple (nonlinear) error control code that protects 0 and 1 against one single error, and can detect a single error in the transmission of 20 other symbols $\{2, \ldots, 21\}$. This code is constructed with the help of two smaller codes.

Denote by $\mathbf{F}_2$ the binary finite field. Let C_1 be the set of all codewords in $\mathbf{F}_2^6$ of (Hamming) weight 0, 1, 5, and 6. Let $\mathcal{A}_1$ be the code $(\mathbf{F}_2, C_1, e_1, d_1)$, where $e_1(0) = 000000$ and $e_1(1) = 111111$, and the decoding relation d_1 maps all codewords of weight 0 or 1 to the plaintext symbol 0, and maps all codewords of weight 5 or 6 to the plaintext symbol 1.

Let C_2 be the set of all codewords in $\mathbf{F}_2^6$ of weight 3. The plaintext symbol set is given by $P_2 = \{2, \ldots, 21\}$. The encoding relation e_2 maps the symbols

2,...,21 to the codewords in C_2 in lexical order respectively, and the decoding relation is given by the inverse function $d_2 = e_2^{-1}$. Then $\mathcal{A}_2 = (P_2, C_2, e, e^{-1})$.

The code $\mathcal{R}$ is given by the union of the codes $\mathcal{A}_1$ and $\mathcal{A}_2$, that is,

$$\mathcal{R} = \left(\{0, \ldots, 22\}, \mathbf{F}_2^6, e_1 \cup e_2, d_1 \cup d_2\right).$$

This code is (isomorphic to) the **coproduct** (as defined in section 6) of the codes $\mathcal{A}_1$ and $\mathcal{A}_2$.

Example 4 (Codes over p-adic Integers). The famous explanation of the nonlinear Kerdoc and Preparata error control codes as linear codes over $\mathbf{Z}/4\mathbf{Z}$ gave rise to other explorations of Hensel lifting in coding theory. In [4], Calderbank and Sloane investigated a series of Hamming codes over the symbol sets $\mathbf{Z}/2^n\mathbf{Z}$. The familiar binary [7,4] Hamming code has generator polynomial $x^3 + x + 1$. Hensel lifting of this generator polynomial to $\mathbf{Z}/4\mathbf{Z}$ gives a unique monic irreducible polynomial that divides $x^7 - 1$ in $\mathbf{Z}/4\mathbf{Z}[x]$. Proceeding further, one obtains a series of cyclic codes over $\mathbf{Z}/8\mathbf{Z}$, $\mathbf{Z}/16\mathbf{Z}$, $\mathbf{Z}/32\mathbf{Z}$, etc. The 2-adic lift of the binary Hamming code is then the error control code over the ring of 2-adic integers with generator matrix

$$\begin{pmatrix} 1 & \lambda & \lambda^* & -1 & 0 & 0 & 0 \\ 0 & 1 & \lambda & \lambda^* & -1 & 0 & 0 \\ 0 & 0 & 1 & \lambda & \lambda^* & -1 & 0 \\ 0 & 0 & 0 & 1 & \lambda & \lambda^* & -1 \end{pmatrix},$$

where λ is the 2-adic integer $(1 - \sqrt{-7})/2$, and $\lambda^* = \lambda - 1$.

The code $(\mathbf{Z}_2^4, \mathbf{Z}_2^7, e, d)$ corresponding to this Hamming code over the 2-adic integers $\mathbf{Z}_2$ is a special case of the **limit** construction of codes described in Section 5.

8 Subprecode Lattice.

The following question was posed in [3]: Does there exist a Jordan-Hölder-Schreier theory of codes? We give an affirmative answer to this question in this section.

Denote by $\mathrm{Lat}(\mathcal{R})$ the set of subprecodes of a precode $\mathcal{R}$. The subprecode relation defines a partial order $\leq$ on $\mathrm{Lat}(\mathcal{R})$, namely, $\mathcal{R}_i \leq \mathcal{R}_j$ if and only if $\mathcal{R}_i$ is a subprecode of $\mathcal{R}_j$.

