UNIVERZITET U BEOGRADU
MATEMATIČKI FAKULTET

MASTER RAD

Ergodičnost i entropija dinamičkih sistema

Mentor: prof. dr Nebojša Lažetić

Autor: Marija Boričić

Beograd, septembar 2011.

P r e d g o v o r

Ovaj rad je pisan tokom školske 2010/11. godine na master studijama studijskog programa Matematika, modul Teorijska matematika i primene. Rad se sastoji iz tri dela. U uvodnom delu, pored istorijskih činjenica koje se odnose na evoluciju tretmana pojma neodređenosti u fizici i matematici, dajemo i precizne definicije osnovnih pojmova teorije dinamičkih sistema i teorije mere, koji se pojavljuju u nastavku rada. Drugi deo rada sadrži najvažnije činjenice iz ergodičke teorije koje, po pravilu, osiguravaju egzistenciju pojmova koje definišemo kasnije. Završni, i najobimniji deo, tretira problem merenja neodređenosti sistema. Manjim delom to je osvrt na prirodno pojavljivanje entropije u termodinamici, teoriji informacija i teoriji mere, dok centralno mesto zauzima definicija, osobine i sam smisao entropije u kontekstu teorije dinamičkih sistema, zasnovano na prvim radovima Kolmogorova i Sinaia koji se pojavljuju u drugoj polovini XX veka, kao i radovima koji su usledili do današnjeg vremena. Dominantan je utisak da su radovi pomenutih matematičara apsolutno promenili mesto i ulogu pojma entropije u nauci, premeštajući je iz konteksta stohastičkih sistema u kontekst determinističkih dinamičkih sistema, pridajući joj, na taj način, značaj jedne univerzalne kategorije koja karakteriše neki sistem.

Koristim ovu priliku da zahvalim svom mentoru, prof. dr Nebojši Lažetiću, na nesebičnoj i presudnoj stručnoj podršci, te nizu korisnih sugestija i komentara učinjenih tokom pisanja ovog rada. Pored toga, svu odgovornost za eventualne propuste u radu snosim lično, kao autor ovog rada.

Beograd, septembar 2011.

Marija Boričić

SADRŽAJ

§1. Uvod

Teorija dinamičkih sistema predstavlja granu primenjene matematike koja matematičkim sredstvima opisuje ponašanje složenih, u početku fizičkih i hemijskih, a danas i društvenih, računarskih, bioloških i sličnih sistema koji se menjaju u prostoru i vremenu. Prve korene ove teorije svakako možemo naći u prvim matematičkim opisima realnog sveta koji nas okružuje i koji su omogućili ustanovljenje kalendara i predviđanja određenih kosmičkih pojava poput pomračenja Sunca i Meseca. Početke savremene teorije dinamičkih sistema nalazimo u Newton–ovoj[1] mehanici. Dinamički sistem se može posmatrati kao matematička formalizacija neke zakonitosti po kojoj se, zavisno od vremena, može opisati kretanje svih tačaka nekog skupa u odgovarajućem prostoru.

Ipak, potreba za egzaktnim opisom stvarnosti koji bi nam bio od koristi u njenom boljem tumačenju, ali i predviđanju i prognoziranju, dovodi do potrebe razvijanja odgovarajuće matematičke aparature. Stoga je, moglo bi se reći, teorija dinamičkih sistema, koja ostvaruje najveću interakciju između matematike i sveta koji nas okružuje, najzaslužnija za ubrzan razvoj pojedinih matematičkih disciplina i savremene metodologije. Tu pre svega imamo u vidu primere poput ekspanzivnog razvoja teorije upravljanja, odnosno kibernetike, tokom XX veka, ali i razvoja telekomunikacionih sistema tokom poslednjih decenija, kakav ne bi bio moguć bez odgovarajuće matematičke podrške. S druge strane, svaki napredak tehnologije je zahtevao razvoj adekvatnog matematičkog konteksta, čime se ostvaruje pomenuta interakcija. Razvoj računarstva, te mogućnost numeričkog rešavanja širokog spektra matematičkih problema daju znatan podsticaj razvoju ove teorije.

Kao svoje osnovno sredstvo, teorija dinamičkih sistema koristi diferencijalne i diferencne jednačine, pomoću kojih opisuje ponašanje posmatranih realnih sistema, zavisno od toga da li je menjanje sistema posmatrano u neprekidnom ili diskretnom vremenu. Jedno od glavnih pitanja apstraktne teorije dinamičkih sistema jeste pitanje klasifikacije sistema. U tom cilju je neophodno obavljati odgovarajuća merenja, radi kvantifikacije i upoređivanja, a za ta merenja je nužno snabdeti sistem adekvatnim sistemima mera. Među najznačajnije osobine pogodne za klasifikacije, kakvih, inače nema mnogo, jesu one osobine sistema koje su *invarijantne,* poput, na primer, periodičnosti i neodređenosti.

Dinamički sistemi sa invarijantnom merom direktno vode ergodičkoj teoriji, a ergodička teorija omogućava uvođenje i posmatranje jednog sasvim drugačijeg koncepta entropije od onog koji je karakterističan za stohastičke sisteme. Entropija u determinističkim dinamičkim sistemima biće centralna tema ovog rada.

Problem merenja, kao stalna opsesija matematičara, kroz merenje dužine duži, u jednodimenzionom prostoru, površine figure, u dvodimenzionom, zapremine tela u trodimenzionom, i uopšte određivanje mere nekog skupa u višedimenzionom prostoru ili u jednom sasvim apstraktnom smislu, dovodi do konstituisanja *teorije mere,* kao zasebne

[1]Isaac Newton (1643–1727) slavni elngleski fizičar, matematičar, astronom, filozof, alhemičar i teolog.

matematičke discipline, koja danas predstavlja osnovu savremenog integralnog računa. S druge strane u statistici, pa delom i teoriji verovatnoće, dva su dominantna pojma koji zahtevaju stalna ocenjivanja i merenja. To su: *pouzdanost* i *neodređenost.* Ova dva pojma jesu u određenom smislu suprotna, ali ne i komplementarna.

U ovom radu dajemo pregled nekih rezultata proisteklih iz sprege *teorije mere, teorije dinamičkih sistema* i *teorije informacija.* Naime, *entropija* kao bazični pojam teorije informacija, definisan kao *mera neodređenosti slučajne promenljive,* prirodno se, posle termodinamike, pojavila u analizi funkcionisanja telekomunikacionih sistema, povezuje ove tri matematičke teorije i definiše sam naslov ovog rada. Centralno mesto našeg rada biće zapravo povezivanje entropije sa dinamičkim sistemima u kontekstu ergodičke teorije, ali i prikaz razvoja te ideje od termodinamike, preko teorije informacija i teorije mere, do teorije dinamičkih sistema.

