Appendices

A. Generalized entropies as regularizers

We present information measures which are often used as regularizers.

Example A.1 (Shannon entropy)**.** *Let* $R(x,y) = -H_S(x,y)$*, where*

$$H_S(x,y) = -x\log x - y\log y.$$

Then $H_S(x,1-x)$ *is the Shannon entropy of a probability distribution* $(x,1-x)$ *and*

$$\begin{aligned} r(x) = R(x,1-x) &= -H_S(x,1-x) \\ &= x\log x + (1-x)\log(1-x). \end{aligned}$$

From

$$r'(x) = \log\frac{x}{1-x}$$

we observe that $r \in \mathbf{SSC}$.[5]

Example A.2 (Arimoto entropies)**.** *We consider the class of Arimoto entropies (Csiszár, 2008), that is functions defined as*

$$H_\eta(x,y) = \eta(x) + \eta(y),$$

where $\eta \in \mathcal{C}^2((0,1))$ *is a concave function.*[6]

We define

$$R(x,y) = -H_\eta(x,y).$$

Then, by its definition, $R \in \mathcal{C}^2((0,1)^2)$ *and* $R(y,x) = R(x,y)$*. Moreover,*

$$r(x) = R(x,1-x) = -H_\eta(x,1-x) = -\eta(x) - \eta(1-x),$$

$$r'(x) = -\eta'(x) + \eta'(1-x) \text{ and } r''(x) = -\eta''(x) - \eta''(1-x).$$

Thus, r *is convex and the limit* $\lim_{x\to 1^-}\eta'(x)$ *is finite. Therefore, the condition for* $R = -H_\eta \in \mathbf{SSC}$ *is steepness of* η *at zero:*

$$\lim_{x\to 0^+}\eta'(x) = \infty. \tag{11}$$

Hence, $R = -H_\eta \in \mathbf{SSC}$ *if and only if* η *satisfies* (11)*.*

Several well-known regularizers are given by (negative) Arimoto entropies satisfying (11). For instance, the Shannon entropy from Example A.1 is an Arimoto entropy for $\eta(x) = -x\log x$, as well as log-barrier regularizer obtained from $\eta(x) = \log x$. Another widely used (especially in statistical physics) example of Arimoto entropy is the Havrda-Charvát-Tsallis entropy.[7]

Example A.3 (Havrda-Charvát-Tsallis entropies)**.** *The Havrda-Charvát-Tsallis entropy for* $q \in (0,\infty)$ *is defined as*

$$H_q(x,y) = \begin{cases} \frac{1}{1-q}(x^q + y^q - 1) & \text{for } q \neq 1 \\ H_S(x,y) & \text{for } q = 1 \end{cases}. \tag{12}$$

H_q *is an Arimoto entropy for* $\eta(x) = \frac{1}{1-q}\left(x^q - \frac{1}{2}\right)$*, satisfying* (11) *for* $0 < q < 1$*. If* $R(x,y) = -H_q(x,y)$ *then*

$$r(x) = R(x,1-x) = \frac{1}{q-1}(x^q + (1-x)^q - 1),$$

$$r'(x) = \frac{q}{q-1}\left(x^{q-1} - (1-x)^{q-1}\right),$$

and $r \in \mathbf{SSC}$ *for* $q \in (0,1]$*.*

For $q > 1$ *the Havrda-Charvát-Tsallis entropy does not satisfy* (11) *and, consequently, the regularizer* R *emerging from the Havrda-Charvát-Tsallis entropy does not belong to* $\mathbf{SSC}$*. Standard non-example is Euclidean norm, which we get from* (12) *when* $q = 2$*. Then*

$$r(x) = R(x,1-x) = -H_2(x,1-x) = x^2 + (1-x)^2 - 1$$

and as $\lim_{x\to 0^+} r'(x) = -2$*,* R *doesn't belong to* $\mathbf{SSC}$*.*

Evidently there exist functions which are not Arimoto entropies but also generate regularizers that belong to $\mathbf{SSC}$, one of them being the Rényi entropy of order $q < 1$.

Example A.4 (Rényi entropies)**.** *The Shannon entropy represents an expected mean of individual informations of the form* $I_k = -\log p_k$*. Rényi (Rényi, 1961) introduced alternative information measures, namely generalized means* $g^{-1}(\sum p_k g(I_k))$*, where* g *is a continuous, strictly monotone function.Then, the Rényi entropy of order* $q \neq 1$ *correspond to* $g(x) = \exp((1-q)x)$*, namely:*

$$H_q^R(x,y) = \begin{cases} \frac{1}{1-q}\log\left(x^q + y^q\right), & \text{for } q \neq 1 \\ H_S(x,y), & \text{for } q = 1 \end{cases}.$$

As the variables x *and* y *are not separable, this is not an Arimoto entropy. However, for* $R(x,y) = -H_q^R(x,y)$*,* $R \in$

[5] By substituting the (negative) Shannon Entropy as R into (4) we obtain the Multiplicative Weights Update algorithm.

[6] In the decision theory Arimoto entropies correspond to separable Bregman scores (Grünwald & Dawid, 2004).

[7] This entropy (called also entropy of degree q) was first introduced by Havrda and Charvát (Havrda & Charvát, 1967) and used to bound probability of error for testing multiple hypotheses. In statistical physics it is known as Tsallis entropy, referring to (Tsallis, 1988).

$\mathcal{C}^2((0,1)^2)$ *and* $R(y,x) = R(x,y)$. *Moreover,*

$$r(x) = R(x, 1-x) = -H_q^R(x, 1-x) = \frac{1}{q-1}\log\left(x^q + (1-x)^q\right)$$

and

$$r'(x) = \frac{q}{q-1} \cdot \frac{x^{q-1} - (1-x)^{q-1}}{x^q + (1-x)^q}.$$

Thus, for $q \in (0,1)$ *we know that* $r''(x) > 0$ *on* $(0,1)$ *and* $\lim_{x\to 0^+} r'(x) = -\infty$. *Because* $H_1^R = H_S$ *we infer that* $R \in \mathbf{SSC}$ *for* $q \in (0,1]$.

