MATHEMATICS OF COMPUTATION
https://doi.org/10.1090/mcom/3382
Article electronically published on October 30, 2018

ITERATION ENTROPY

JOACHIM VON ZUR GATHEN

Abstract. We apply a common measure of randomness, the entropy, in the context of iterated functions on a finite set with n elements. For a permutation, this entropy turns out to be asymptotically (for a growing number of iterations) close to $\log_2 n$ minus the entropy of the vector of its cycle lengths. For general functions, a similar approximation holds.

1. Introduction

Arithmetic dynamics deals with discrete dynamical systems given by an (arithmetic) function on a finite set. Of particular interest are polynomials over a finite field or ring. Their iterations form a well-studied subject with many applications. In the area of cryptography, one is interested in showing some randomness properties of such iterations. For lack of entropy, iterations of a function on a finite set, beginning with a uniformly random starting value, cannot provide uniformly random values. But some specific power maps (Example 6.1) have been shown to produce pseudorandom sequences. It seems hard to extend this to more general situations.

More modestly, one tries to show certain randomness properties of such a function such as (approximate) equidistribution. The *functional graph* of f has the base set as its nodes and a directed edge from x to y if $f(x) = y$. One may consider certain graph parameters like the numbers and sizes of connected components or cycles and ask whether they are (approximately) distributed for the functions under consideration as they are for general functions.

Beginning with Flajolet and Odlyzko [10], also Flynn and Garton [11], Bellah et al. [3], Bridy and Garton [7] studied functions and polynomials from this perspective. For a uniformly random map on n points, the expected size of the giant component (an undirected component of largest size) in its functional graph is μn with $\mu \approx 0.75788$ (Flajolet and Odlyzko [10], Theorem 8 (ii)). Certain classes of polynomials over finite fields, mostly of small degree, are considered in Martins and Panario [18]. Konyagin et al. [14] presents theoretical and experimental results on maps given by random quadratic polynomials over a finite prime field. The expected size of the giant component coincides with that for random maps. However, the number of cyclic points (points on a cycle in the functional graph) is much smaller. In their experiments with the ten primes following 500 000, this is only about 885. Ostafe and Sha [20] extends some of this to certain rational functions.

Arratia and Tavaré [1] show that the distribution of cycle lengths in a random permutation is closely approximated by a Poisson distribution, with mean $1/i$ for

Received by the editor December 31, 2017, and, in revised form, May 10, 2018.
2010 *Mathematics Subject Classification.* Primary 11K45, 37P25, 68W20, 94A60.

cycle length i. Arratia et al. [2] discuss a general framework, including the factorization patterns of univariate polynomials over finite fields. Burnette and Schmutz [8] provide precise results on the distribution of cycle lengths for polynomials over a finite field, and also for rational functions. The least common multiple T of all cycle lengths is the order for a permutation and might be called the asymptotic order for a general function. They prove a lower bound $\frac{d}{2}(1+o(1))$ for $\log T$. Martins et al. [19] considers the distribution of this value, and also the number of cyclic points, for special types of polynomials.

In uniformly random permutations, the expected length of a longest cycle is τn with $\tau \approx 0.62433$; Shepp and Lloyd [24] gives an exact expression for τ, and much statistical information about the cycle length of random permutations, including the moments of the rth shortest and longest lengths, for $r = 1, 2, \ldots$. Mans et al. [17] find the average number of cyclic points for quadratic polynomials to be about the average size of a longest cycle, namely, close to $\sqrt{2n/\pi}$. For $n = 500\,000$, this evaluates to about 564. Thus there is a substantial difference of the expected largest cycle lengths between random permutations and general functions.

In this paper, we take a different route. We define a general notion of *iteration entropy*, applicable to any function from a finite set to itself. For a growing number of iterations, it approaches a limit which forms the central concept of this paper, the *asymptotic iteration entropy*. This measure abstracts from individual values like number or size of components or cycles by including them in a single parameter. It enjoys some natural properties like convexity for disjoint unions of functions. One can compare different functions under this measure. For example, when we fix the component sizes (summing to n), then permutations have a larger asymptotic iteration entropy than other functions.

For the connected components of size t containing a cycle of size c, the values $t \log_2 c$ make up the asymptotic iteration entropy (up to a factor of n). This suggests as an open question the study of this parameter, or, more generally, the joint distribution of (t, c) in functional graphs.

2. The iteration entropy

We let X be a finite set with n elements, $f\colon X \to X$ a map, and for a nonnegative integer j, $f^{(j)} = f \circ f \circ \cdots \circ f$ (with j copies of f) its jth iteration. Thus $f^{(0)} = \mathrm{id}$. For a positive integer k and $x, y \in X$, we denote as

$$N_{f,k}(x,y) = \#\{j \in \mathbb{N}\colon 0 \le j < k, f^{(j)}(x) = y\}$$

the number of times that x is mapped to y by an iterate of f, before the kth one. Then

$$\begin{aligned} \sum_{x,y \in X} N_{f,k}(x,y) &= kn, \\ N_{f,k}(x,y) &\le k \text{ for all } x, y, \end{aligned}$$

and

$$p_{f,k}(x,y) = \frac{N_{f,k}(x,y)}{kn}$$

defines a probability distribution on X^2, with all $p_{f,k}(x,y)$ at most $1/n$. The usual Shannon entropy $H^*(p_{f,k})$ of this distribution is

$$H^*(p_{f,k}) = \sum_{x,y\in X} p_{f,k}(x,y)\log_2 p_{f,k}^{-1}(x,y) = \sum_{x,y\in X} \frac{N_{f,k}(x,y)}{kn}\log_2(\frac{kn}{N_{f,k}(x,y)}).$$

Throughout this paper, we employ the convention that $z\log_2 z^{-1}$ is taken as 0 when $z = 0$. The general upper bound on the entropy implies that $0 \le H^*(p_{f,k}) \le 2\log_2 n$.