Ergodička teorija, ili metrička teorija dinamičkih sistema, kao integralni deo teorije dinamičkih sistema, ima svoje korene još u radovima Boltzmann–a[2] u oblasti statističke mehanike, a svoje matematičke temelje dobija početkom XX veka u radovima von Neumann–a[3], Birkhoff–a[4], Hinčina[5] i Koopman[6]–a. Centralna tema ergodičke[7] teorije su dinamički sistemi sa invarijantnom merom koji, grubo govoreći, imaju isto 'prosečno ponašanje' kako u vremenu, tako i u prostoru, tokom protoka dugog vremenskog perioda.

Imajući u vidu ogromnu zavisnost ostalih grana matematike od razvoja integralnog računa, teorija mere dobija još i više na opštem značaju u matematici. Začetnike teorije mere nalazimo, krajem XIX veka, u slavnoj francuskoj matematičkoj školi. Prvi rad potiče od C. Jordan–a[8] 1892. godine, dok za savremeni oblik teorije mere najviše dugujemo E. Borel–u[9] i H. Lebesgue–u[10].

Teorija informacija je matematička disciplina koja se razvija od početka prošlog veka, i danas se, kao deo teorije verovatnoća, uz teoriju sistema i teoriju upravljanja, može

[2]Ludwig Eduard Boltzmann (1844–1906) austrijski fizičar poznat po svom doprinosu razvoju statističke mehanike i statističke termodinamike.

[3]John von Neumann (1903–1957) američki matematičar mađarskog porekla. Bavio se sa velikim uspehom širokim spektrom matematičkih disciplina poput teorije skupova, funkcionalne analize, ergodičke teorije, matematičke ekonomije i teorije igara, numeričke analize, statistike, računarstva itd. Mnogi ga smatraju jednim od najvećih matematičara XX veka.

[4]George David Birkhoff (1884-1944), američki matematičar, poznat upravo po ergodičkoj teoremi koju razmatramo u ovom radu.

[5]Aleksandr Yakovlevich Khinchin (1894–1959) sovjetski matematičar, poznat po svojim radovima iz teorije verovatnoće.

[6]Bernard Osgood Koopman (1900–1981) američki matematičar francuskog porekla.

[7]Reč 'ergodičnost' je uveo Boltzmann kao izvedenu iz starogrčkih reči '$\acute{\varepsilon}\varrho\gamma o\nu$' i '$o\delta\acute{o}\varsigma$' koje označavaju 'rad' i put'.

[8]C. Jordan (1838–1922), francuski matematičar, poznat i po radovima iz teorije grupa, linearne algebre i matematičke analize.

[9]E. Borel (1871–1957), francuski matematičar i političar.

[10]H. L. Lebesgue (1875–1941), francuski matematičar, tvorac jednog esencijalno drugačijeg pristupa teoriji mere i integracije.

posmatrati kao bazični činilac kibernetike. Neposredne primene teorija informacija nalazi u mnogim oblastima nauke, kao, na primer, u informatici, lingvistici, psihologiji, pedagogiji, estetici, teoriji igara, statistici, kombinatorici, logici, a da ne govorimo o tome kako se funkcionisanje savremenih telekomunikacionih sistema ne bi bez nje ni moglo zamisliti. Ako bi se u igri asocijacija na pominjanje 'teorije informacija' postavilo pitanje koja je to ličnost na koju se tom prilikom pomisli, to bi svakako morao biti Shannon[11]. Ipak, istorijski, pojam entropije se prvo pojavio u termodinamici, a u vezi sa definisanjem gubitka energije pri njenom transformisanju iz jednog njenog pojavnog oblika u drugi. S druge strane Kolmogorov[12] i Sinai[13] uvode entropiju u jednom opštijem kontekstu, čemu ćemo ovde posvetiti više pažnje u drugom delu rada.

§1.1. Osnovni pojmovi teorije dinamičkih sistema

Dinamički sistem predstavlja način na koji se opisuje evolucija svih tačaka nekog skupa S tokom vremena. U slučaju proizvoljnog fizičkog sistema, skup S jeste skup svih mogućih stanja tog sistema, pri čemu pod stanjem sistema podrazumevamo informaciju koja karakteriše sistem u datom trenutku. Sa matematičke tačke gledišta, skup S je neki konačnodimenzioni euklidski prostor, otvoren skup na takvom prostoru, mnogostrukost ili otvoren skup u Banach–ovom prostoru.

Najčešći pristup opisu neprekidnog dinamičkog sistema jeste posredstvom diferencijalne jednačine

$$y' = f(t, y)$$

uz početni uslov $y(t_0) = y_0$, i pretpostavku da su f i y vektor–funkcije n–dimenzionog euklidskog prostora, a f neprekidna i $t \geq 0$. Analogno se, posredstvom diferencne jednačine predstavlja diskretan dinamički sistem.

Formalnije, u novijoj literaturi, neprekidni, odnosno diskretni dinamički sistem definiše se na sledeći način:

Neka je S otvoren skup u prostoru R^n. Pod *neprekidnim dinamičkim sistemom* na skupu S podrazumevamo preslikavanje $T : R \times S \to S$ za koje važi:

(1) $T \in C^1(R \times S; S)$

(2) $T_0 : S \to S$ je identičko preslikavanje

(3) $T_t \circ T_s = T_{t+s}$, za sve $t, s \in R$, pri čemu je preslikavanje T_t definisano kao $T_t(x) = T(t, x)$.

Pod *diskretnim dinamičkim sistemom* podrazumevamo preslikavanje $g \in C^1(S; E)$, gde je S otvoren skup u konačnodimenzionom normiranom prostoru E.

[11] C. E. Shannon (1916–2001), američki matematičar poznat po svom doprinosu razvoju teorije informacija.

[12] A. N. Kolmogorov (1903–1987) slavni sovjetski matematičar, jedan od najvećih matematičara XX veka. Poznat je po svojim doprinosima u oblasti teorije verovatnoće, topologije, matematičke logike, klasične mehanike i teorije izračunljivosti.

[13] Y. G. Sinai (1935–) jedan od najplodnijih Kolmogorovljevih učenika.

Dakle, familija $\{T_t | t \in R\}$ opisuje kretanje fiksiranog stanja $x \in S$ u vremenu $t \in R$ i ona se naziva *trajektorijom* tačke x. Sa druge strane, ta familija, za fiksirani trenutak t, pokazuje gde se nalazi skup S posle proteklog vremena t. U slučaju diskretnog dinamičkog sistema, posle n jedinica vremena sistem se nalazi u stanju $g^n(x) = (\underbrace{g \circ \cdots \circ g}_{n-\text{puta}})(x)$.