B. Regularity of log-barrier dynamics

To understand better the phenomenon discussed in Section 8, let us investigate regularity of $f_{a,b}$. Nice properties of interval maps are guaranteed by the negative Schwarzian derivative. Let us recall that the Schwarzian derivative of f is given by the formula

$$Sf = \frac{f'''}{f'} - \frac{3}{2}\left(\frac{f''}{f'}\right)^2.$$

A "metatheorem" states that almost all natural noninvertible interval maps have negative Schwarzian derivative. Note that, by Lemma 3.2.ii, if $a \le -\Psi'(b)$ then $f_{a,b}$ is a homeomorphism, so we should not expect negative Schwarzian derivative for that case. For maps with negative Schwarzian derivative each attracting or neutral periodic orbit has a critical point in its immediate basin of attraction. Thus, if we show that the Schwarzian derivative is negative, then we will know that all periodic orbits can be find by studying behavior of critical points of $f_{a,b}$. Therefore, we want to show that $Sf_{a,b} < 0$ for sufficiently large a for $f_{a,b}$ determined by log-barrier regularizer.

In general, computation of Schwarzian derivative may be very complicated. However, there is a useful formula

$$S(h \circ f) = (f')^2\,((Sh) \circ f) + Sf. \tag{13}$$

The function $f_{a,b}$ is given by (7). Consider

$$g(x) := (\Psi \circ f_{a,b})(x) = \Psi(x) + a(x-b).$$

By (13) we have that

$$Sg = (f'_{a,b})^2\,((S\Psi) \circ f_{a,b}) + Sf_{a,b}.$$

At the same time

$$Sg(x) = S(\Psi(x) + a(x-b)).$$

Therefore,

$$(f'_{a,b}(x))^2\,((S\Psi) \circ f_{a,b}(x)) + Sf_{a,b}(x) = S(\Psi(x) + a(x-b)). \tag{14}$$

Direct computations yield

$$S\Psi(x) = \frac{6}{\left(x^2 + (1-x)^2\right)^2} > 0$$

and

$$S(\Psi(x) + a(x-b)) = \frac{6\left[1 - a(x^4 + (1-x)^4)\right]}{\left[x^2 + (1-x)^2 - ax^2(1-x)^2\right]^2}.$$

Observe that $x^4 + (1-x)^4 \geqslant \frac{1}{8}$ for all $x \in [0,1]$. Thus $1 - a(x^4 + (1-x)^4) \leqslant 1 - \frac{a}{8}$, and

$$S(\Psi(x) + a(x-b)) < 0 \quad \text{for } a > 8. \tag{15}$$

Therefore, $Sf_{a,b} < 0$ for $a > 8$. Moreover,

$$\max_{b\in[0,1]} \Psi'(b) = \Psi'(1/2) = -8.$$

Thus,

$$Sf_{a,b}(x) < 0 \text{ for all } a > -\Psi'(b) \ge 8.$$

C. Proofs

Proof of Proposition 3.1. Due to the condition $R(x,y) = R(y,x)$, we have that $\frac{\partial R}{\partial x}(x, 1-x) = \frac{\partial R}{\partial y}(1-x, x)$. Thus, if $\varphi(x) = 1 - x$, then:

$$\begin{aligned}\Psi(1-x) &= -[r(\varphi(x))]' = -\frac{\partial R}{\partial x}(1-x, x) + \frac{\partial R}{\partial y}(1-x, x)\\ &= -\frac{\partial R}{\partial y}(x, 1-x) + \frac{\partial R}{\partial x}(x, 1-x)\\ &= r'(x) = -\Psi(x).\end{aligned}$$

This implies (i). Moreover $\Psi'(x) = -r''(x) < 0$. Thus, Ψ is decreasing.

$$\lim_{x\to 0^+} \Psi(x) = -\lim_{x\to 0^+} r'(x) = \infty.$$

From (i) we obtain that $\lim_{x\to 1^-} \Psi(x) = -\infty$. □

Proof of Lemma 3.2. We obtain (i) directly from (7) and the fact that Ψ is decreasing.

Ψ is a homeomorphism, thus if $y = \Psi(x)$ for some $x \in (0,1)$, then

$$y = \Psi(\Psi^{-1}(y)) = -\Psi(1 - \Psi^{-1}(y)).$$

Hence,

$$\Psi^{-1}(-y) = 1 - \Psi^{-1}(y).$$

Now let $x \in (0,1)$. Then

$$\begin{aligned}(\varphi \circ f_{a,b})(x) &= 1 - f_{a,b}(x) = 1 - \Psi^{-1}\left(\Psi(x) + a(x-b)\right) \\ &= \Psi^{-1}\left(-\Psi(x) - a(x-b)\right) \\ &= \Psi^{-1}\left(\Psi(1-x) + a\big((1-x)-(1-b)\big)\right) \\ &= (f_{a,1-b} \circ \varphi)(x),\end{aligned}$$

and (ii) follows.

By (i), $f_{a,b}(x) > x$ for $x \in (0,b)$ and $f_{a,b}(x) < x$ for $x \in (b,1)$. Therefore, there exists $0 < \delta_1 < \min\{b, 1-b\}$ such that $|\frac{1}{2} - x| > |\frac{1}{2} - f_{a,b}(x)|$ for $x \in (0,1)\setminus(\delta_1, 1-\delta_1)$. There exists also $\delta_2 > 0$ such that $f_{a,b}([\delta_1, 1-\delta_1]) \subset (\delta_2, 1-\delta_2)$. Set $\delta = \min\{\delta_1, \delta_2\}$. Then, the interval $I = [\delta, 1-\delta]$ is invariant.