An observation by Igor Shparlinski leads to the following simplification: we subtract $\log_2 n$ from this value.

Definition 2.1. The (shifted kth) *iteration entropy* $H_{f,k}$ of f is:

$$H_{f,k} = H^*(p_{f,k}) - \log_2 n = \frac{1}{kn}\sum_{x,y\in X} N_{f,k}(x,y)\log_2(\frac{k}{N_{f,k}(x,y)}).$$

Thus

$$0 \le H_{f,k} = H^*(p_{f,k}) - \log_2 n \le \log_2 n.$$

We usually leave out the "shifted" in the following, although H is not defined as an entropy.

For any $x \in X$, we have $\sum_{y\in X} N_{f,k}(x,y) = k$, and $N_{f,k}(x,y)/k$ defines a probability distribution on the $y \in X$, with Shannon entropy

$$E(x) = \sum_{y\in X}\frac{N_{f,k}(x,y)}{k}\log_2(\frac{k}{N_{f,k}(x,y)}).$$

We might call this the *value entropy.* Then

$$H(p_{f,k}) = \frac{1}{n}\sum_{x,y\in X}\frac{N_{f,k}(x,y)}{k}\log_2(\frac{k}{N_{f,k}(x,y)}) = \frac{1}{n}\sum_{x\in X} E(x)$$

is the average value entropy.

We start with three examples.

Example 2.2. If f is the identity function on X, then

$$N_{f,k}(x,y) = \begin{cases} k & \text{if } x = y,\\ 0 & \text{otherwise,}\end{cases}$$
$$H_{f,k} = 0.$$

Example 2.3. If f is a cyclic permutation, then $N_{f,n}(x,y) = 1$ for all $x, y \in X$, $p_{f,n}(x,y) = 1/n^2$ is the uniform distribution on X^2 with Shannon entropy $2\log_2 n$, and

$$H_{f,n} = \log_2 n.$$

If k is an integer multiple of n, we have for all $x, y \in X$

$$N_{f,k}(x,y) = k/n,$$
$$H_{f,k} = \log_2 n.$$

Example 2.4. We take X to be a field with n elements, $a, b \in X$ with $a(a-1) \ne 0$, and consider the *linear congruential generator* f given by $f(x) = ax + b$. Then

$$f^{(j)}(x) = a^j x + \frac{(a^j-1)b}{a-1}$$

for all $j \geq 0$. Furthermore, let ℓ be the order of a in the multiplicative group of X. Then the functional graph of f consists of one cycle $C_0 = \{x_0\}$ with the fixed point $x_0 = -b/(a-1)$ and length $c_0 = 1$, plus $(n-1)/\ell$ cycles $C_1, \ldots, C_{(n-1)/\ell}$ of length ℓ. For a positive integer m, $k = \ell m$, $1 \leq i \leq (n-1)/\ell$ and $x, y \in C_i$, we have $N_{f,k}(x,y) = m$, and also $N_{f,k}(x_0, x_0) = k$. Thus

$$\begin{aligned} H_{f,k} &= \frac{1}{kn}\big(k \log_2 1 + \sum_{1 \leq i \leq (n-1)/\ell} \; \sum_{x,y \in C_i} m \log_2(\frac{k}{m})\big) \\ &= \frac{1}{\ell m \cdot n} \frac{(n-1)\ell^2}{\ell} m \log_2 \ell = (1 - \frac{1}{n}) \log_2 \ell. \end{aligned}$$

3. Combining functions

Given functions $f_i \colon X_i \to X_i$ on pairwise disjoint sets $X_1, \ldots, X_s$, we can combine them into a function $f \colon X \to X$ on their union $X = \bigcup_{1 \leq j \leq s} X_i$ by setting $f(x) = f_i(x)$ for $x \in X_i$. The functional graph of f is the disjoint union of those of the f_i; the same holds for the usual notion of graph as the set of pairs $(x, f(x))$. We write $n_i = \# X_i$ and $n = \sum_{1 \leq i \leq s} n_i = \# X$. The iteration entropy of f turns out to be a convex linear combination of those of the f_i.

Theorem 3.1. *For a positive integer k, we have*

$$H_{f,k} = \sum_{1 \leq i \leq s} \frac{n_i}{n} H_{f_i,k}.$$

Proof. For $x, y \in X$, we have:

$$N_{f,k}(x,y) = \begin{cases} N_{f_i,k} & \text{if } x, y \in X_i \text{ for some } i, \\ 0 & \text{otherwise.} \end{cases}$$

Thus

$$\begin{aligned} H_{f,k} =& \frac{1}{kn} \sum_{x,y \in X} N_{f,k}(x,y) \log_2 \frac{k}{N_{f,k}(x,y)} \\ =& \frac{1}{kn} \sum_{1 \leq i \leq s} \sum_{x,y \in X_i} N_{f_i,k}(x,y) \log_2 \frac{k}{N_{f_i,k}(x,y)} = \sum_{1 \leq i \leq s} \frac{n_i}{n} H_{f_i,k}. \end{aligned}$$

□

4. The asymptotic iteration entropy

The *functional graph* of an arbitrary function $f \colon X \to X$ has X as its set of nodes and a directed edge from x to y if $f(x) = y$. The underlying undirected graph consists of undirected connected components T_i each containing a cycle C_i, for various values of i. We consider these subgraphs as subsets of X, ignoring the order imposed by applications of f. The nodes in $T_i \setminus C_i$ form various *preperiod trees*. The subgraph T_i consists of C_i and all nodes in the preperiod trees attached to C_i, and we let t_i and c_i be the sizes of T_i and C_i, respectively. Figure 1 gives two explicit examples.