Neka je dat dinamički sistem $T : R \times S \to S$. Tada nam familija $T_t : S \to S$ omoguáva da definišemo funkciju $f : S \to S$ na sledeći način:

$$(1) \qquad (\forall x \in S) \quad f(x) = \frac{d}{dt} T_t(x)|_{t=0}.$$

Dakle, dinamički sistem T generiše vektorsko polje na S, pri čemu se vektor $f(x)$ interpretira kao tangentni vektor krive $t \to T_t(x)$ u tački $t = 0$. Ukoliko definišemo $x(t) = T_t(x)$ i ako (1) interpretiramo kao diferencijalnu jednačinu

$$x'(t) = f(x(t)),$$

onda funkcija $x = x(t)$ jeste rešenje te diferncijalne jednačine, koje zadovoljava početni uslov $x(0) = x$. U tom slučaju, pod *trajektorijom* tačke x podrazumevamo rešenje $x = x(t)$, pomenute diferencijalne jednačine.

§1.2. Osnovni pojmovi teorije mere

Neka je X neprazan skup i $\mathcal{P}(X)$ njegov partitivan skup. Pod *σ–algebrom* podrazumevamo svaki skup $\mathcal{B} \subseteq \mathcal{P}(X)$ koji sadrži skup X i zatvoren je za operacije komplementa i najviše prebrojive unije. Uređen par $(X, \mathcal{B})$ tada nazivamo *merljivim prostorom.* Kada je X topološki prostor, onda ćemo uvek pod skupom $\mathcal{B}$ podrazumevati najmanju σ–algebru koja sadrži sve otvorene skupove prostora X i nazivati je *σ–algebrom Borel–ovih skupova.*

Verovatnosnu meru na merljivom prostoru $(X, \mathcal{B})$ definišemo kao funkciju $\mu : \mathcal{B} \to [0,1]$ koja je σ–aditivna, tj. važi: ako $\forall i \forall j (i \neq j \to A_i \cap A_j = \emptyset)$, onda

$$\mu(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$$

i za koju je, takođe, $\mu(X) = 1$. U ovom slučaju ćemo uređenu trojku $(X, \mathcal{B}, \mu)$ nazivati *prostorom verovatnoće.*

Neka je $(X, \mathcal{B}, \mu)$ prostor verovatnoće. Transformaciju $T : X \to X$ nazivamo *merljivom* ako, za svaki $A \in \mathcal{B}$, $T^{-1}(A) \in \mathcal{B}$. Kažemo da T *čuva meru* ako, za svaki $A \in \mathcal{B}$, $\mu(T^{-1}(A)) = \mu(A)$.

Neka su $(X_1, \mathcal{B}_1, \mu_1)$ i $(X_2, \mathcal{B}_2, \mu_2)$ prostori verovatnoće i $T_1 : X_1 \to X_1$ i $T_2 : X_2 \to X_2$ transformacije koje čuvaju meru. Kažemo da su T_1 i T_2 *izomorfne* ako postoje $M_1 \in \mathcal{B}_1$ i $M_2 \in \mathcal{B}_2$, za koje važi

$$\mu_1(M_1) = 1 \wedge T_1(M_1) \subseteq M_1 \wedge \mu_2(M_2) = 1 \wedge T_2(M_2) \subseteq M_2$$

kao i invertibilna transformacija $\varphi : M_1 \to M_2$, koja čuva meru i, za svaki $x \in M_1$, zadovoljava uslov $\varphi T_1 x = T_2 \varphi x$.

Za skupove $A, B \in \mathcal{B}$ kažemo da su *ekvivalentni po meri* ako je $\mu(A \triangle B) = 0$. Prostor ovako definisanih klasa ekvivalencije predstavlja jednu σ–algebru na kojoj je indukovana mera i taj par nazivamo *algebrom mere.*

§2. Elementi ergodičke teorije

U nameri da se sistemi matematički analiziraju, potrebno je definisati strukturu prostora stanja X dinamičkog sistema $T : X \to X$, kao i postaviti određene uslove na preslikavanje T. Izdvajamo tri najvažija slučaja:

(1) Slučaj *diferencijabilna dinamika* : X je mnogostrukost i T je difeomorfizam;

(2) Slučaj *topološka dinamika* : X je topološki prostor i T je homeomorfizam;

(3) Slučaj *ergodička teorija* : X je merljiv prostor i T transformacije koja čuva meru;

U ovom radu ćemo se prvenstveno bavati slučajem (3).

U statističkoj mehanici, teoriji informacija i drugim oblastima primene dinamičkih sistema, od značaja je proučavati ponašanje vremenske sredine $\frac{1}{N}\sum_{k=0}^{N-1} f(T^k x)$ tokom dugog vremenskog perioda, za velike vrednosti N. Osnovno pitanje ergodičke teorije jeste: kada granična vrednost $\lim_{N\to\infty} \frac{1}{N}\sum_{k=0}^{N-1} f(T^k x)$ postoji u nekom smislu? Konvergencija pomenutih vremenskih sredina je dokazana 1909. godine (Borel–ov jaki zakon velikih brojeva) u specijalnom slučaju kada su fT^k međusobno nezavisne jednako raspodeljene slučajne veličine. L^2–konvergenciju je dokazao von Neumann 1931. godine, a skoro sigurnu konvergenciju dokazao je Birkhoff iste godine. Ovim rezultatima ćemo se baviti u ovom radu.

Drugo važno pitanje ergodičke teorije jeste pod kojim uslovima se granična vrednost vremenskih sredina $\lim_{N\to\infty} \frac{1}{N}\sum_{k=0}^{N-1} f(T^k x)$ poklapa sa prostornom sredinom $\int_X f d\mu$? Ukoliko se vremenska sredina svake merljive funkcije poklapa skoro svuda sa prostornom sredinom, sistem (odnosno transformacija T), se naziva ergodičkim. Ispostavlja se da je sistem ergodički ako i samo ako orbita skoro svake tačke prolazi kroz svaki skup strogo pozitivne mere, odnosno, ako je $\mu(E) > 0$ i $\mu(F) > 0$, onda je i $\mu(T^n E \cap F) > 0$.

Treća tema kojom se bavi ergodička teorija jeste problem klasifikacije. Tačnije, kako zaključiti da li su dva data sistema međusobno izomorfna. Za dva sistema $(X, \mathcal{B}, \mu, T)$ i $(Y, \mathcal{M}, \nu, S)$ kažemo da su *izomorfni* ako postoje skupovi mere nula $X_0 \subset X$ i $Y_0 \subset Y$ i injektivno preslikavanje $\phi : X \setminus X_0 \to Y \setminus Y_0$ takvo da je $\phi T = S\phi$ na $X \setminus X_0$ i $\mu(\phi^{-1}E) = \nu(E)$, za sve merljive skupove $E \subset Y \setminus Y_0$. Jedan od načina za ustanovljavanje da li su dva sistema međusobno (metrički) izomorfna, jeste uvođenje invarijanti u odnosu na izomorfizme u sisteme, čiji jedan važan primer predstvalja ergodičnost.

