BULLETIN (New Series) OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 51, Number 4, October 2014, Pages 669–674
S 0273-0979(2014)01445-4
Article electronically published on May 13, 2014

Entropy in dynamical systems, by Tomasz Downarowicz, New Mathematical Monographs, Vol. 18, Cambridge University Press, Cambridge, 2011, xii+391 pp., ISBN 978-0-521-88885-1, US $101.00

Since the introduction of dynamical entropy by Kolmogorov in 1958 [Ko58], entropy has played a central role in dynamics. Today there are many branches of active research in this area. The reader is encouraged to peruse [Ka07] for a more complete survey.

1. Measure entropy

Let us begin with a brief review of measure-theoretic entropy. Imagine you are observing some dynamical system by making and recording observations at equal time intervals, say at integer times. Because your measuring instruments and recording devices have finite capacity, each individual observation can be encoded by a symbol from a fixed finite alphabet. So when you finish observing, a sequence of symbols from a finite alphabet is produced that represents all the information collected. This process is called *digitalization*.

Now let us introduce some mathematical formalism. The set of all possible states of the dynamical system is represented by a probability space (X, μ). We assume that our system is deterministic, which implies the existence of a transformation $T : X \to X$ representing time (so if x is the present state, then Tx is the state one moment into the future). We also assume our system is in equilibrium, which means T preserves the measure μ. An *observable* is a measurable map ϕ from X to a finite alphabet $\mathcal{A}$. This induces a measurable digitalization map Φ from X into the space $\mathcal{A}^{\mathbb{N}}$ of all infinite sequences with values in $\mathcal{A}$; namely, a state $x \in X$ is mapped to its itinerary as "seen through ϕ". So $\Phi(x) = (\phi(x), \phi(Tx), \phi(T^2x), \ldots)$.

Entropy is used to measure how predictable the sequence $\Phi(x)$ is and also how efficient the encoding is. To be precise, suppose $Y \subset X$ is known to us and $x \in X$ is random and unknown to us. How much information do we gain by learning that $x \in Y$? Let us say that the answer is a real number (of *bits*) denoted by $I(Y)$. We assume the following: that $I(Y)$ is always nonnegative; that it depends only on the measure $\mu(Y)$; and that if Y_1, Y_2 are independent, then $I(Y_1 \cap Y_2) = I(Y_1) + I(Y_2)$, and if $\mu(Y) = 1/2$, then $I(Y) = 1$ bit. These assumptions and measurability imply $I(Y) = -\log_2(\mu(Y))$.

The *Shannon entropy* of the observable ϕ is the average amount of information gained by learning $\phi(x)$ when $x \in X$ is random. Precisely, it is $H_\mu(\phi) := \sum_{a\in\mathcal{A}} \mu(\phi^{-1}(a)) I(\phi^{-1}(a))$. The *entropy rate* of ϕ with respect to T, denoted $h_\mu(T, \phi)$, is the average amount of information generated by the system per unit of time: $h_\mu(T, \phi) = \lim_{n\to\infty} \frac{1}{n} H(\bigvee_{i=0}^{n-1} \phi \circ T^i)$. Intuitively, if $h_\mu(T, \phi)$ is large, then, on average, it is difficult to predict $\phi(T^n x)$ even if we know the values of $\phi(x), \phi(Tx), \ldots, \phi(T^{n-1}x)$.

In order to obtain an invariant of the dynamical system that does not depend on ϕ, we consider *lossless* observables. That is, we say ϕ is lossless if the digitalization $\Phi(x)$ determines x for almost every x. It was A. Kolmogorov's great insight that all

2010 *Mathematics Subject Classification.* Primary 37A15, 37A20, 03E15.

lossless observables have the same entropy rate. He therefore defined the entropy of the dynamical system by $h_\mu(T) := h_\mu(T, \phi)$, where ϕ is any lossless observable. But what if there are no lossless observables? No problem: Y. Sinai showed that the following new definition $h_\mu(T) := \sup_\phi h_\mu(T, \phi)$ agrees with Kolmogorov's definition, and it clearly does not require the existence of a lossless observable.