Definition 4.1. Let f be a function with a functional graph of component sizes t_i and cycle sizes c_i for $1 \leq i \leq s$, as above. Then

$$H_{f,\infty} = \frac{1}{n} \sum_{1 \leq i \leq s} t_i \log_2 c_i \tag{4.1}$$

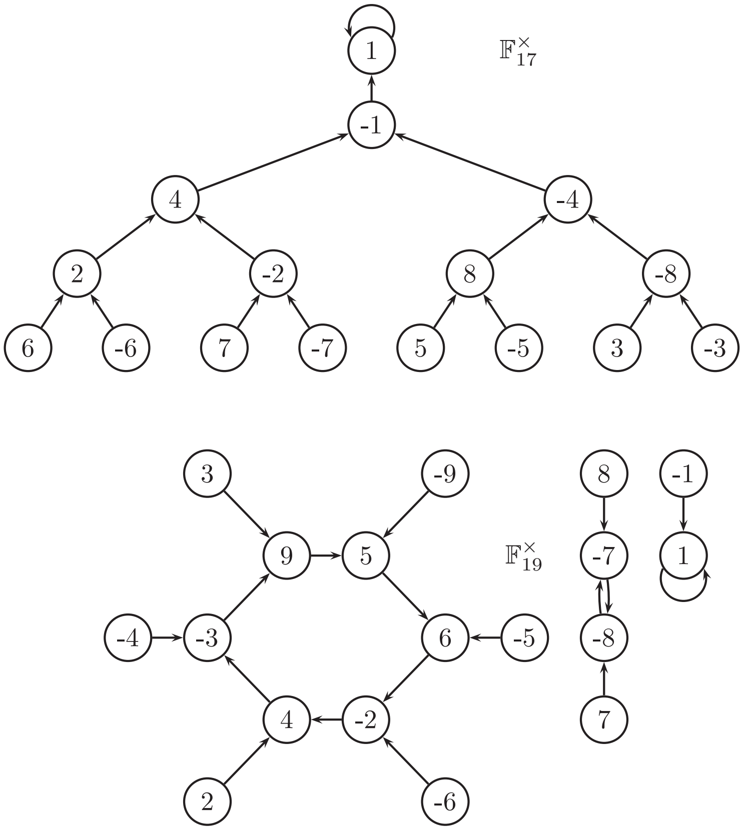

FIGURE 1. The function $x \mapsto x^2$ on the units modulo 17 and 19.

is the *asymptotic* (*shifted*) *iteration entropy* of f.

If f_i denotes the restriction of f to T_i, operating on t_i values, then

$$\begin{aligned} H_{f_i,\infty} &= \log_2 c_i, \\ H_{f,\infty} &= \sum_{1 \le i \le s} \frac{t_i}{n} H_{f_i,\infty}, \end{aligned} \tag{4.2}$$

similar to Theorem 3.1. We now show that the finite iteration entropies of a function converge to its asymptotic iteration entropy, for a growing number of iterations.

Theorem 4.2. *For $k \ge 4n \ge 77$, the following hold:*

(i)

$$|H_{f,k} - H_{f,\infty}| \le \frac{4n \log_2 k}{k}. \tag{4.3}$$

(ii) *If f is a permutation, then $c_i = t_i$ for all i and*

$$|H_{f,k} - H_{f,\infty}| \le \frac{3n \log_2 n}{k},$$

$$H_{f,\infty} + H^*(\frac{c_1}{n}, \ldots, \frac{c_s}{n}) = \log_2 n. \tag{4.4}$$

Here $H^*(c_1/n, \ldots, c_s/n)$ is the Shannon entropy of the distribution on s elements (the cycles) with probabilities $c_1/n, \ldots, c_s/n$. If f is cyclic, then $H_{f,\infty} = \log_2 n$. If f is the identity function, then $H_{f,\infty} = 0$.

(iii) *If f is a permutation and k an integer multiple of the order* $\operatorname{lcm}(c_1, \ldots, c_s)$ *of f, then*

$$H_{f,k} = H_{f,\infty}.$$

(iv) *For any f, we have $0 \le H_{f,\infty} \le \log_2 n$, and $H_{f,\infty} = \log_2 n$ if and only if f is a cyclic permutation.*

Proof. (i) We start with a single connected component X containing a single cycle $C \subseteq X$ of size c. The depth d of the functional graph on X is the maximal number of edges on a directed path within it that terminates in its first point on the cycle; this equals the maximal number of nodes on such paths minus 1. Cyclic points do not contribute to this depth. In Figure 1, we have $d = 4$ in the graph at the top, and $d = 1$ at the bottom. We consider the division with remainder

$$k = mc + r, \tag{4.5}$$

with $0 \le r < c$. The first k iterations of f send each initial value on a cycle m times around the cycle, and then r steps further. Thus if x and y are on the same cycle, then the orbit of x includes y m times, plus possibly one more time, namely if the distance (in the directed functional graph) from x to y is less than r. An off-cycle value spends at most d steps before reaching its root on the cycle, and then cycles around for at least $k - d$ steps. Thus for $x, y \in X$, there is an integer $u(x, y)$ so that

$$N_{f,k}(x,y) = \begin{cases} m + u(x,y) \text{ with } -\lceil \frac{d}{c} \rceil \le u(x,y) \le 1 & \text{if } y \in C, \\ u(x,y) \text{ with } 0 \le u(x,y) \le 1 & \text{if } y \in X \setminus C, \\ 0 & \text{otherwise,} \end{cases} \tag{4.6}$$

$$H_{f,k} = \frac{1}{kn} \sum_{\substack{x \in X \\ y \in C}} (m + u(x,y)) \log_2 \frac{k}{m + u(x,y)} \tag{4.7}$$

$$+ \frac{1}{kn} \sum_{\substack{x \in X \\ y \in X \setminus C \\ u(x,y)=1}} u(x,y) \log_2 \frac{k}{u(x,y)}. \tag{4.8}$$