Sada ćemo se baviti osnovnim tvrđenjima ergodičke teorije.

§2.1. Ergodičke teoreme

Opšta pretpostavka čije zadovoljenje podrazumevamo u tvrđenjima koja slede jeste da transformacija $T : X \to X$ čuva meru verovatnosnog prostora $(X, \mathcal{B}, \mu)$, što, dakle, znači da, za svaki $B \in \mathcal{B}$, važi $T^{-1}B \in \mathcal{B}$ i $\mu(T^{-1}B) = \mu(B)$.

Dokazaćemo sledeće tvrđenje:

Teorema. (H. Poincaré[14]) *Neka je $B \in \mathcal{B}$ takav da je $\mu(B) > 0$. Tada za skoro svaki $x \in B$ postoji niz $n_1 < n_2 < \dots$ prirodnih brojeva takav da $T^{n_i}(x) \in B$, za svaki $i \in \{1, 2, \dots\}$.*

Dokaz. Neka su skupovi E i F definisani na sledeći način:

$$E = \{x \in B | T^{n_i}(x) \in B, \forall i \in \{1, 2, \dots\}\}$$

$$F = \{x \in B | T^n x \notin B, \forall n \geq 1\}$$

Da bismo dokazali teoremu, dovoljno je dokazati da je $\mu(B \backslash E) = 0$. Za ovako definisane skupove E i F važi $B \backslash E = \cup_{k=0}^{\infty}(T^{-k}F \cap B)$, odakle sledi

$$\mu(B \backslash E) = \mu(\cup_{k=0}^{\infty}(T^{-k}F \cap B)) \leq \mu(\cup_{k=0}^{\infty}(T^{-k}F)) \leq \sum_{k=0}^{\infty} \mu(T^{-k}F).$$

Kako preslikavanje T čuva meru, za svaki $k \geq 1$ važi $\mu(T^{-k}F) = \mu(F)$.

Pretpostavimo da su $m, n \in N$ takvi da je $n > m$ i $T^{-n}F \cap T^{-m}F \neq \emptyset$. Tada postoji $y \in T^{-n}F \cap T^{-m}F$ i $T^m y \in F$. Takođe važi $T^{n-m}(T^m y) = T^n y \in F \subset B$, što protivreči definiciji skupa F. Dakle $\{T^n\}_{n \geq 0}$ je familija međusobno disjunktnih skupova, pa važi

$$\sum_{k=0}^{\infty} \mu(T^{-k}F) = \mu(\cup_{k=0}^{\infty}(T^{-k}F)) \leq \mu(X) = 1.$$

Iz prethodne jednakosti i činjenice da je za svaki $k \geq 1$ važi $\mu(T^{-k}F) = \mu(F)$ zaključujemo da je $\mu(T^{-k}F) = \mu(F) = 0$ za svaki $k \geq 1$, odakle dobijamo $\mu(B \backslash E) \leq 0$, odnosno $\mu(B \backslash E) = 0$ što je i trebalo dokazati.$\square$

Transformaciju T nazivamo *ergodičnom* ukoliko, za svaki $B \in \mathcal{B}$, za koji je $T^{-1}B = B$, važi ili $\mu(B) = 0$, ili $\mu(B) = 1$. Primetimo da, ako, za neki $B \in \mathcal{B}$, važi $T^{-1}B = B$, onda takođe važi i $T^{-1}(X \setminus B) = X \setminus B$, pa se izučavanje dinamike transformacije T na X može svesti na njeno izučavanje na samo jednom od podskupova B ili $X \setminus B$. Slučaj kada je $0 < \mu(B) < \mu(X) = 1$ postaje jednostavniji, jer kada je $\mu(B) = 0$ ili $\mu(B) = 1$, onda se razmatranje kada je skup B ili $X \setminus B$ mere nula može izostaviti, jer se tada ne menja ponašanje transformacije T sa stanovišta teorije mere.

Teorema. *Transformacija T je ergodična ako i samo ako svaka funkcija $f \in L^2(\mu)$, koja, za skoro svaki $x \in X$, zadovoljava uslov $f(Tx) = f(x)$, je konstantna skoro svuda.*

[14]H. Poincaré (1854–1912), slavni francuski matematičar, fizičar, inženjer i filozof nauke.

Dokaz. Potrebnost uslova. Posmatraćemo skupove

$$B(k,n) = \{x \in X | 2^{-n}k \leq f(x) < 2^{-n}(k+1)\}$$

za $k \in \mathbf{Z}$ i $n \in \mathbf{N}$. Ovi skupovi su invarijantni u odnosu na T do na skupove mere nula, ako je $f(Tx) = f(x)$, za skoro svaki $x \in X$, pa je ili $\mu(B(k,n)) = 0$, ili $\mu(B(k,n)) = 1$. Kako je, za fiksirani $n \in \mathbf{N}$, X unija međusobno disjunktinih skupova $B(k,n))$, za $k \in \mathbf{Z}$, tačno jedan od njih ima meru 1, i to sa indeksom $k = k(n)$. Tada je i skup

$$B = \bigcap_{n \in \mathbf{N}} B(k(n), n)$$

mere 1, f je konstantna na B, pa i skoro svuda na X.

Dovoljnost uslova. Neka je, za neki $B \in \mathcal{B}$, $T^{-1}B = B$. Tada karakteristična funkcija χ_B pripada klasi $L^2(\mu)$, i, za sve $x \in X$, važi $\chi_B(Tx) = \chi_B(x)$. Dakle, χ_B je konstantna i ima vrednost 0 ili 1. Stoga $\mu(B) = \int \chi_B d\mu$ ima vrednost 0 ili 1.$\square$

Navodimo i nekoliko primera:

Primer. Neka je $S^1 = \{|z| = 1\} \subseteq \mathbf{C}$ snabdeven Lebesgue–ovom merom. Transformacija $T : S^1 \to S^1$ definisana kao $Tz = az$, za $a \in S^1$, je ergodična ako i samo ako a nije koren jedinice, tj. ako T nije periodična.$\triangle$