Kolmogorov was motivated by a problem posed by von Neumann in the 1920s about classifying measure-preserving transformations. To be precise, if (Y, ν) is a probability space and $S : Y \to Y$ is a measure-preserving transformation, then T and S are said to be *measurably conjugate* or *isomorphic* if there is a measure-space isomorphism $\Psi : X \to Y$ intertwining the two transformations (so $\Psi T = S\Psi$). We would like to classify transformations up to measure-conjugacy. There is a special class of transformations, called Bernoulli shifts, which has played an important role in this history. They are defined as follows. Given a probability space (K, κ), we let $K^{\mathbb{Z}}$ denote the set of all functions from $\mathbb{Z}$ to K with the product measure $\kappa^{\mathbb{Z}}$. The shift map $\sigma : K^{\mathbb{Z}} \to K^{\mathbb{Z}}$ is defined by $\sigma(x)_i = x_{i+1}$ for $i \in \mathbb{Z}$. This transformation is called the *Bernoulli shift* with base space (K, κ). In the special case in which K is finite of cardinality $n \geq 1$ and κ is the uniform probability measure, we say that σ is the *full n-shift*. Von Neumann asked whether the full 2-shift is measurably conjugate to the full 3-shift. At the time, only spectral invariants of dynamical systems were known, but these do not distinguish Bernoulli shifts. Kolmogorov solved von Neumann's problem with a short computation of the entropy of a Bernoulli shift (e.g., the full n-shift has entropy $\log(n)$).

A celebrated result of Ornstein [Or70a, Or70b] obtains the converse and thereby shows that Bernoulli shifts are completely classified by entropy. Ornstein's machine, as it is now known, led to many other breakthroughs. For example, many transformations of classical interest are measurably conjugate to Bernoulli (e.g., mixing Markov chains, hyperbolic toral automorphisms, the time 1 map of the geodesic flow on a hyperbolic surface). These facts cemented the absolutely fundamental role of entropy in measurable dynamics. Ornstein theory also led to uniform proofs of Sinai's Theorem (any system factors onto a Bernoulli system with equal entropy) and Krieger's Theorem (any transformation with entropy less than $\log(n)$ admits a lossless observable $\phi : X \to \mathcal{A}$ with alphabet $\mathcal{A}$ of cardinality n). In other words, there is an *efficient* digitalization. Further research has shown that Bernoulli shifts with the same entropy over finite alphabets are *finitarily isomorphic*, which is to say there exists an almost-continuous measure-conjugacy Ψ with respect to the product topology [KS79, Se06]. In other words, a Bernoulli shift can be effectively recoded into an arbitrary second Bernoulli shift of the same entropy. Ornstein's ideas also played an important role in the construction of counterexamples, such as non-Bernoulli K-automorphisms [Or73] and the recent anti-classification theorems [FW04, FRW11] that study the complexity of the classification problem for measure-preserving transformations up to measure-conjugacy. For example, [FW04] proves it is impossible to classify all measure-preserving transformations up to measure-conjugacy by countable structures.

As an aside, we would like to mention that entropy theory was extended to actions of amenable groups in the 1970s and 1980s [OW80, OW87] and more recently to sofic groups [Bo10, KL11]. Ornstein's converse has been extended to all amenable groups [OW87], to all groups containing an infinite amenable subgroup [St75], and to all countable groups provided the base measures of the Bernoulli shifts are supported on more than two elements [Bo12].