Since $n \ge c$ and $k \ge 4n > 4d$, we have $m - \lceil d/c \rceil > 0$. We write

$$H_{f,\infty} = \log_2 c = \sum_{\substack{x \in X \\ y \in C}} \frac{\log_2 c}{cn}. \tag{4.9}$$

For the error bound, we first bound the difference of the contributions of $(x, y) \in X \times C$ to (4.7) and (4.9). This proceeds in two steps, first ignoring the logarithmic factors. We use

$$\delta(x,y) = \frac{m + u(x,y)}{kn} - \frac{1}{cn} = \frac{cm + cu(x,y) - k}{ckn} = \frac{cu(x,y) - r}{ckn}, \tag{4.10}$$

$$\delta(x,y) \le \frac{c - r}{ckn} \le \frac{c}{ckn},$$

$$\delta(x,y) \ge \frac{-c(d/c + 1) - r}{ckn} > \frac{-d - c - r}{ckn},$$

$$|\delta(x,y)| \le \frac{d + c + r}{ckn} \le \frac{2}{ck}, \tag{4.11}$$

since $r < c \le d + c \le n$. For the logarithms we consider, again for $x \in X$ and $y \in C$,

$$\epsilon(x,y) = \frac{\delta(x,y)}{1/cn} = cn\delta(x,y),$$

so that

$$\begin{aligned} |\epsilon(x,y)| &\le \frac{2n}{k} \le \frac{1}{2}, \\ |\log_2(1+\epsilon(x,y))| &\le |2\epsilon(x,y)| \le \frac{4n}{k}, \qquad (4.12) \\ \frac{m+u(x,y)}{kn} &= \frac{1}{cn}\cdot(1+\epsilon(x,y)). \end{aligned}$$

The difference between the contributions of $(x,y) \in X \times C$ to $H_{f,k}$ and to $H_{f,\infty}$ is

$$\begin{aligned} \alpha(x,y) &= \frac{m+u(x,y)}{kn}\log_2\frac{k}{m+u(x,y)} - \frac{\log_2 c}{cn} \\ &= (\frac{1}{cn}+\delta(x,y))\big(\log_2 c - \log_2(1+\epsilon(x,y))\big) - \frac{\log_2 c}{cn} \qquad (4.13) \\ &= \delta(x,y)\log_2 c - \frac{1}{cn}\log_2(1+\epsilon(x,y)) - \delta(x,y)\log_2(1+\epsilon(x,y)). \end{aligned}$$

From (4.11) and (4.12), we have, as in a Cauchy-Schwartz inequality,

$$\begin{aligned} |\alpha(x,y)| &\le \frac{2\log_2 c}{ck} + \frac{1}{cn}\cdot\frac{4n}{k} + \frac{2}{ck}\cdot\frac{4n}{k} \\ &= \frac{1}{ck}(2\log_2 c + 4 + \frac{8n}{k}) \le \frac{3\log_2 n}{ck}. \end{aligned}$$

In total, we find

$$\begin{aligned} |H_{f,k} - H_{f,\infty}| &\le \sum_{\substack{x\in X\\ y\in C}} |\alpha(x,y)| + \Big|\sum_{\substack{x\in X\\ y\in X\setminus C\\ u(x,y)=1}} \frac{u(x,y)}{kn}\log_2\frac{k}{u(x,y)}\Big| \qquad (4.14) \\ &\le \sum_{x\in X}(\sum_{y\in C}\frac{3\log_2 n}{ck} + \sum_{y\in X}\frac{\log_2 k}{kn}) \qquad (4.15) \\ &\le \sum_{x\in X}(\frac{3\log_2 n}{k} + \frac{\log_2 k}{k}) \qquad (4.16) \\ &\le \frac{3n\log_2 n + n\log_2 k}{k} \le \frac{4n\log_2 k}{k}. \qquad (4.17) \end{aligned}$$

We now turn to the general case, with connected components T_i of size t_i containing a cycle C_i of size c_i, and let f_i be the restriction of f to T_i, for $1 \le i \le s$. Thus $\sum_{1\le i\le s} t_i = n$, the graph of each f_i contains just one component T_i, and f is the combination of all f_i in the sense of Section 3. From Theorem 3.1 and (4.2), we have

$$H_{f,k} - H_{f,\infty} = \sum_{1\le i\le s}\frac{t_i}{n}(H_{f_i,k} - H_{f_i,\infty}).$$

Since for the single-cycle function f_i, t_i plays the role of n in (4.1) and $H_{f_i,\infty} = \log_2 c_i$, it follows from (4.3) that

$$(4.18)\quad |H_{f,k} - H_{f,\infty}| \leq \sum_{1\leq i\leq s} \frac{t_i}{n}|H_{f_i,k} - H_{f_i,\infty}| \leq \sum_{1\leq i\leq s} \frac{t_i}{n}\cdot\frac{4t_i\log_2 k}{k} \leq \frac{4n\log_2 k}{k}.$$

(ii) As in the proof of (i), we first assume f to be a cyclic permutation. Then $d = 0$ and $0 \leq u(x,y) \leq 1$ in the first line of (4.6), and $u(x,y) = 0$ in the second line, since $X = T = C$. In the equation (4.8), the last summand vanishes in (4.14) through (4.16), and the bound in (4.17) becomes $(3n\log_2 n)/k$.