Proposition 1. *The partially ordered set* $\mathrm{Lat}(\mathcal{R})$ *of subprecodes of a precode* $\mathcal{R}$ *is a lattice. In particular, the subcodes of a code form a lattice.*

Proof. Define the **join** $\mathcal{R}_i \vee \mathcal{R}_j$ of two precodes $\mathcal{R}_i$ and $\mathcal{R}_j$ by their union $\mathcal{R}_i \vee \mathcal{R}_j = (P_i \cup P_j, C_i \cup C_j, e_i \cup e_j, d_i \cup d_j)$; and define the **meet** $\mathcal{R}_i \wedge \mathcal{R}_j$ of two precodes $\mathcal{R}_i$ and $\mathcal{R}_j$ by their intersection $\mathcal{R}_i \wedge \mathcal{R}_j = (P_i \cap P_j, C_i \cap C_j, e_i \cap e_j, d_i \cap d_j)$. Clearly, $\mathcal{R}_i \vee \mathcal{R}_j$ is the smallest precode containing $\mathcal{R}_i$ and $\mathcal{R}_j$, and $\mathcal{R}_i \wedge \mathcal{R}_j$ is the largest subprecode contained in both $\mathcal{R}_i$ and $\mathcal{R}_j$. Thus, $\mathrm{Lat}(\mathcal{R})$ is indeed a lattice

with respect to those meet and join operations. The second statement follows immediately, since every subprecode of a code is again a code. □

Notice that the lattice $\mathrm{Lat}(\mathcal{R})$ is bounded, since all subprecodes $\mathcal{A}$ of $\mathcal{R}$ satisfy $0 \leq \mathcal{A} \leq 1$, where the bounds 0 and 1 are given by $0 = (\emptyset, \emptyset, \emptyset, \emptyset)$ and $1 = \mathcal{R}$. The lattice $\mathrm{Lat}(\mathcal{R})$ is distributive, since the distributive laws of the meet and join operations

$$\begin{aligned} &\forall\, \mathcal{A}, \mathcal{B}, \mathcal{C} \in \mathrm{Lat}(\mathcal{R}): \quad \mathcal{A} \wedge (\mathcal{B} \vee \mathcal{C}) = (\mathcal{A} \wedge \mathcal{B}) \vee (\mathcal{A} \wedge \mathcal{C}), \\ &\forall\, \mathcal{A}, \mathcal{B}, \mathcal{C} \in \mathrm{Lat}(\mathcal{R}): \quad \mathcal{A} \vee (\mathcal{B} \wedge \mathcal{C}) = (\mathcal{A} \vee \mathcal{B}) \wedge (\mathcal{A} \vee \mathcal{C}), \end{aligned}$$

follow immediately from the set theoretic union and intersection properties. Thus, we can strengthen the statement of Proposition 1 as follows:

Proposition 2. *The partially ordered set* $\mathrm{Lat}(\mathcal{R})$ *of subprecodes of a precode* $\mathcal{R}$ *is a bounded distributive lattice.*

A Schreier refinement theorem can be derived for any modular lattice, and thus in particular for the distributive lattice $\mathrm{Lat}(\mathcal{R})$. We need to introduce some terminology to state this result. Let $\mathcal{A}$ and $\mathcal{B}$ be two precodes in $\mathrm{Lat}(\mathcal{R})$ such that $\mathcal{A} \leq \mathcal{B}$. The subset $[\mathcal{A}, \mathcal{B}] = \{\mathcal{C} \in \mathrm{Lat}(\mathcal{R}) \,|\, \mathcal{A} \leq \mathcal{C} \leq \mathcal{B}\}$ is called the **interval** between the precodes $\mathcal{A}$ and $\mathcal{B}$. Two **chains** in a subprecode lattice $\mathrm{Lat}(\mathcal{R})$,

$$\mathcal{A} = \mathcal{A}_0 \leq \ldots \leq \mathcal{A}_m = \mathcal{B}, \tag{1}$$

$$\mathcal{A} = \mathcal{B}_0 \leq \ldots \leq \mathcal{B}_n = \mathcal{B}, \tag{2}$$

between the same subprecodes $\mathcal{A}$ and $\mathcal{B}$ of $\mathcal{R}$ are said to be **isomorphic** if and only if $m = n$ and there is a permutation π of $1, \ldots, n$ such that the interval $[\mathcal{A}_{i-1}, \mathcal{A}_i]$ is lattice-isomorphic to the interval $[\mathcal{B}_{\pi(i)-1}, \mathcal{B}_{\pi(i)}]$. Defining the precodes $\mathcal{A}_{01} = \mathcal{B}_{01} = \mathcal{A}$, and

$$\mathcal{A}_{ij} = (\mathcal{A}_i \wedge \mathcal{B}_j) \vee \mathcal{B}_{j-1}, \qquad \mathcal{B}_{ji} = (\mathcal{B}_j \wedge \mathcal{A}_i) \vee \mathcal{A}_{i-1},$$

for $i = 1, \ldots, n$ and $j = 1, \ldots, m$, we obtain a refinement of chain (1) and (2) by