2. Topological entropy

Topological entropy was introduced by Adler et al. [AKM65]. Given a homeomorphism T of a compact metrizable space X, the topological entropy $h(T)$ measures the complexity of the dynamics, interpreted as the amount of information transmitted by the system per unit of time. This quote from Downarowicz provides an intuitive explanation:

> The initial state carries complete information about the evolution ..., but the observer cannot "read" all this information immediately. Since we do not fix any particular measure, we want to use the metric (or more generally, the topology) to describe the "amount of information" about the initial state, acquired by the observer in one step (one measurement). A reasonable intepretation relies on the notion of *topological resolution*. Intuitively, resolution is a parameter measuring the ability of the observer to distinguish between points. A resolution is topological, when this ability agrees with the topological structure of the space. The simplest such resolution is based on the metric and a positive number ϵ: two points are "indistinguishable" if they are less than ϵ apart. Another way to define a topological resolution (applicable in all topological spaces) refers to an open cover of X. Points cannot be distinguished when they belong to a common cell of the cover.
>
> –Tomasz Downarowicz, *Entropy in dynamical systems*

For example, the topological entropy with respect to an open cover $\mathcal{U}$ can be defined by $h(T, \mathcal{U}) = \lim_{n\to\infty} \frac{1}{n} \log N(\mathcal{U}^n)$, where $\mathcal{U}^n$ is the open cover $\bigvee_{i=0}^{n-1} T^{-i}\mathcal{U}$ and $N(\mathcal{U}^n)$ is the cardinality of the smallest subcover of $\mathcal{U}^n$. The entropy $h(T)$ is defined to be the supremum of $h(T, \mathcal{U})$ over all open covers $\mathcal{U}$.

The space of all T-invariant Borel probability measures on X is a Choquet simplex denoted $M(T)$: it is a compact convex subset of topological vector space with the property that every $\mu \in M(T)$ is the barycenter of a unique probability measure supported on the extreme points of $M(T)$, which are precisely the ergodic measures. It is natural then to obtain invariants of T from the space $M(T)$. A classical example of this is the variational principle: the topological entropy of T is the supremum of the measure-theoretic entropies $h_\mu(T)$ over $\mu \in M(T)$. More recently [DS03] shows the set of all pairs (M, h), where M is the simplex of invariant probability measures for a system (X, T) and $h : M \to [0, \infty]$ is the entropy function is, up to isomorphism, precisely the set of all pairs (M', h') where M' is a metrizable Choquet simplex, h' is affine, and h' is a non-decreasing limit of upper semi-continuous functions.

Returning to the theme of digitalization, recall that a symbolic dynamical system has the following form. Let $\mathcal{A}$ denote a finite alphabet, and let $\mathcal{A}^{\mathbb{Z}}$ be the set of all sequences with terms in $\mathcal{A}$ with the product topology (where $\mathcal{A}$ is given the discrete topology). The shift-map $\sigma : \mathcal{A}^{\mathbb{Z}} \to \mathcal{A}^{\mathbb{Z}}$ defined by $\sigma(x)_i = x_{i+1}$ is a homeomorphism. If $X \subset \mathcal{A}^{\mathbb{Z}}$ is a closed shift-invariant subspace, then σ restricted to X is called a *symbolic dynamical system.* These systems possess nice properties: they are expansive (this means that for any continuous metric ρ on X there is a constant $\epsilon_0 > 0$ such that if $x \neq y$ are any distinct points in X, then $\rho(T^n x, T^n y) > \epsilon_0 > 0$ for some $n \in \mathbb{Z}$), the entropy function $\mu \mapsto h_\mu(T)$ is upper-semicontinuous on the simplex $M(T)$ and the topological entropy $h_{\text{top}}(T)$ is finite (indeed, it is bounded by $\log |\mathcal{A}|$). Quoting now from Downarowicz and Newhouse [DN05]:

> Due to the convenient "digital" form, these systems allow an abundance of applications in more practical areas such as information theory, signal processing, and computer science. The same form makes them also relatively easy for abstract studies. For these reasons building a symbolic model has been a key tool in the investigation of dynamical systems since the beginning of the 20th century. Classical examples of such approach are:
> (1) describing a homotopy class of a trajectory of a geodesic flow on a surface of negative curvature by a sequence of labels of certain closed curves (Hadamard, Morse),
> (2) parameterizing a unimodal map on $[0, 1]$ by the kneading sequence, obtained by labeling the trajectory of the critical point c with respect to the partition into $[0, c]$ and $(c, 1]$—the key notion in the study of chaos, bifurcations, etc.