Primer. Neka je $Y = \{0, 1, \ldots, l-1\}$ prostor stanja i na njemu definisana verovatnosna mera p koja zadovoljava sledeće uslove: $p_i > 0$, za $i = 0, 1, \ldots, l-1$, i $\sum_{i=0}^{l-1} p_i = 1$. Neka je, dalje, $X = Y^{\mathbf{Z}}$ i $\mathcal{B}$ generisan skupovima oblika $A = \{(x_n)_{n \in \mathbf{Z}} | x_m \in Y, \text{za } m < n_1 \vee m > n_2, x_{n_1} \in A_{n_1}, \ldots, x_{n_2} \in A_{n_2}\}$, za $n_1 \leq n_2$ i $A_i \subseteq Y$, sa merom generisanom verovatnoćom p:

$$\mu\{x | x_{i_1} = j_1 \wedge \cdots \wedge x_{i_k} = j_k\} = p_{j_1} \cdot \ldots \cdot p_{j_k}$$

gde je $j_i \in Y$, za sve i, $i = 0, 1, \ldots, l-1$.$\triangle$

Lema. *Pomeranje $\sigma : X \to X$, $(x_n)_{n \in \mathbf{Z}} \mapsto (y_n)_{n \in \mathbf{Z}}$, gde je skup X dat u prethodnom primeru, definisano kao $y_n = x_{n+1}$, je ergodično.*

Ovako definisano pomeranje σ se naziva još i *Bernoulli*[15]*–jevim pomeranjem.*

Dokaz. Neka je $B \in \mathcal{B}$, takav da važi $\sigma^{-1}B = B$. Pokazaćemo, zapravo, da ako je u tački i vrednost svih nizova sadržanih u E restrikovana, npr. $x_i = j$, onda invarijantnost skupa E u odnosu na pomeranje povlači takvu invarijantnost za svaku drugu tačku. Svaki takav skup koji vrši restrikciju vrednosti svih članova niza, ustvari, briše meru. Sa druge strane, ukoliko ne bi bilo takvih restrikcija, E bi morao biti ceo prostor. Formalno, za svaki $\varepsilon > 0$ možemo naći skup A, u prethodno opisanom obliku, takav da je $\mu(E \triangle A) < \varepsilon$, jer $E \in \mathcal{B}$. Stoga $|\mu(E) - \mu(A)| < \varepsilon$. Pretpostavimo sada da je A skup sa datim svojstvima i da je $n_0 > n_2 - n_1$ i $B = \sigma^{-n_0}A$. Tada će biti $\mu(A) = \mu(B)$, pa i $\mu(A \cap B) = \mu(A)\mu(B) = (\mu(A))^2$. Kako je $\sigma^{-1}E = E$, imamo takođe da je $\mu(E \triangle B) = \mu(\sigma^{-n_0}E \triangle \sigma^{-n_0}A) =$

[15]Jacob Bernoulli (1654-1705) jedan od najznačajnijih članova slavne porodice matematičara Bernoulli.

$\mu(E \triangle A)$. Štaviše, važi i: $\mu(E \triangle (A \cap B)) < 2\varepsilon$, $|\mu(E) - \mu(A \cap B)| < 2\varepsilon$ i

$$\begin{aligned} |\mu(E) - (\mu(E))^2| &\leq |\mu(E) - \mu(A \cap B)| + |\mu(A \cap B) - (\mu(E))^2| \\ &\leq 2\varepsilon + |(\mu(A))^2 - (\mu(E))^2| \\ &\leq 2\varepsilon + \mu(A)|\mu(A) - \mu(E)| + \mu(E)|\mu(A) - \mu(E)| < 4\varepsilon \end{aligned}$$

odakle, dalje, sledi da je $\mu(E) = (\mu(E))^2$, tj. $\mu(E) = 0 \vee \mu(E) = 1$.□

Primer. Neka je $X = Y^{\mathbf{Z}}$ i σ sa svojstvima kao u prethodnoj lemi, osim što sada drugačije definišemo meru. Neka je $A = (a_{ij})$ matrica tipa $(n-1)\times(n-1)$ sa nenegativnim elementima, takvim da je $\sum_{j=0}^{n-1} a_{ij} = 1$, za svaki i $(1 \leq i \leq n-1)$. Tada postoji $p = p(p_0, \ldots, p_{n-1})$, takav da je $p_i > 0$, za svaki i $(1 \leq i \leq n-1)$, $\sum_{i=0}^{n-1} p_i = 1$ i $pA = p$. Neka je, dalje, $\mu_A\{x | x_i = j_0 \wedge x_{i+1} = j_1, \ldots, x_{i+k} = j_k\} = p_{j_0} a_{j_0 j_1} \ldots a_{j_{k-1} j_k}$, gde p_{j_0} označava verovatnoću pojavljivanja simbola j_0, a $a_{j_l j_{l+1}}$ verovatnoću prelaza sa simbola j_l na simbol j_{l+1}. Mera μ_A se može produžiti do verovatnosne mere na $(X, \mathcal{B})$. Pomeranje σ ostavlja meru μ_A invarijantnom. Ovako definisano pomeranje σ se naziva još i *Markovljevim*[16] *pomeranjem.* Primetimo da se Bernoulli–jevo pomeranje može dobiti kao poseban slučaj Markovljevog pomeranja. Može se dokazati da je pomeranje σ ergodično na $(X, \mathcal{B}, \mu_A)$ ako i samo ako A je ireducibilan, što znači da za sve i, j, postoji $m \in \mathbf{N}$, takav da je $a_{ij}^{(m)} > 0$, gde je, za $m \in \mathbf{N}$, $A^{(m)} = (a_{ij}^{(m)})$.△

Navodimo sledeća tvrđenja koja zauzimaju centralno mesto u ergodičkoj teoriji.

Wiener[17]—Yoshida[18]—Kakutani[19]–jeva ergodička teorema. *Neka je* $f \in L^1(\mu)$ *i* $f^* = \sup\{\frac{1}{n}\sum_{k=0}^{n-1} f(T^k x) | n \geq 1\}$. *Tada važi*

$$\int_{f^*>0} f d\mu \geq 0.$$

Birkhoff—Khinchin–ova ergodička teorema. *Neka je* $(X, \mathcal{B}, \mu)$ *prostor verovatnoća,* $T : X \to X$ *transformacija koja čuva meru i* $f \in L^1(X, \mathcal{B}, \mu)$. *Tada postoji funkcija* $\widetilde{f}$ *za koju važi:*

(1) $\lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} f(T^i x) = \widetilde{f}(x)$ *skoro svuda;*

(2) $\widetilde{f}(Tx) = \widetilde{f}(x)$ *skoro svuda;*

(3) $\widetilde{f} \in L^1$ *i* $\|\widetilde{f}\|_1 \leq \|f\|_1$*;*

(4) $\frac{1}{n}\sum_{i=0}^{n-1} f(T^i x)$ *konvergira ka* $\widetilde{f}$ *u* L^1*;*

[16] Andrej Andrejevič Markov (1856–1922) ruski matematičar poznat po svom doprinosu teoriji stohastičkih procesa.