To complete the proof of (iii) we need to show that I is attracting. Assume that $x \in (0,1) \setminus I$ is such that its $f_{a,b}$-trajectory never enters I. Since $\delta \le \delta_1$, the distance between $f^n_{a,b}(x)$ and I (that is, $d_I(f^n_{a,b}(x))$, where $d_I(z) = \delta - z$ for $z \in [0,\delta]$ and $d_I(z) = z - (1-\delta)$ for $z \in [1-\delta, 1]$) is decreasing and $\delta < f(\delta) < 1-\delta$. Sequence $d_I(f^n_{a,b}(x))$ is decreasing and bounded from below by 0, so it is convergent to some $\epsilon \ge 0$. Therefore, the ω-set of the trajectory of x is a non-empty subset of $d_I^{-1}(\{\epsilon\}) = I_\epsilon = \{\delta - \epsilon, 1 - \delta + \epsilon\}$. However, no non-empty subset of I_ϵ can be invariant (and thus, can be an ω-set of a trajectory), because $\delta - \epsilon \le \delta_1$ and thus $f_{a,b}(I_\epsilon) \subset (\delta - \epsilon, 1 - \delta + \epsilon)$, and $f_{a,b}(I_\epsilon) \cap I_\epsilon = \emptyset$. By this contradiction, such x does not exist, thus I is globally attracting.

□

Proof of Theorem 4.2. Fix $x_0 \in (0,1)$ and let $x_k = f^k_{a,b}(x_0)$.

From (7) we get by induction that

$$x_n = f_{a,b}(x_{n-1}) = \Psi^{-1}\left(\Psi(x_0) + a\left(\sum_{k=0}^{n-1}(x_k - b)\right)\right). \tag{16}$$

By Lemma 3.2.iii there is $\delta > 0$ such that there exists a closed, globally absorbing and invariant interval $I \subset (\delta, 1-\delta)$. Thus, for sufficiently large n

$$\delta < x_n = \Psi^{-1}\left(\Psi(x_0) + a\left(\sum_{k=0}^{n-1}(x_k - b)\right)\right) < 1 - \delta.$$

Ψ is decreasing, thus

$$\Psi(\delta) > \Psi(x_0) + a\left(\sum_{k=0}^{n-1}(x_k - b)\right) > \Psi(1-\delta).$$

Therefore

$$\frac{1}{an}\left(\Psi(\delta) - \Psi(x_0)\right) > \frac{1}{n}\sum_{k=0}^{n-1} x_k - b > \frac{1}{an}\left(\Psi(1-\delta) - \Psi(x_0)\right),$$

so

$$\left|\frac{1}{n}\sum_{k=0}^{n-1} x_k - b\right| < \frac{1}{an}\max\{|\Psi(\delta) - \Psi(x_0)|, |\Psi(1-\delta) - \Psi(x_0)|\}.$$

Thus, (8) follows. □

Proof of Lemma 6.1. Without loss of generality we can assume that $x < y$. Then

$$\begin{aligned}&\frac{y-x}{2}g'\left(\frac{x+y}{2}\right) - \int_x^{\frac{x+y}{2}} g'(t)\,\mathrm{dt} = \int_x^{\frac{x+y}{2}}\int_t^{\frac{x+y}{2}} g''(s)\,\mathrm{ds}\,\mathrm{dt} \\ &> \int_{\frac{x+y}{2}}^{y}\int_{\frac{x+y}{2}}^{t} g''(s)\,\mathrm{ds}\,\mathrm{dt} = \int_{\frac{x+y}{2}}^{y} g'(t)\,\mathrm{dt} - \frac{y-x}{2}g'\left(\frac{x+y}{2}\right),\end{aligned}$$

where the inequality follows from the fact that $g''(s)$ is smaller in the latter region while the integration is over the set of the same size. Therefore,

$$g'\left(\frac{x+y}{2}\right) > \frac{1}{y-x}\int_x^y g'(t)\,\mathrm{dt} = \frac{g(y) - g(x)}{y-x},$$

which completes the proof of lemma. □

Proof of Theorem 6.2. In order to prove this theorem it is sufficient to show that $f_{a,b}$ doesn't have periodic orbits of period 2.

Suppose that $\{x_0, x_1\} \in (0,1)$ is a periodic orbit of $f_{a,b}$ of period 2.

$$\frac{x_0 + x_1}{2} = b.$$

We have that $x_1 = \Psi^{-1}\left(\Psi(x_0) + a(x_0 - b)\right)$, and therefore,

$$\Psi(x_1) = \Psi(x_0) + a(x_0 - b).$$

Thus, $\Psi(x_1) - \Psi(x_0) = -\frac{a}{2}(x_1 - x_0)$, or equivalently

$$a = -2 \cdot \frac{\Psi(x_1) - \Psi(x_0)}{x_1 - x_0}. \tag{17}$$

By Lemma 6.1

$$\Psi'(b) > \frac{\Psi(x_1) - \Psi(x_0)}{x_1 - x_0} = -\frac{a}{2}, \tag{18}$$

but the point b is attracting if and only if $\Psi'(b) < -\frac{a}{2}$, which contradicts the inequality (18). Therefore, f has no periodic point of period 2.

Now, by (Block & Coppel, 2006), Chapter VI, Proposition 1, every trajectory of f converges to a fixed point.

□

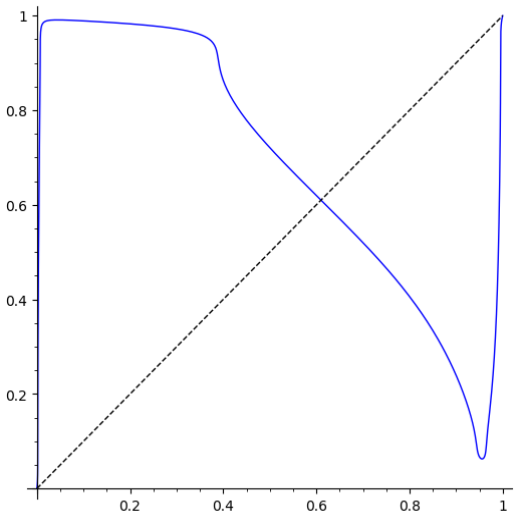


Figure 3. Graph of $f_{a,b}$ for $a = 3.25$, $b = 0.61$ generated by the regularizer $r(x) = (1-x)\log(1-x) + x\log x - 0.4167\log(-x^2+x+0.11)$.