This leads to the questions, what is a good "symbolic model" and which systems have them? In the measure-theoretic setting, by a "symbolic model" one usually means an isomorphism $\pi : X \to \mathcal{A}^{\mathbb{Z}}$ (defined on a conull set). In the topological setting, where we require continuity of the map π, it is clear that non-trivial maps do not exist if, for example, X is connected. Therefore, it reasonable to look for *symbolic extensions*. In other words, given a topological system (X, T) we seek a symbolic system $(\tilde{X}, \tilde{T})$ and a continuous surjection $\pi : \tilde{X} \to X$ with $\tilde{T}\pi = \pi T$. The amount of "imprecision" in the model is measured by the amount of entropy added to each invariant measure $\mu \in M(X, T)$. So we define $h^{\pi}_{\mu}(T) = \sup_{\nu} h_{\nu}(\tilde{T})$, where the sup is over all $\tilde{T}$-invariant probability measures ν that project to μ and the *symbolic extension entropy* is defined by $h_{\rm sex}(\mu) = \inf_{\pi} h^{\pi}_{\mu}(T)$, where the infimum is over all symbolic extensions π. The symbolic extension entropy of the system is $h_{\rm sex}(T) := \inf h(\tilde{T})$, where the infimum is over all symbolic extensions $(\tilde{X}, \tilde{T})$.

Of course, an infinite-entropy system does not have any symbolic systems. Around 1990 Mike Boyle gave the first examples of finite-entropy systems with no symbolic extensions [BFF02]. In [Do01], a family of examples is constructed showing that any positive number can be the residual entropy, which is defined by $h_{\rm res}(T) := h_{\rm sex}(T) - h(T)$. By contrast, C^{∞}-systems behave very nicely: Buzzi [Bu97] proved C^{∞}-systems are asymptotically h-expansive and [BFF02] showed that asymptotic h-expansiveness is equivalent to the existence of a *principal symbolic extension*, meaning an extension whose residual entropy is zero. C^r-diffeomorphisms are more complicated: the authors of [DN05] conjectured an explicit upper bound on the symbolic extension entropy of a C^r-system ($r > 1$) in terms of Lyapunov exponents. This conjecture is confirmed in [DM09] for maps of the interval and circle and in [Bu11] for C^2-surface diffeomorphisms, but it remains open in higher dimensions.

In [Do05] Downarowicz introduced a new invariant, called the *entropy structure* of a topological dynamical system. This "master invariant" consists of an equivalence class of sequences of functions $\{h_k\}_k$ on the simplex $M(X, T)$ of invariant Borel probability measures. Intuitively, $h_k(\mu)$ is the entropy of μ as measured at a certain scale parametrized by k. In fact, Downarowicz considers several established methods for computing entropy and shows that they give rise to equivalent sequences (with a few exceptions). The equivalence relation on such sequences captures the "type of non-uniformity" in the convergence of h_k to the entropy function. Downarowicz shows that symbolic extension entropy and Misuirewicz's topological conditional entropy can be derived from entropy structure. More generally, entropy structure determines the entire set of possible entropy extension functions of a system. These results are applied in [DM09] to obtain bounds on the symbolic extension entropy of interval and circle diffeomorphisms.

3. This book

The book under review is a pleasure to read as each new concept is well motivated and presented in an intuitively clear manner before being made rigorous. The first part of the book is on measure entropy. Aside from standard material, the book includes the intriguing Ornstein–Weiss return times theorem and the shortest known proof of Ornstein's Theorem classifying Bernoulli shifts (based on [DS12]). There is an "optional" section on the ergodic law of series, which seeks to provide a mathematical basis to explain the curious phenomenon that random events (usually extremely rare) may be observed suprisingly often throughout a relatively short period of time (based on [DL11, Do11b]).