[17] Norbert Wiener (1894–1964), osnivač kibernetike i jedan od najplodnijih američkih matematičara XX veka sa naročito značajnim doprinosima teoriji stohastičkih procesa, računarstvu i teoriji sistema upravljanja.

[18] Kosaku Yoshida (1909–1990), japanski matematičar.

[19] Shizuo Kakutani (1911–2004), američki matematičar japanskog porekla poznat po teoremi o nepokretnoj tački i doprinosu matematičkoj ekonomiji.

(5) ako za $A \in \mathcal{B}$ važi $T^{-1}(A) = A$, onda $\int_A f d\mu = \int_A \widetilde{f} d\mu$;

Dokaz. (1) Za α i β, za koje je $\alpha < \beta$, definišemo skupove

$$E_{\alpha,\beta} = \{x \in X | \liminf_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} f(T^i x) < \alpha < \beta < \limsup_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} f(T^i x)\}.$$

Dokazaćemo da važi $\mu(E_{\alpha,\beta}) = 0$, za sve α i β. Samim tim će i skup koji se dobija uniranjem skupova $E_{\alpha,\beta}$ po svim racionalnim brojevima α i β biti mere 0, što znači da granična vrednost $\lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} f(T^i x)$ postoji skoro svuda.

Skup $E_{\alpha,\beta}$ je invarijantan podskup skupa $\{f^* > \beta\}$. Primenjujući prethodnu teoremu na restrikciju transformacije T na skup $E_{\alpha,\beta}$, imamo

$$\int_{E_{\alpha,\beta}} f d\mu \geq \beta\mu(E_{\alpha,\beta}).$$

Ako $x \in E_{\alpha,\beta}$, postoji $n \geq 1$ takav da je $\frac{1}{n}\sum_{i=0}^{n-1} f(T^i x) < \alpha$ i važi $E_{\alpha,\beta} \subset \{(-f^*) > -\alpha\}$. Po prethodnoj teoremi imamo $\int_{E_{\alpha,\beta}} (-f) d\mu \geq -\alpha\mu(E_{\alpha,\beta})$, odnosno

$$\int_{E_{\alpha,\beta}} f d\mu \leq \alpha\mu(E_{\alpha,\beta}).$$

Dakle,

$$\beta\mu(E_{\alpha,\beta}) \leq \int_{E_{\alpha,\beta}} f d\mu \leq \alpha\mu(E_{\alpha,\beta}).$$

Kako je $\alpha < \beta$, gornja nejednakost važi jedino u slučaju da je $\mu(E_{\alpha,\beta}) = 0$.

(2) Očigledno važi $\widetilde{f}T = \widetilde{f}$ skoro svuda.

(3) U dokazu ovog tvrđenja se koristi Fatuova lema. Iz

$$\left| \frac{1}{n}\sum_{i=0}^{n-1} f(T^i) \right| \leq \frac{1}{n}\sum_{i=0}^{n-1} |f|(T^i)$$

imamo $|\widetilde{f}| \leq |f|^-$, gde je $|f|^- = \liminf_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} |f|(T^i)$, odakle sledi

$$\int |\widetilde{f}| d\mu \leq \int |f|^- d\mu \leq \liminf_{n\to\infty} \int \frac{1}{n}\sum_{i=0}^{n-1} |f(T^i x)| = \int |f| d\mu < \infty.$$

(4) U slučaju ograničenih funkcija, L^1 konvergencija sledi iz teoreme o konvergenciji ograničenih funkcija, po kojoj ako za merljive funkcije $f_1, f_2, \ldots$ i pozitivnu konstantu M važi $\|f_n\|_\infty \leq M$, za $n = 1, 2, \ldots$ i $\lim_{n\to\infty} f_n(x) = f(x)$ skoro svuda, onda je $\|f_n - f\|_p \to 0$ za sve $0 < p < \infty$. Opšti slučaj se može dokazati aproksimiranjem ograničenim funkcijama i pomoću dela (3) ove teoreme. Bez gubitka opštosti, pretpostavimo da je

$f \geq 0$, jer u suprotnom možemo posmatrati funkcije $f^+ = \max\{f, 0\}$ i $f^- = -\min\{f, 0\}$, gde je $f = f^+ - f^-$. Ako je funkcija g ograničena i važi $0 \leq g \leq f$, onda

$$\left\| \frac{1}{n} \sum_{i=0}^{n-1} fT^i - \widetilde{f} \right\|_1 \leq \left\| \frac{1}{n} \sum_{i=0}^{n-1} (fT^i - gT^i) \right\|_1 + \left\| \frac{1}{n} \sum_{i=0}^{n-1} gT^i - \widetilde{g} \right\|_1 + \left\| \widetilde{g} - \widetilde{f} \right\|_1 .$$

Prvi i treći sabirak gornje sume su ograničeni normom $\|g - f\|_1$, koja odgovarajućim izborom funkcije g može biti proizvoljno mala, dok drugi sabirak teži 0 kad $n \to \infty$, prema teoremi o konvergenciji ograničenih funkcija. Time je dokazano tvrđenje (4).

(5) Ovo tvrđenje dokazujemo pomoću tvrđenja (4). Važi:

$$\begin{aligned} \left| \int_A f d\mu - \int_A \widetilde{f} d\mu \right| &= \left| \int_A \left(\frac{1}{n} \sum_{i=0}^{n-1} fT^i - \widetilde{f} \right) d\mu \right| \\ &\leq \int_A \left| \frac{1}{n} \sum_{i=0}^{n-1} fT^i - \widetilde{f} \right| d\mu \\ &= \left\| \frac{1}{n} \sum_{i=0}^{n-1} fT^i - \widetilde{f} \right\|_{L^1(A)} \to 0. \square \end{aligned}$$

Napomenimo da ako je prostor X beskonačne mere tvrđenja (1), (2) i (3) prethodne teoreme važe, dok tvrđenja (4) i (5) važe u slučaju invarijantnih skupova konačne mere.

Teorema. *Transformacija T je ergodična ako i samo ako za svaku funkciju $f \in L^1(\mu)$ važi $\widetilde{f}(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x) = \int_X f d\mu$ skoro svuda.*

Dokaz. U ovom dokazu ćemo koristiti tvrđenje prema kojem je transformacija T ergodička ako i samo ako svaka funkcija $f \in L^1(\mu)$, koja, za skoro svaki $x \in X$, zadovoljava uslov $f(Tx) = f(x)$, je konstantna skoro svuda. Pomenuto tvrđenje navedeno je i dokazano u ovom radu u slučaju da $f \in L^2(\mu)$.