Proof of Proposition 6.4. Let us take $b = 0.61$ and $a = 3.25$ (see the graph of $f_{a,b}$ in Figure 3). Set

$$\begin{aligned}\xi(x) &:= \Psi(x) + a(x-b)\\ &= \log(1-x) - \log x + 0.4167\left(\frac{1}{1.1-x} - \frac{1}{x+0.1}\right)\\ &\quad + 3.25\cdot(x-0.61).\end{aligned}$$

Since $a < 3.282596521095308 \approx -2\Psi'(b)$, the fixed point b is attracting.

To show that $f_{a,b}$ is chaotic, we will prove that $f_{a,b}$ has a periodic point of period 6. With this aim, we will show that $(f^2)^3(x) < x < f^2(x)$ for any $x \in [0.9559, 0.956]$. We start by showing that $\xi(x)$ is monotone on $[0.9559, 0.956]$. Formula for the derivative of $\xi(x)$ is

$$\begin{aligned}\xi'(x) = -\frac{1}{1-x} - \frac{1}{x} + 0.4167\cdot\left(\frac{1}{(1.1-x)^2} + \frac{1}{(x+0.1)^2}\right)\\ + 3.25.\end{aligned}$$

Set $z = (x-0.5)^2$. Then $\xi'(x) = 0$ if and only if $g(z) = 0$, where

$$g(z) = 3.25\cdot z^3 - 1.3191\cdot z^2 + 0.377874\cdot z - 0.050706.$$

We have

$$g'(z) = 9.75\cdot z^2 - 2.6382\cdot z + 0.377874,$$

and the discriminant of this quadratic polynomial is negative. Therefore, g has only one zero (approximately 0.2077259768645677), so ξ' has only two zeros, symmetric with respect to 0.5. Thus, as

$$\xi'(0.9559) \approx -0.03051955745677404,$$

$$\xi'(0.956) \approx -0.05413532613604133$$

there is no zero of ξ' between these two points. Moreover, those computations give us an approximation to both zeros of ξ': 0.0442303467050842 and 0.9557696532949158.

Now we look at the first six images of $[0.9559, 0.956]$:

$$\begin{aligned}\Psi(0.063) &\approx 0.5449390463486314\\ \xi(0.956) &\approx 0.5450794481395858\\ \xi(0.9559) &\approx 0.5450836794177281\\ \Psi(0.062) &\approx 0.5458384284441133\end{aligned}$$

$$\begin{aligned}\Psi(0.991) &\approx -1.26049734857964\\ \xi(0.062) &\approx -1.235161571555887\\ \xi(0.063) &\approx -1.232810953651368\\ \Psi(0.99) &\approx -1.189231609934426\end{aligned}$$

$$\begin{aligned}\Psi(0.52) &\approx -0.03369120600501601\\ \xi(0.991) &\approx -0.02224734857964017\\ \xi(0.99) &\approx 0.04576839006557365\\ \Psi(0.47) &\approx 0.05052025169168718\end{aligned}$$

$$\begin{aligned}\Psi(0.76) &\approx -0.4116261583651984\\ \xi(0.47) &\approx -0.4044797483083129\\ \xi(0.52) &\approx -0.3261912060050159\\ \Psi(0.69) &\approx -0.3112461911278587\end{aligned}$$

$$\begin{aligned}\Psi(0.54) &\approx -0.06732925721803665\\ \xi(0.69) &\approx -0.05124619112785883\\ \xi(0.76) &\approx 0.07587384163480165\\ \Psi(0.45) &\approx 0.08411125490271053\end{aligned}$$

$$\begin{aligned}\Psi(0.8) &\approx -0.4602943611198909\\ \xi(0.45) &\approx -0.4358887450972894\\ \xi(0.54) &\approx -0.2948292572180365\\ \Psi(0.65) &\approx -0.2486392084062237.\end{aligned}$$

We have $f_{a,b}(x) = \Psi^{-1}(\xi(x))$, so $\Psi(f_{a,b}(x)) = \xi(x)$. Write $\langle x, y\rangle$ for $[x, y]$ or $[y, x]$. If $\langle \xi(x), \xi(y)\rangle \subset \langle \Psi(z), \Psi(w)\rangle$ and ξ is monotone on $\langle x, y\rangle$, then $\langle f_{a,b}(x), f_{a,b}(y)\rangle \subset \langle z, w\rangle$. Thus, the computations show that

$$f_{a,b}^2([0.9559, 0.956]) \subset [0.99, 0.991]$$

and

$$f_{a,b}^6([0.9559, 0.956]) \subset [0.65, 0.8].$$

Therefore, for any $x \in [0.9559, 0.956]$ we have

$$(f_{a,b}^2)^3(x) < x < (f_{a,b}^2)(x),$$

so by theorem from (Li et al., 1982), $f_{a,b}^2$ has a periodic point of period 3 and $f_{a,b}$ has a periodic point of period 6. Thus, because $f_{a,b}$ has a periodic point of period that is not a power of 2, the topological entropy $h(f_{a,b})$ is positive (see (Misiurewicz, 1979)) and it is Li-Yorke chaotic. □

Proof of Theorem 7.1. By Lemma 3.2.ii, without loss of generality, we may assume $b \in (0, \frac{1}{2})$. We will show that there exists $x_0 \in (0,1)$ such that $f_{a,b}^3(x_0) < x_0 < f_{a,b}(x_0)$.