$$\begin{aligned} &\mathcal{A} = \mathcal{A}_{01} \leq \cdots \leq \mathcal{A}_{m1} \leq \mathcal{A}_{12} \leq \cdots \leq \mathcal{A}_{m2} \leq \mathcal{A}_{13} \leq \cdots \mathcal{A}_{mn} = \mathcal{B}, \\ &\mathcal{A} = \mathcal{B}_{01} \leq \cdots \leq \mathcal{B}_{n1} \leq \mathcal{B}_{12} \leq \cdots \leq \mathcal{B}_{n2} \leq \mathcal{B}_{13} \leq \cdots \mathcal{B}_{nm} = \mathcal{B}, \end{aligned}$$

respectively. Since the lattice $\mathrm{Lat}(\mathcal{R})$ is modular, these two chains are isomorphic, cf. [5, p. 70]. Therefore, one obtains the following Schreier-type refinement proposition for precodes:

Proposition 3. *Any two chains between two precodes in* $\mathrm{Lat}(\mathcal{R})$ *have isomorphic refinements.*

As a consequence, we obtain a Jordan-Hölder-type proposition:

Proposition 4. *Suppose that the precode* $\mathcal{R} = (P, C, e, d)$ *has finite symbol sets. Then any chain can be refined to a maximal chain and any two maximal chains between two given end-points have the same length.*

Unfortunately, this proposition is not as useful as its group theoretic analogue, since a maximal chain reflects the size of the precode and not its structure. For example, suppose that the cardinalities of P, C, e and d are α, β, γ and δ. Then all maximal chains are of length $\alpha + \beta + \gamma + \delta$.

9 Conclusions

A cryptanalyst who breaks a monoalphabetic substitution cipher by uncovering successively the plaintext value of various codetext symbols (e.g. by means of frequency analysis) does what amounts to forming an increasing sequence of subcodes of the cipher under attack. Similarly, an attack on a polyalphabetic substitution cipher which recovers one alphabet after another can be viewed as discovering homomorphic images of that cipher. It may often make sense to approach a cryptanalytic problem as a sequence of breaks of a sequence of homomorphic images of, or subobjects of, a code which is a key-setting of a cryptosystem.

The four propositions above form one schema for the first of these two approaches, but can involve lengthy maximal chains. A complementary – perhaps more incisive – Jordan-Hölder-Schreier theory might be obtained by recourse to a different partial order on a collection of precodes, such as an order based on homomorphic images (or even the strong homomorphic images suggested by the three isomorphim theorems in [3]).

This paper has shown that the general theory of codes introduced in [2, 3] can be formulated in category-theoretical terms. It has presented constructions such as limits, colimits, equalizers, and has showed that special cases are in fact already present in existing codes in current use.

Acknowledgements. A.K. thanks the Santa Fe Institute for support through their Fellow-at-Large program.

References

1. M.A. Arbib and E. Manes. *Arrows, Structures, and Functors – The Categorical Imperative*. Academic Press, New York, 1975.
2. G.R. Blakley and I. Borosh. A general theory of codes, I: Basic concepts. In D. Dorninger, G. Eigenthaler, H.K. Kaiser, H. Kautschistsch, and W. More, editors, *Proceedings of the Klagenfurt Conference*, volume 10 of *Contributions to General Algebra*, pages 1–29. Verlag Johannes Heyn, Klagenfurt, Austria, 1998.
3. G.R. Blakley and I. Borosh. A general theory of codes, II: Paradigms and homomorphisms. In E. Okamoto, G. Davida, and M. Mambo, editors, *Information Security, First International Workshop, ISW '97*, volume 1396 of *LNCS*, pages 1–30. Springer Verlag, Berlin, 1998.
4. A.R. Calderbank and N.J.A. Sloane. Modular and p-adic cyclic codes. *Designs, Codes, and Cryptography*, 6:21–35, 1995.
5. P.M. Cohn. *Universal Algebra*. D. Reidel Publishing Company, Dordrecht, revised edition, 1981.

6. W. Diffie and M.E. Hellman. New directions in cryptography. *IEEE Trans. Inform. Theory*, 22:644–654, 1976.
7. S. Mac Lane. *Categories for the Working Mathematician*. Springer Verlag, Berlin, 1971.
8. S.C. Pohlig and M.E. Hellman. An improved algorithm for computing logarithms over GF(p) and its cryptographic significance. *IEEE Trans. Inform. Theory*, 24:106–110, 1978.
9. G.P. Purdy. A high-security log-in procedure. *Comm. of the ACM*, 17(4):442–445, 1974.
10. R.L. Rivest, A. Shamir, and Adleman L.M. A method for obtaining digital signatures and public-key cryptosystems. *Comm. ACM*, 21:120–126, 1978.