Pretpostavimo da je T ergodična. Tada, kako je $\widetilde{f}$ invarijatna u odnosu na T, imamo da je $\widetilde{f}$ konstantna skoro svuda i iz prethodne teoreme sledi da je $\int_X f d\mu = \int_X \widetilde{f} d\mu = \widetilde{f}(x)$ skoro svuda.

Sada, pretpostavimo da je za svaku funkciju $f \in L^1(\mu)$, funkcija $\widetilde{f}$ konstantna skoro svuda. Ako je f invarijantna funkcija u $L^1(\mu)$, onda je $\frac{1}{n} \sum_{i=0}^{n-1} f(T^i x) = f(x)$ skoro svuda, pa je $f = \widetilde{f}$ skoro svuda. Dakle, i funkcija f je konstantna, pa je T ergodična transformacija.$\square$

Slično, za jedan neprekidan ergodički proces $(T_t)_{t\geq 0}$ važi:

$$\lim_{T\to\infty} \frac{1}{T} \int f(T^t x) dt = \int_X f d\mu$$

za skoro svaki x.

Ovo će značiti da kada posmatramo jednu vremensku seriju, u slučaju kada je ona ergodična, njena srednja vrednost će konvergirati odgovarajućoj očekivanoj vrednosti.

Ergodička teorema o srednjoj vrednosti u Hilbert[20]–ovom prostoru. *Neka je U linearna kontrakcija na Hilbertovom prostoru H, $M = \{f \in H | Uf = f\}$ i neka je $P : H \to M$ projekcija prostora H na prostor M. Tada niz $\frac{1}{n}\sum_{i=0}^{n-1} U^i f$ konevrgira u H ka Pf, za svaku $f \in H$.*

Dokaz. Neka je N zatvorenje lineala skupa $\{g - Ug | g \in H\}$. Dokazaćemo da za potprostore M i N, prostora H, važi $N^\perp = \{h \in H | (h, s) = 0, \ \forall s \in N\} = M$. Neka je $h \in N^\perp$. Tada za sve $g \in H$ važi $0 = (h, g - Ug) = (h, g) - (h, Ug) = (h, g) - (U^*h, g) = (h - U^*h, g)$, odnosno $h = U^*h$. Važi:

$$\begin{aligned}\|Uh - h\|^2 &= \|Uh\|^2 - (h, Uh) - (Uh, h) + \|h\|^2 \\ &= \|Uh\|^2 - (U^*h, h) - (h, U^*h) + \|h\|^2 \\ &\le \|h\|^2 - (h, h) - (h, h) + \|h\|^2 \\ &= 0.\end{aligned}$$

Dakle, $Uh = h$, što znači da $h \in M$, odnosno $N^\perp \subset M$.

Obrnuto, neka $h \in M$. Tada je $Uh = h$, odakle prema prethodnom izvođenju primenjenom na U^* i činjenici da je U^* takođe kontrakcija, važi $U^*h = h$. Dakle, za sve $g \in H$ važi $(h, g - Ug) = (h, g) - (h, Ug) = (h - U^*h, g) = 0$. Time smo dokazali da $h \in N^\perp$, a samim tim i $N^\perp = M$.

Sledeći korak jeste da dokažemo da ako $f \in N$, onda red $\frac{1}{n}\sum_{i=0}^{n-1} U^i f$ konvergira ka 0. Razmotrimo prvo slučaj kada je $f = g - Ug$, za neki $g \in H$. Tada važi

$$\left\| \frac{1}{n}\sum_{i=0}^{n-1} U^i f \right\| = \frac{1}{n}\|g - U^n g\| \le \frac{2}{n}\|g\| \to 0, \quad n \to \infty.$$

Ako $f \in N$, onda postoji niz $\{g_n\}$ u H, takav da niz $f_n = g_n - Ug_n$ konvergira ka f. Tada važi

$$\left\| \frac{1}{n}\sum_{i=0}^{n-1} U^i f \right\| \le \left\| \frac{1}{n}\sum_{i=0}^{n-1} U^i (f - f_k) \right\| + \left\| \frac{1}{n}\sum_{i=0}^{n-1} U^i f_k \right\|,$$

za sve k. Neka je $\varepsilon > 0$ proizvoljan i izaberimo k tako da $\|f - f_k\| < \frac{\varepsilon}{2}$. Neka je n takav da važi $\|\frac{1}{n}\sum_{i=0}^{n-1} U^i f_k\| < \frac{\varepsilon}{2}$. Tada je $\|\frac{1}{n}\sum_{i=0}^{n-1} U^i f_k\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$.

Konačno, ako je $f \in H$ proizvoljna, onda postoji jedinstvena $f_0 \in N$ takva da je $f = f_0 + Pf$. Tada je

$$\left\| \frac{1}{n}\sum_{i=0}^{n-1} U^i f - Pf \right\| = \left\| \frac{1}{n}\sum_{i=0}^{n-1} U^i (f_0 + Pf) - Pf \right\| = \left\| \frac{1}{n}\sum_{i=0}^{n-1} U^i f_0 \right\| \to 0, \quad n \to \infty. \square$$

Bez dokaza navodimo i sledeće tvrđenje:

[20]David Hilbert (1862–1943), nemački matematičar, jedan od najuticajnijih matematičara s kraja XIX i početka XX veka.

Von Neumann–ova ergodička teorema. *Ako je* $f \in L^p(\mu)$ $(1 \leq p < \infty)$, *onda postoji* $\widetilde{f} \in L^p(\mu)$ *za koju važi* $\widetilde{f} \circ T = \widetilde{f}$, *tako da* $\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x) = \widetilde{f}$ *u* $L^p(\mu)$.

Imajući u vidu da transformacija $T : X \to X$ definiše linearni operator $U : L^2 \to L^2$ kao $(Uf)(x) = f(Tx)$, koji je ujedno i invertibilna izometrija, kao neposredna posledica ergodičke teoreme o srednjoj vrednosti u Hilbert–ovom prostoru, može se zaključiti da važi von Neumann–ova ergodička teorema u specijalnom slučaju $p = 2$, što predstavlja oblik ove teoreme koji se najčešće navodi u literaturi.

Direktna posledica ergodičkih teorema je sledeće tvrđenje: *T je ergodična ako i samo ako za sve $A, B \in \mathcal{B}$ važi:*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mu(T^{-i}(A \cap B)) = \mu(A)\mu(B)$$

Kao dobre izvore o ergodičkoj teoriji imamo u vidu P. R. Halmos (1956), A. Katok, B. Hasselblatt (1995) i K. Petersen (1983).