Fix $a > 0$ and $b, x \in (0,1)$. We set $x_n = f_{a,b}^n(x)$, then formula (16) holds. Hence $f_{a,b}(x) > x$ if and only if $x < b$ and, because Ψ^{-1} is decreasing, $f_{a,b}^3(x) < x$ is equivalent to $x + f_{a,b}(x) + f_{a,b}^2(x) > 3b$.

From the fact that $b \in (0, \frac{1}{2})$ we have that $3b - 1 < b$. So we can take $x_0 > 0$ such that $3b - 1 < x_0 < b$. Then $f_{a,b}(x_0) > x_0$. Moreover

$$\lim_{a\to\infty} f_{a,b}(x_0) = \lim_{a\to\infty} \Psi^{-1}\left(\Psi(x_0) + a(x_0 - b)\right) = 1.$$

Thus, since $3b - x_0 < 1$, there exists $a_b > 0$ such that if $a > a_b$, then $f_{a,b}(x_0) > 3b - x_0$, so $x_0 + f_{a,b}(x_0) + f_{a,b}^2(x_0) > 3b$. Hence, if $a > a_b$, then $f_{a,b}^3(x_0) < x_0$.

Now we conclude that $f_{a,b}$ has a periodic point of period 3 for $a > a_b$, from theorem from (Li et al., 1982), which implies that if $f^n(x) < x < f(x)$ for some odd $n > 1$, then f has a periodic point of period n.

□

Proof of 7.3. By Lemma 3.2.ii the maps f_a and φ commute. Set $g_a = \varphi \circ f_a = f_a \circ \varphi$. Since φ is an involution, we have $g_a^2 = f_a^2$. We show that the dynamics of f_a is simple, no matter how large a is.

We aim to find fixed points and points of period 2 of f_a and g_a. Clearly,

$$f_a(0) = 0, \quad f_a(1) = 1, \quad g_a(0) = 1, \quad g_a(1) = 0.$$

By (16) we have

$$f_a^2(x) = \Psi^{-1}(\Psi(x) + a(x + f_a(x) - 1)),$$

so the fixed points of f_a^2 are 0, 1 and the solutions to $x + f_a(x) - 1 = 0$, that is, to $g_a(x) = x$. Thus, the fixed points of g_a^2 (which, as we noticed, is equal to f_a^2) are the fixed points of g_a and 0 and 1.

We can choose the invariant interval $I_a = I_{a,1/2}$ symmetric, so that $\varphi(I_a) = I_a$. Let us look at $G_a = g_a|_{I_a} : I_a \to I_a$. All fixed points of G_a^2 are also fixed points of G_a, so G_a has no periodic points of period 2. By the Sharkovsky Theorem, G_a has no periodic points other than fixed points. For such maps it is known (see, e.g., (Block & Coppel, 2006)) that the ω-limit set of every trajectory is a singleton of a fixed point, that is, every trajectory converges to a fixed point. If $x \in (0,1) \setminus I_a$, then the g_a-trajectory of x after a finite time enters I_a, so g_a-trajectories of all points of $(0,1)$ converge to a fixed point of g_a in I_a. Observe that a fixed point of g_a can be a fixed point of f_a (other that 0, 1) or a periodic point of f_a of period 2. Thus, the f_a-trajectory of every point of $(0,1)$ converges to a fixed point or a periodic orbit of period 2 of f_a, other than 0 and 1.

Observe now that $1/2$ is a fixed point of both f_a and g_a. The fixed points of g_a in $[0, 1/2]$ are the solutions of the equation $g_a(x) = x$, which is equivalent to $f_a(x) = 1 - x$, further to

$$\Psi(x) + a(x - 1/2) = \Psi(1 - x)$$

and finally, by Proposition 3.1.i, to

$$2\Psi(x) = -a(x - 1/2).$$

Define $\gamma_a(x) = -a/2(x - 1/2)$. We look for $\sigma_a \in (0, 1/2)$ such that

$$\Psi(\sigma_a) = \gamma_a(\sigma_a). \tag{19}$$

We know that $\Psi(1/2) = \gamma_a(1/2) = 0$ and $\gamma_a'(1/2) = -\frac{a}{2}$. As $\Psi'(1/2) < 0$ (Ψ is strictly decreasing) there is no solution of (19) in $(0, 1/2)$ for sufficiently small a. Then, $1/2$ is the only fixed point of g_a in $(0,1)$. Thus, $1/2$ will attract all points from $(0,1)$.

If $\gamma_a'(1/2) < \Psi'(1/2)$, then there exists $x \in (0, 1/2)$ such that $\gamma_a(x) > \Psi(x)$. Because $\lim_{x\to 0^+} \Psi(x) = +\infty$ and $\lim_{x\to 0^+} \gamma_a(x) = a/4$ and both functions are continuous, there exists $\sigma_a \in (0, 1/2)$ such that $\Psi(\sigma_a) = \gamma_a(\sigma_a)$. Finally, $\gamma_a'(1/2) < \Psi'(1/2)$ if and only if $a > -2\Psi'(1/2)$.

Lastly, if Ψ is convex on some neighborhood of zero, that is $(0, \delta)$, then we can choose a (sufficiently large) such that all solutions of (19) in $(0, \frac{1}{2})$ lay in $(0, \delta)$. From the fact that Ψ is convex on this interval and γ_a is an affine function we obtain uniqueness of σ_a. □

D. Figures

In this Section we include the figures described in Section 8 of the main part of the text. As already discussed, we observe a plethora of complex phenomena such as the simultaneous creation and destruction of different attractors, locally complex behavior and different period doubling phenomena (not always leading to chaos) for different type of regularizers.