Representing a general permutation as a combination of cyclic ones gives this bound also in (4.18). Furthermore, we have

$$H_{f,\infty} + H^*(\frac{c_1}{n},\ldots,\frac{c_s}{n}) = \frac{1}{n}\sum_{1\leq i\leq s} c_i\log_2 c_i + \sum_{1\leq i\leq s}\frac{c_i}{n}\log_2\frac{n}{c_i} = \log_2 n.$$

(iii) In addition to the properties in (ii), now $r = 0$ in (4.5) and $u(x,y) = 0$ in the first line of (4.6). Therefore $\delta(x,y) = 0$ in (4.10) and $\alpha(x,y) = 0$ in (4.13).

(iv) Using (4.4) and $c_i \leq t_i$ for all i, we have

$$H_{f,\infty} = \frac{1}{n}\sum_{1\leq i\leq s} t_i\log_2 c_i \leq \frac{1}{n}\sum_{1\leq i\leq s} t_i\log_2 t_i = \log_2 n - H^*(\frac{t_1}{n},\ldots,\frac{t_s}{n}) \leq \log_2 n.$$

The first inequality is strict unless $t_i = c_i$ for all i, and $H^*(t_1/n,\ldots,t_s/n) = 0$ if and only if $s = 1$ and thus $t_1 = n$. Hence $H_{f,\infty} = \log_2 n$ if and only if f is a cyclic permutation. □

The main term $H_{f,\infty}$ in Theorem 4.2 (i) is independent of k, and the error bound goes to zero with growing k. Sinkov [25] calls the expression $\sum_{1\leq i\leq s} t_i\log_2 c_i$ a *cross-entropy*, but does not discuss it further. It plays a role in modern cryptanalysis of classical ciphers, as in Lasry [16]. We are not aware of other sources for this cross-entropy.

While Definition 2.1 of the shifted iteration entropy is stated as a sum over n^2 terms, the number of summands in the asymptotic shifted iteration entropy is only the number of cycles. Of course, the different cycle lengths seem, in general, hard to compute.

If the functional graph of a function f on n elements contains a connected component of size $t_1 = \tau n$ with a cycle of length $c_1 = n^\gamma$, plus possibly other components, then

$$(4.19)\qquad H_{f,\infty} \geq \tau\gamma\log_2 n.$$

If f is, in addition, a permutation, then

$$(4.20)\qquad H_{f,\infty} \geq \tau\log_2 n + \tau\log_2\tau.$$

5. Tree surgery

How do the asymptotic iteration entropies of two distinct but closely related functions compare? We discuss three ways of slightly modifying a functional graph and their effect on the asymptotic iteration entropy.

Suppose we remove one "leaf" (most outlying node) from one of the preperiod trees. Thus we consider components and cycles $T_i \supseteq C_i$, and set $t'_i = t_i$ and $c'_i = c_i$

for all i, except that $t'_s = t_s - 1$, assuming $t_s > c_s$. We take a new function f' on a set with $n-1$ elements whose graph has these parameters. Then

$$\Delta = H_{f,\infty} - H_{f',\infty} = (\frac{1}{n} - \frac{1}{n-1}) \sum_{1 \le i < s} t_i \log_2 c_i + (\frac{t_s}{n} - \frac{t_s - 1}{n-1}) \log_2 c_s$$
$$= \frac{1}{n(n-1)}((\sum_{1 \le i < s} t_i \log_2 c_i) - (n - t_s)\log_2 c_s) = \frac{1}{n(n-1)} \sum_{1 \le i < s} t_i \log_2 \frac{c_i}{c_s}.$$

If $s = 1$, then $\Delta = 0$, and if $s \ge 2$ and C_s is a smallest cycle, then $\Delta \ge 0$.

An alternative is to enlarge C_s at the expense of T_s, by moving one node in T_s, at distance 1 from C_s, into C_s. Thus $c'_i = c_i$ and $t'_i = t_i$ for all i, except that $c'_s = c_s + 1$ and $t'_s = t_s - 1$, assuming $t_s > c_s$.

We take a new function f' on X whose graph has these parameters. Then

$$\Delta = H_{f,\infty} - H_{f',\infty} = \frac{1}{n}(t_s \log_2 c_s - (t_s - 1)\log_2(c_s + 1))$$
$$= \frac{1}{n}(t_s \log_2 \frac{c_s}{c_s + 1} + \log_2(c_s + 1)).$$

Now Δ may be positive, negative, or zero. If we replace $\log_2(1 - \frac{1}{c_s+1})$ by $-1/(c_s+1)$, then the value is positive if and only if $t_s < (c_s + 1)\log_2(c_s + 1)$.

For a more general result, we can at least compare two functions one of which is obtained from the other one by amalgamating components and cycles. We take four sequences of positive integers representing component and cycle sizes:

$$t = (t_1, \ldots, t_s),$$
$$c = (c_1, \ldots, c_s),$$
$$t' = (t'_1, \ldots, t'_r),$$
$$c' = (c'_1, \ldots, c'_r),$$

with $r < s$, $n = \sum_{1 \le i \le s} t_i = \sum_{1 \le j \le r} t'_j$, and $c_i \le t_i$ and $c'_i \le t'_i$ for all i. We say that $(t, c) \prec (t', c')$ if there exist pairwise disjoint sets $S_1, \ldots, S_r \subseteq \{1, \ldots, s\}$ such that $t'_j = \sum_{i \in S_j} t_i$ and $c'_j = \sum_{i \in S_j} c_i$ for $1 \le j \le r$. For example, if $r = s - 1$, $S_j = \{j\}$ for $j < r$, and $S_r = \{s - 1, s\}$, then we may imagine the corresponding cycles C_{s-1} and C_s cut open at one point and then joined to form one cycle, with all preperiod trees remaining attached.