The second chapter covers topological entropy. While some of the material here is classical, the heart of this chapter comes from recent research related to the author's interests. There is a large variety of different kinds of relative entropy for topological systems which are carefully explained along with how they relate to one another. Three different variational principles are presented. It is really a welcome contribution to the literature to have all of these concepts explained in one place in a uniform and systematic matter. The chapter also includes a section on Downarowicz' entropy structure. This is a complicated notion to absorb, and the author deserves credit for a careful presentation with helpful examples. Symbolic extension entropy and tail entropy are also carefully explained as well as their relationships with entropy structure.

The last chapter covers the entropy theory of stochastic and Markov operators. These correspond to random dynamical systems. The treatment follows the author's work [DF05].

References

[AKM65] R. L. Adler, A. G. Konheim, and M. H. McAndrew, *Topological entropy*, Trans. Amer. Math. Soc. **114** (1965), 309–319. MR0175106 (30 #5291)

[BFF02] Mike Boyle, Doris Fiebig, and Ulf Fiebig, *Residual entropy, conditional entropy and subshift covers*, Forum Math. **14** (2002), no. 5, 713–757, DOI 10.1515/form.2002.031. MR1924775 (2003g:37024)

[Bo10] Lewis Bowen, *Measure conjugacy invariants for actions of countable sofic groups*, J. Amer. Math. Soc. **23** (2010), no. 1, 217–245, DOI 10.1090/S0894-0347-09-00637-7. MR2552252 (2011b:37010)

[Bo12] Lewis Bowen, *Every countably infinite group is almost Ornstein*, Dynamical systems and group actions, Contemp. Math., vol. 567, Amer. Math. Soc., Providence, RI, 2012, pp. 67–78, DOI 10.1090/conm/567/11234. MR2931910

[Bu97] Jérôme Buzzi, *Intrinsic ergodicity of smooth interval maps*, Israel J. Math. **100** (1997), 125–161, DOI 10.1007/BF02773637. MR1469107 (99g:58071)

[Bu11] David Burguet, *$\mathcal{C}^2$ surface diffeomorphisms have symbolic extensions*, Invent. Math. **186** (2011), no. 1, 191–236, DOI 10.1007/s00222-011-0317-8. MR2836054 (2012k:37049)

[DF05] Tomasz Downarowicz and Bartosz Frej, *Measure-theoretic and topological entropy of operators on function spaces*, Ergodic Theory Dynam. Systems **25** (2005), no. 2, 455–481, DOI 10.1017/S014338570400032X. MR2129106 (2006k:47024)

[DL11] T. Downarowicz and Y. Lacroix, *The law of series*, Ergodic Theory Dynam. Systems **31** (2011), no. 2, 351–367, DOI 10.1017/S0143385709001217. MR2776379 (2012h:37009)

[DM09] Tomasz Downarowicz and Alejandro Maass, *Smooth interval maps have symbolic extensions: the antarctic theorem*, Invent. Math. **176** (2009), no. 3, 617–636, DOI 10.1007/s00222-008-0172-4. MR2501298 (2010d:37038)

[DN05] Tomasz Downarowicz and Sheldon Newhouse, *Symbolic extensions and smooth dynamical systems*, Invent. Math. **160** (2005), no. 3, 453–499, DOI 10.1007/s00222-004-0413-0. MR2178700 (2006j:37021)

[Do01] Tomasz Downarowicz, *Entropy of a symbolic extension of a dynamical system*, Ergodic Theory Dynam. Systems **21** (2001), no. 4, 1051–1070, DOI 10.1017/S014338570100150X. MR1849601 (2002e:37016)

[Do05] Tomasz Downarowicz, *Entropy structure*, J. Anal. Math. **96** (2005), 57–116, DOI 10.1007/BF02787825. MR2177182 (2006g:37016)