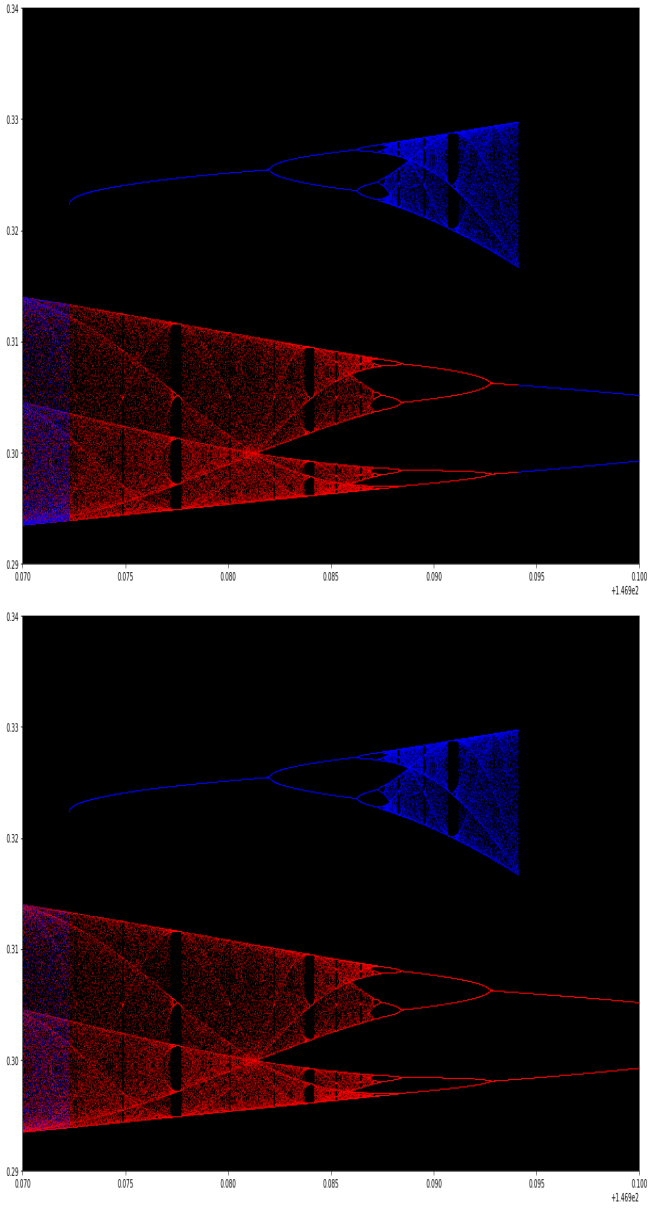


Figure 4. **Simultaneous creation and destruction of different attractors.** The bifurcation diagrams for $f_{a,b}$ where the dynamics is determined by taking (negative) log-barrier regularizer with parameter $b = 0.61$. On the horizontal axis the parameter a is between 146.97 and 147, and on the vertical axis values of $f_{a,b}$ between 0.27 and 0.34 are shown. As starting points for bifurcation diagrams two critical points of $f_{a,b}$ are taken (regularity of this map, see Appendix B, guarantees that their trajectories detect all attractors). — red refers to the critical point in $(0, 0.5)$ and blue to the critical point in $(0.5, 1)$. Each critical point is iterated 4000 times, visualizing the last 200 iterates. On the top picture first red and then blue trajectories are drawn and on the bottom one first blue and then red. We observe the collapse of the red attractor (built on the left critical point) with the simultaneous creation of the blue one (built on the right critical point).

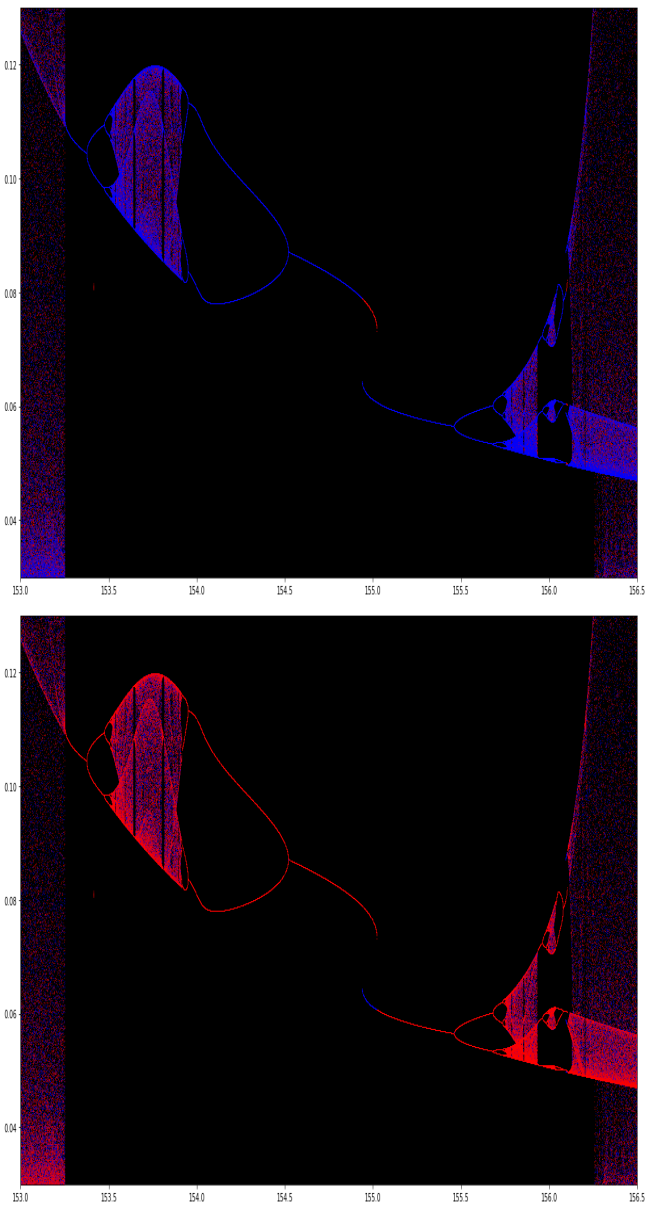
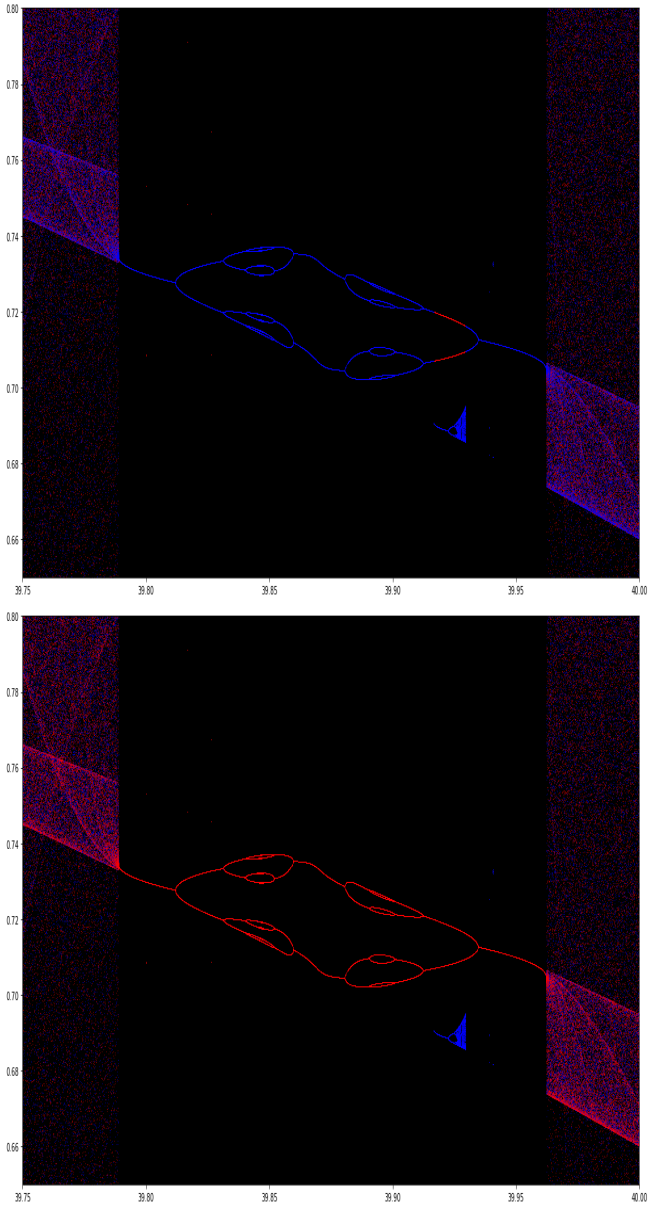