Theorem 5.1. *Let f and f' be functions on a set of n elements whose functional graphs have component and cycle sizes t and c and t' and c', respectively. If $(t, c) \prec (t', c')$, then $H_{f,\infty} < H_{f',\infty}$.*

Proof. Inductively, it is sufficient to consider the example above with $r = s - 1$, $S_j = \{j\}$ for $j \le s - 2$ and $S_r = \{s - 1, s\}$. Thus $c'_j = c_j$ and $t'_j = t_j$ for $j < r$, and $t'_r = t_{s-1} + t_s$ and $c'_r = c_{s-1} + c_s$. Then

$$\begin{aligned} n(H_{f,\infty} - H_{f',\infty}) = &\sum_{1 \le i \le s} t_i \log_2 c_i \\ &- \Big(\sum_{1 \le i \le s-2} t_i \log_2 c_i + (t_{s-1} + t_s)\log_2(c_{s-1} + c_s)\Big) \\ = &\, t_{s-1}(\log_2 c_{s-1} - \log_2(c_{s-1} + c_s)) \\ &+ t_s(\log_2 c_s - \log_2(c_{s-1} + c_s)) < 0. \end{aligned}$$ □

In other words, amalgamating components as above increases the asymptotic iteration entropy.

6. Examples

We present some examples.

Example 6.1. The *power map* $x \mapsto x^e$ in a finite field or a ring $\mathbb{Z}/N\mathbb{Z}$, for fixed $e, N \geq 2$, is of cryptographic interest. Its iterations include the *power generator* for pseudorandom sequences and, with $e = 2$, the Blum-Blum-Shub and Hofheinz-Kiltz-Shoup cryptosystems (Blum et al. [4], Hofheinz et al. [13]). Friedlander et al. [12] exhibits lower bounds on the order (or period) of this function, that is, the lcm of all cycle lengths. Kurlberg and Pomerance [15] show that the maximal value (over all initial points) equals the order of e modulo M, where M is the largest divisor of the Carmichael value $\lambda(N)$ that is coprime to e. They prove a lower bound of about $N^{1/2}$ for a "Blum integer", which is the product of two primes p and q for which $p - 1$ and $q - 1$ have a large prime divisor. Sha and Hu [23] shows a similar result in finite fields, and Sha [22] for the case where N is a prime power. Pomerance and Shparlinski [21] proves several results about the number of cycles in the functional graph, among them a lower bound of $p^{5/12+o(1)}$ for infinitely many primes p. Chou and Shparlinski [9] computes the number of cyclic points, the average cycle length, and other quantities for such maps, extending the work of Vasiga and Shallit [26] on $e = 2$ (using the Extended Riemann Hypothesis). Corollary (4.19) gives a lower bound on the asymptotic iteration entropy of a permutation whose graph has a large cycle.

Example 6.2. Let n be a power of 2 and suppose that the functional graph of f contains a complete binary tree whose root is mapped under f to the only cycle in the graph, consisting of one point. Figure 1 illustrates this with the squaring function $x \mapsto x^2$ for the Fermat prime $p = 17$ on the unit group $X = \mathbb{F}_{17}^\times$ with $n = 16$ elements. Thus $s = 1$, $T_1 = X$, $t_1 = n$, $c_1 = 1$, and $H_{f,\infty} = 0$. Theorem 4.2 says that $H_{f,k} \leq (4n \log_2 k)/k$; the latter value tends to zero with growing k. Under our measure, this function exhibits "small" iteration entropy.

Example 6.3. Suppose that n is even and f has one cycle C_1 of size $n/2$, with a one-node tree attached to each point on the cycle. Thus $s = 1$, $c_1 = n/2$, $t_1 = n$. The *benzene ring* on $\mathbb{F}_{19}^\times$ at the lower left in Figure 1 is an example on $n = 12$ points. Then

$$H_{f,\infty} = \frac{t_1}{n} \log_2 \frac{n}{2} = (\log_2 n) - 1.$$

If we combine it with a single fixed point $C_2 = \{x_0\}$, then the functional graph of some quadratic function f on a field with $q = n + 1$ elements might look like this. Then

$$H_{f,\infty} = \frac{q-1}{q} \log_2((q-1)/2) + \frac{1}{q} \log_2 1 \approx (1 - \frac{1}{q}) \cdot (\log_2 q - 1).$$

Under our measure, both are "large" asymptotic iteration entropies.

Example 6.4. In Boppré et al. [5], the *ElGamal function* $f : x \mapsto g^x$ on $\mathbb{F}_p$ is studied, where p is a prime number and g is a generator of the multiplicative group of $\mathbb{F}_p$. This function occurs in some cryptographic protocols. For the two primes 1009 and 10 009, there are 288 and 3312 generators, respectively. Figures 2

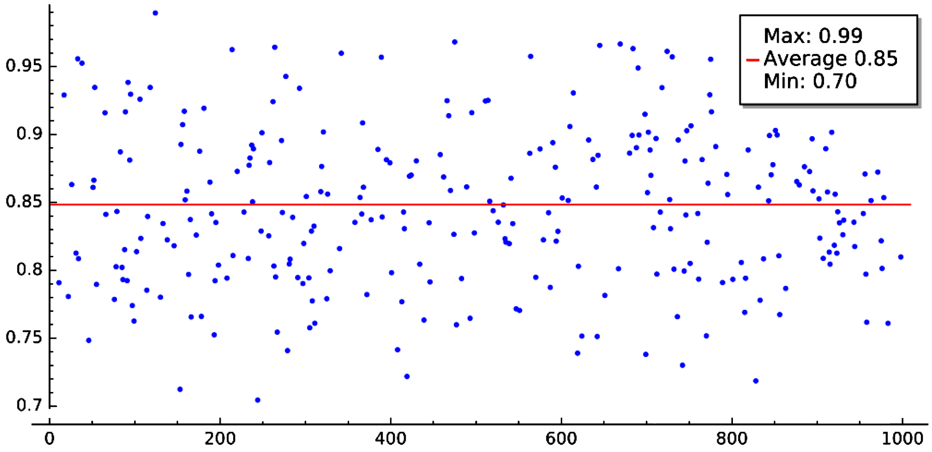

FIGURE 2. The asymptotic iteration entropies for the 288 ElGamal functions on $\mathbb{F}_{1009}$. The red line at 0.85 indicates the average.