[Do11b] Tomasz Downarowicz, *A mathematical approach to the law of series* (Polish), Wiad. Mat. **47** (2011), no. 1, 1–16. MR2961805

[DS03] Tomasz Downarowicz and Jacek Serafin, *Possible entropy functions*, Israel J. Math. **135** (2003), 221–250, DOI 10.1007/BF02776059. MR1997045 (2004c:37013)

[DS12] T. Downarowicz and J. Serafin, *A short proof of the Ornstein theorem*, Ergodic Theory Dynam. Systems **32** (2012), no. 2, 587–597, DOI 10.1017/S0143385711000265. MR2901361

[FRW11] Matthew Foreman, Daniel J. Rudolph, and Benjamin Weiss, *The conjugacy problem in ergodic theory*, Ann. of Math. (2) **173** (2011), no. 3, 1529–1586, DOI 10.4007/annals.2011.173.3.7. MR2800720 (2012k:37006)

[FW04] Matthew Foreman and Benjamin Weiss, *An anti-classification theorem for ergodic measure preserving transformations*, J. Eur. Math. Soc. (JEMS) **6** (2004), no. 3, 277–292. MR2060477 (2005c:37006)

[Ka07] Anatole Katok, *Fifty years of entropy in dynamics: 1958–2007*, J. Mod. Dyn. **1** (2007), no. 4, 545–596, DOI 10.3934/jmd.2007.1.545. MR2342699 (2008i:37001)

[Ko58] A. N. Kolmogorov, *A new metric invariant of transient dynamical systems and automorphisms in Lebesgue spaces* (Russian), Dokl. Akad. Nauk SSSR (N.S.) **119** (1958), 861–864. MR0103254 (21 #2035a)

[KL11] David Kerr and Hanfeng Li, *Entropy and the variational principle for actions of sofic groups*, Invent. Math. **186** (2011), no. 3, 501–558, DOI 10.1007/s00222-011-0324-9. MR2854085

[KS79] Michael Keane and Meir Smorodinsky, *Bernoulli schemes of the same entropy are finitarily isomorphic*, Ann. of Math. (2) **109** (1979), no. 2, 397–406, DOI 10.2307/1971117. MR528969 (80f:28024)

[Or70a] Donald Ornstein, *Bernoulli shifts with the same entropy are isomorphic*, Advances in Math. **4** (1970), 337–352 (1970). MR0257322 (41 #1973)

[Or70b] Donald Ornstein, *Two Bernoulli shifts with infinite entropy are isomorphic*, Advances in Math. **5** (1970), 339–348 (1970). MR0274716 (43 #478a)

[Or73] Donald S. Ornstein, *An example of a Kolmogorov automorphism that is not a Bernoulli shift*, Advances in Math. **10** (1973), 49–62. MR0316682 (47 #5229)

[OW80] Donald S. Ornstein and Benjamin Weiss, *Ergodic theory of amenable group actions. I. The Rohlin lemma*, Bull. Amer. Math. Soc. (N.S.) **2** (1980), no. 1, 161–164, DOI 10.1090/S0273-0979-1980-14702-3. MR551753 (80j:28031)

[OW87] Donald S. Ornstein and Benjamin Weiss, *Entropy and isomorphism theorems for actions of amenable groups*, J. Analyse Math. **48** (1987), 1–141, DOI 10.1007/BF02790325. MR910005 (88j:28014)

[Se06] Jacek Serafin, *Finitary codes, a short survey*, Dynamics & stochastics, IMS Lecture Notes Monogr. Ser., vol. 48, Inst. Math. Statist., Beachwood, OH, 2006, pp. 262–273, DOI 10.1214/lnms/1196285827. MR2306207 (2008a:37009)

[St75] A. M. Stepin, *Bernoulli shifts on groups* (Russian), Dokl. Akad. Nauk SSSR **223** (1975), no. 2, 300–302. MR0409769 (53 #13521)

L. Bowen

Department of Mathematics

University of Texas at Austin

E-mail address: `lpbowen@math.utexas.edu`