Figure 5. **Locally complex behavior.** The bifurcation diagrams for $f_{a,b}$ where the dynamics is determined by taking (negative) log-barrier as the regularizer for $b = 0.61$. On the horizontal axis the parameter a is between 153 and 156.5, and on the vertical axis values of $f_{a,b}$ are between 0.03 and 0.13. As starting points for bifurcation diagrams two critical points of $f_{a,b}$ are taken — red refers to the critical point in $(0, 0.5)$ and blue the critical point in $(0.5, 1)$. Each critical point is iterated 4000 times, then visualizing the last 200 iterates. On the top picture first red and then blue trajectories are drawn, and on the bottom one the order is reversed. As a increases chaotic behavior of orbits disappears (around 153.25). Then, within the window $[153.25, 156.3]$, chaos emerges at $[153.5, 154]$ and vanishes. Then trajectories jump, one after the other, and then generate a chaotic attractor which then spreads, vanishes, and finally spreads onto the whole interval.



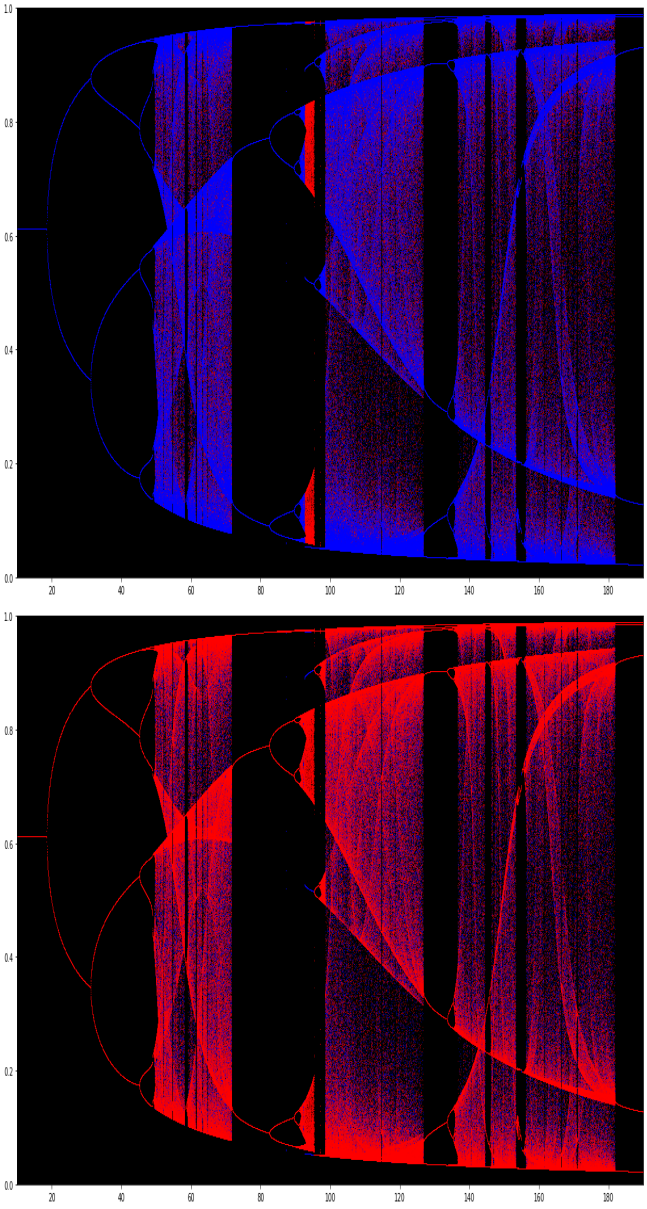


Figure 6. **Period-doubling not always lead to chaos.**The bifurcation diagrams for $f_{a,b}$ where the dynamics is determined by taking (negative) Havrda-Charvát-Tsallis entropy with $q = 0.5$ as the regularizer, that is, $r(x) = \frac{1}{\sqrt{x}} - \frac{1}{\sqrt{1-x}}$. We fix $b = 0.61$. On the horizontal axis the parameter a is between 39.75 and 40, and on the vertical axis values of $f_{a,b}$ are between 0.65 and 0.8. As starting points for bifurcation diagrams two critical points of $f_{a,b}$ are taken — red refers to the critical point in $(0, 0.5)$ and blue the critical point in $(0.5, 1)$. Each critical point is iterated 4000 times, then visualizing the last 200 iterates. On the top picture first red and then blue trajectories are drawn, and on the bottom one the order is reversed. As a increases both trajectories go through the same forward and backward period doubling steps. Then, as a increases from 39.915 to 39.93, the trajectory of the right critical point escapes the attractor which she shared with the trajectory of the left critical point, and builds separate chaotic attractor. Then it jumps back to the red attractor.