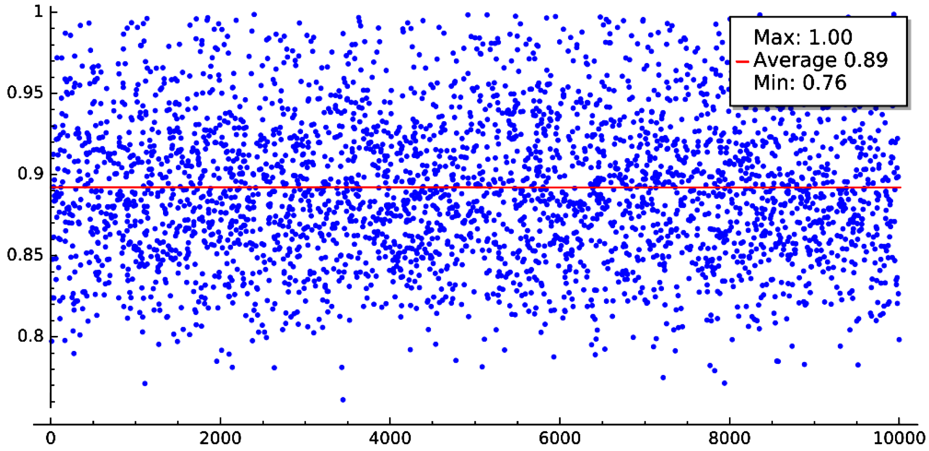

FIGURE 3. The asymptotic iteration entropies for the 3312 ElGamal functions on $\mathbb{F}_{10009}$. The red line near 0.89 indicates the average.

and 3 show the asymptotic iteration entropies of all these functions, normalized as $H_{f,\infty}/\log_2 p$, so that all values lie between 0 and 1. The values lie, on average, about 10 to 15% below the maximal value of 1. The variances were also computed but are too small to be shown.

7. Open questions

- What is the average asymptotic iteration entropy of a random permutation? Or a random function? Equation (4.20) provides a lower bound for individual permutations. Is this, with the proper value of $\tau \approx 0.62433$, also a (lower or upper) bound on the average? The results of Arratia and Tavaré [1] may help to answer this question. What is the average value of

$\max\{t \log_2 c\}$ for random functions, where t runs through the component sizes and c is the size of the component's cycle? The joint distribution of (t, c) does not seem to have been studied. The average size μn with $\mu \approx 0.75788$ (Flajolet and Odlyzko [10]) of the giant component might be a lower bound, except that components with a fixed point ($c = 1$) would have to be ruled out.

- The linear congruential generator of Example 2.4 is well known to be insecure (see Boyar [6]) and hence does not provide pseudorandom values by iteration. For large ℓ, say $\ell = n - 1$, its asymptotic iteration entropy is close to the maximal value of $\log_2 n$. Thus large iteration entropy does not imply pseudorandomness. Is the converse true in some sense?
- What is the relation of the (asymptotic) iteration entropy to usual notions of random generation? A function on a finite set contains only a finite amount of information (or Shannon entropy) and its iterates, from a uniformly random starting value, do not generate a statistically random sequence of elements. But one may ask for a modest amount of equidistribution (see Boppré [5] for the ElGamal function, and other works cited earlier) or whether some form of pseudorandomness is obtainable. The Diffie-Hellman problem in a finite cyclic group G of order d is to find g^c with $g^c = g^{ab}$, given g^a and g^b. Any function on G translates into a function on the exponent group $\mathbb{Z}_d$, with the same cycle structure. The Diffie-Hellman problem becomes the trivial task of finding c with $c = ab$, given a and b. This illustrates the general observation that the computational difficulty may lie in the presentation (here: group elements vs. exponents). But for the ElGamal function, the translation to the exponent group does not seem to simplify the issue, say for finding a preimage or for understanding randomness properties. Conversely, does pseudorandomness imply that the function is a permutation? For example, the squaring function on the set of quadratic residues with Jacobi symbol 1 modulo a special type of RSA modulus is used in Hofheinz et al. [13]; it is pseudorandom under the assumption that such moduli are hard to factor, it is a permutation, and in general not cyclic.

Acknowledgments

The images for Figures 2 and 3 were produced by Lucas Perin, whose help is much appreciated. Many thanks go to Alina Ostafe and to Igor Shparlinski for valuable suggestions, corrections, and pointers to the literature, and to an anonymous referee.

References

[1] R. Arratia and S. Tavaré, *The cycle structure of random permutations*, Ann. Probab. **20** (1992), no. 3, 1567–1591. MR1175278

[2] R. Arratia, A. D. Barbour, and S. Tavaré, *On random polynomials over finite fields*, Math. Proc. Cambridge Philos. Soc. **114** (1993), no. 2, 347–368, DOI 10.1017/S0305004100071620. MR1230136

[3] E. Bellah, D. Garton, E. Tannenbaum, and N. Walton, *A probabilistic heuristic for counting components of functional graphs of polynomials over finite fields*, Involve **11** (2018), no. 1, 169–179, DOI 10.2140/involve.2018.11.169. MR3681355

[4] L. Blum, M. Blum, and M. Shub, *A simple unpredictable pseudorandom number generator*, SIAM J. Comput. **15** (1986), no. 2, 364–383, DOI 10.1137/0215025. MR837589

[5] L. Boppré, J. von zur Gathen, L. Perin, and A. Zumalacárregui, *Sidon sets and statistics of the ElGamal function*, 2017. Preprint, `https://arxiv.org/abs/1708.04395`.

[6] J. Boyar, *Inferring sequences produced by pseudo-random number generators*, J. Assoc. Comput. Mach. **36** (1989), no. 1, 129–141, DOI 10.1145/58562.59305. MR1072416

[7] A. Bridy and D. Garton, *Dynamically distinguishing polynomials*, Res. Math. Sci. **4** (2017), Paper No. 13, 17, DOI 10.1186/s40687-017-0103-3. MR3669394

[8] C. Burnette and E. Schmutz, *Periods of iterated rational functions*, Int. J. Number Theory **13** (2017), no. 5, 1301–1315, DOI 10.1142/S1793042117500713. MR3639698

[9] W.-S. Chou and I. E. Shparlinski, *On the cycle structure of repeated exponentiation modulo a prime*, J. Number Theory **107** (2004), no. 2, 345–356, DOI 10.1016/j.jnt.2004.04.005. MR2072394

[10] P. Flajolet and A. M. Odlyzko, *Random mapping statistics*, Advances in cryptology—EUROCRYPT '89 (Houthalen, 1989), Lecture Notes in Comput. Sci., vol. 434, Springer, Berlin, 1990, pp. 329–354, DOI 10.1007/3-540-46885-4_34. MR1083961

[11] R. Flynn and D. Garton, *Graph components and dynamics over finite fields*, Int. J. Number Theory **10** (2014), no. 3, 779–792, DOI 10.1142/S1793042113501224. MR3190008

[12] J. B. Friedlander, C. Pomerance, and I. E. Shparlinski, *Period of the power generator and small values of Carmichael's function*, Math. Comp. **70** (2001), no. 236, 1591–1605, DOI 10.1090/S0025-5718-00-01282-5. Corrigendum in **71** (2002), 1803–1806. MR1836921

[13] D. Hofheinz, E. Kiltz, and V. Shoup, *Practical chosen ciphertext secure encryption from factoring*, J. Cryptology **26** (2013), no. 1, 102–118, DOI 10.1007/s00145-011-9115-0. MR3016825

[14] S. V. Konyagin, F. Luca, B. Mans, L. Mathieson, M. Sha, and I. E. Shparlinski, *Functional graphs of polynomials over finite fields*, J. Combin. Theory Ser. B **116** (2016), 87–122, DOI 10.1016/j.jctb.2015.07.003. MR3425238

[15] P. Kurlberg and C. Pomerance, *On the periods of the linear congruential and power generators*, Acta Arith. **119** (2005), no. 2, 149–169, DOI 10.4064/aa119-2-2. MR2167719

[16] G. Lasry, *A Methodology for the Cryptanalysis of Classical Ciphers with Search Metaheuristics*, Universität Kassel, Germany, 2017, Ph.D. Thesis.

[17] B. Mans, M. Sha, I. E. Shparlinski, and D. Sutantyo, *On functional graphs of quadratic polynomials*, 2017, `arXiv:1706.04734`, to appear in Experimental Mathematics.

[18] R. S. V. Martins and D. Panario, *On the heuristic of approximating polynomials over finite fields by random mappings*, Int. J. Number Theory **12** (2016), no. 7, 1987–2016, DOI 10.1142/S1793042116501219. Erratum on pp 2041-2042. MR3544423

[19] R. S. V. Martins, D. Panario, C. Qureshi, and E. Schmutz, *Periods of iterations of mappings over finite fields with restricted preimage sizes*, AofA 2018, LIPIC (Leibniz International Proceedings in Informatics), Vol. 110, Dagstuhl Publishing, 2018. arXiv:1701.09148

[20] A. Ostafe and M. Sha, *Counting dynamical systems over finite fields*, Dynamics and numbers, Contemp. Math., vol. 669, Amer. Math. Soc., Providence, RI, 2016, pp. 187–203, DOI 10.1090/conm/669/13429. MR3546669

[21] C. Pomerance and I. E. Shparlinski, *Connected components of the graph generated by power maps in prime finite fields*, Integers **18A** (2018), Paper No. A16, 8. MR3777538

[22] M. Sha, *On the cycle structure of repeated exponentiation modulo a prime power*, Fibonacci Quart. **49** (2011), no. 4, 340–347. MR2852007

[23] M. Sha and S. Hu, *Monomial dynamical systems of dimension one over finite fields*, Acta Arith. **148** (2011), no. 4, 309–331, DOI 10.4064/aa148-4-1. MR2800698

[24] L. A. Shepp and S. P. Lloyd, *Ordered cycle lengths in a random permutation*, Trans. Amer. Math. Soc. **121** (1966), 340–357, DOI 10.2307/1994483. MR0195117

[25] A. Sinkov, *Elementary cryptanalysis*, 2nd ed., Anneli Lax New Mathematical Library, vol. 22, Mathematical Association of America, Washington, DC, 2009. A mathematical approach; Revised, updated and with a preface by Todd Feil. MR2530836

[26] T. Vasiga and J. Shallit, *On the iteration of certain quadratic maps over* GF(p), Discrete Math. **277** (2004), no. 1-3, 219–240, DOI 10.1016/S0012-365X(03)00158-4. MR2033734

B-IT, Universität Bonn, Endenicher Allee 19c, 53115 Bonn, Germany
Email address: `gathen@bit.uni-bonn.de`