Figure 7. **Period-doubling road to chaos with log-barrier regularizer.** The bifurcation diagrams for $f_{a,b}$ where the dynamics is determined by taking (negative) log-barrier as the regularizer: $r(x) = -\log x - \log(1-x)$ for $b = 0.61$. On the horizontal axis the parameter a is between 10 and 190, and on the vertical axis the value of $f_{a,b}$ ranges between 0 and 1. As starting points for bifurcation diagrams two critical points of $f_{a,b}$ are taken (regularity of this map, see Appendix B, guarantees that by studying their trajectories we visit all attractors) — red refers to the left critical point (in $(0, 0.5)$) and blue to the right critical point (in $(0.5, 1)$). Each critical point is iterated 4000 times, visualizing the last 200 iterates. On the top picture first red and then blue trajectories are drawn, and on the bottom one the order is reversed. The first bifurcation takes place at the moment when the Nash equilibrium b becomes repelling. Then we observe period-doubling route to chaos. In addition two different attractors are visible for $a \in (92, 96)$.

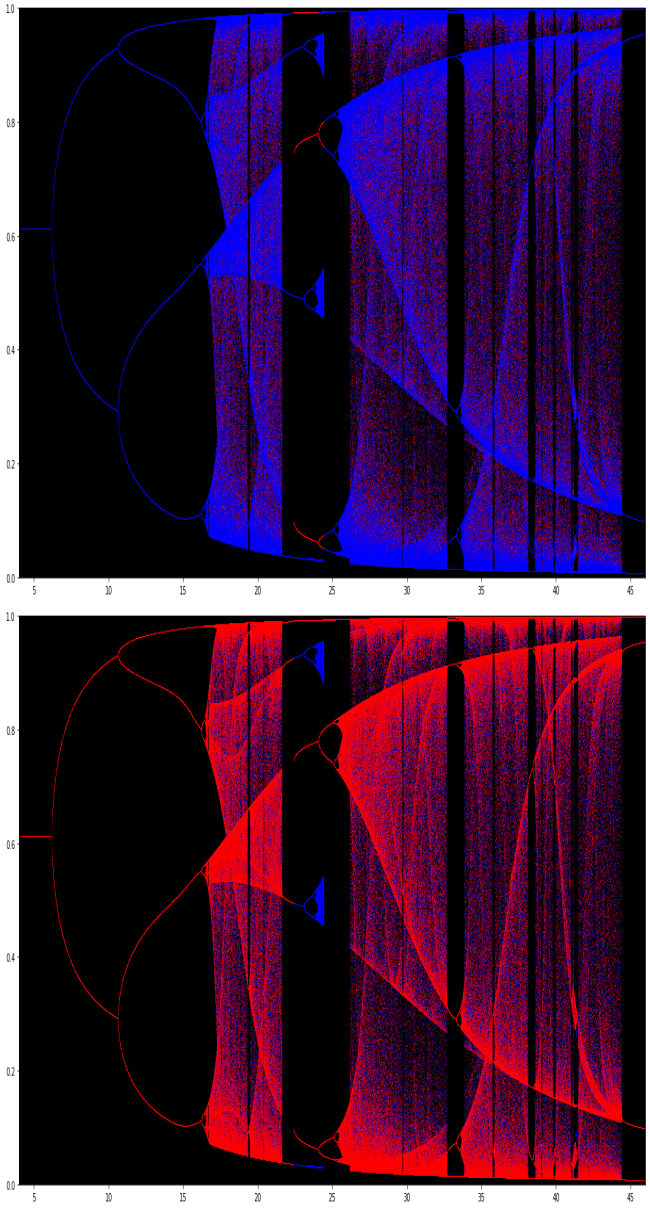


Figure 8. **Period-doubling road to chaos with Havrda-Charvát-Tsallis regularizer.** The bifurcation diagrams for $f_{a,b}$ where the dynamics is determined by taking (negative) Havrda-Charvát-Tsallis entropy with $q = 0.5$ as the regularizer and $b = 0.61$. On the horizontal axis the parameter a is between 4 and 46, and on the vertical axis values of $f_{a,b}$ ranges between 0 and 1. As starting points for bifurcation diagrams two critical points of $f_{a,b}$ are taken — red refers to the critical point in $(0, 0.5)$ and blue the critical point in $(0.5, 1)$. Each critical point is iterated 4000 times, then visualizing the last 200 iterates. On the top picture first red and then blue trajectories are drawn, and on the bottom one the order is reversed. The first bifurcation takes place at the moment when the Nash equilibrium b becomes repelling. Then we observe period-doubling route to chaos. In addition two different attractors are visible for $a \in (22.5, 24.5)$.