§3. Entropija

Entropija je veličina koja karakteriše neuređenost (ili stohastičnost) sistema. Prema Gaspard–u (P. Gaspard (2005)), prve nagoveštaje definicije entropije nalazimo u pionirskim radovima S. Carnot–a iz 1824. godine o osnovama funkcionisanja parne mašine i R. Clausius–a koji, između 1851. i 1865. godine razvija koncept entropije (i uvodi za nju ime prema starogrčkoj reči '$\tau\rho o\pi\acute{\eta}$' označavajući transformaciju). Ipak, u literaturi se pojam entropije, u oznaci S, koji se pojavljuje u termodinamici, vezuje za Ludwig–a Boltzmann–a i period od 1896. do 1898. godine, kada je *entropija* S okarakterisana kao rešenje[21]

$$S = k \log W$$

diferencijalne jednačine

$$dS = \frac{dQ}{W}$$

gde je k Boltzmann–ova konstanta i W broj mikrostanja kompatibilnih sa zadatim makrostanjem sistema koji sadrži n čestica. Uz Boltzmann–a, M. Planck i J. W. Gibbs prvi ističu statističko značenje entropije i uvode je u mehaniku, kao verovatnosnu funkciju.

Drugačiji koncept entropije[22], primeren primenama u teoriji informacija, vezuje se za 1949. godinu i ime Claude–a Shannon–a, i njemu ćemo posvetiti malo više pažnje na

[21]Na Boltzmann–ovom grobu se nalazi epigram $S = k \log W$ (V. Benci, G. Menconi).

[22]Duhovita je anegdota vezana za davanje imena ovom pojmu u kontekstu teorije informacija, koju ćemo, prema A. Garrido (2009), ovde parafrazirati. Naime, na pitanje postavljeno od strane C. Shannon–a i predlog da taj novi pojam koji je definisao i čija svojstva izučava nazove 'neodređenost', J. von Neumann je odgovorio: "Ja bih to nazvao 'entropijom', iz dva razloga. Prvi je, što se vaša funkcija neodređenosti već koristi u statističkoj

stranicama koje slede. U ovom slučaju se radi o temporalnoj meri neodređenosti jednog stohastičkog sistema. Shannon polazi od slučajnog događaja A sa mogućim ishodima $a_1, ..., a_m$ čija je raspodela verovatnoća data kao

$$A : \begin{pmatrix} a_1 & ... & a_m \\ p_1 & ... & p_m \end{pmatrix}, \quad \text{za} \quad (\forall i) p_i \geq 0 \;\; \text{i} \quad p_1 + ... + p_m = 1,$$

i *entropiju* $H(A)$ događaja A definiše kao:

$$H(A) = -\sum_{i=1}^{m} p_i \log p_i$$

Prirodno uopštenje pojma entropije nastalog u teoriji informacija predstavlja entropija sa kakvom se susrećemo u teoriji mere. Njoj ćemo posvetiti deo ovog rada takođe. U okvirima teorije mere, polazeći od proizvoljne merljive particije $\alpha = \{A_i | i \in I\}$, što podrazumeva da je $(\forall i)(1 \leq i \leq n \rightarrow \mu(A_i) > 0)$, $(\forall i, j)(1 \leq i \leq n \wedge 1 \leq j \leq n \wedge i \neq j \rightarrow A_i \cap A_j = \emptyset)$ i $\mu(X \setminus \cup_{i=1}^{n} A_i) = 0$, prostora X snabdevenog merom μ, entropija te particije se definiše kao:

$$H(\alpha) = -\sum_{i \in I} \mu(A_i) \log \mu(A_i)$$

što predstavlja očigledno uopštenje prethodnog koncepta, ali i koncepta u kojem bi mera μ mogla biti i verovatnosna mera. Ovakav tretman entropije vodi, daljim procesom apstrakcije, pojmu topološke entropije, kojim se ovde nećemo baviti.

Ipak, naša centralna tema biće četvrti koncept entropije definisan 1959. godine od strane Andreja N. Kolmogorova i Jakova G. Sinaija, koji se odnosi na determinističke dinamičke sisteme, čemu će biti posvećeno poslednje poglavlje rada. Oni, dakle, polaze od determinističkog dinamičkog sistema, snabdevenog invarijantnom merom, posmatrajući njegova stanja $\omega_1 \omega_2 \ldots \omega_n$ u faznom prostoru tokom vremenskih intervala $\triangle t$ i promene trajektorije sistema u vremenu. U tom slučaju *entropiju* h_{KS} sistema definisu kao:

$$h_{KS} = \sup_{\mathcal{P}} \lim_{n \to \infty} \left(-\frac{1}{n \triangle t} \sum_{\omega_1 \omega_2 \ldots \omega_n} \mu_{\omega_1 \omega_2 \ldots \omega_n} \ln \mu_{\omega_1 \omega_2 \ldots \omega_n} \right)$$

gde se supremum uzima po svim mogućim particijama $\mathcal{P}$, a sa μ je označena data invarijantna mera dinamičkog sistema.

§3.1. Entropija u teoriji informacija

Na početku ovog dela teksta dajemo prikaz pojma entropije na način kako se on istorijski razvijao, da bismo na kraju posvetili pažnju njegovom savremenom i apstraktnom tretmanu u okviru ergodičke teorije.

---

mehanici pod tim imenom, što znači da ona već ima ime. Drugi, još važniji, je taj što niko ne zna šta je zaista entropija, tako da ćete vi u svakoj debati o njoj uvek unapred imati prednost."

Jedan od fundamentalnih pojmova koji je Shannon uveo u teoriju informacija je pojam *entropije* ili mere neodređenosti nekog slučajnog događaja.

Pretpostavićemo da je A slučajan događaj sa ukupno m različitih mogućih ishoda $a_1, ..., a_m$ takvih da je verovatnoća $p(a_i)$ svakog ishoda a_i $(1 \leq i \leq m)$ pozitivna. Dakle:

$$A : \begin{pmatrix} a_1 & ... & a_m \\ p(a_1) & ... & p(a_m) \end{pmatrix}$$

za

$$(\forall i) p(a_i) > 0 \quad \text{i} \quad p(a_1) + ... + p(a_m) = 1$$

Saglasno našoj intuiciji, mera neodređenosti, u oznaci f, posmatranog slučajnog događaja bi, pre svega, morala zadovoljavati neke opšte uslove koje zadovoljava svaka mera, tzv. *aksiome mere:*

$$f(A) \geq 0 \qquad \textit{(nenegativnost)}$$
$$f(AB) = f(A) + f(B) \qquad \textit{(aditivnost)}$$

gde se, u ovakvom kontekstu, za uslov aditivnosti pretpostavlja nezavisnost slučajnih događaja A i B.

Ako već merimo neodređenost slučajnog događaja, onda bi, svakako, neosporna bila i činjenica da u slučaju kada važi $m = 1$, tj. kada događaj A više faktički i nije slučajan događaj, nego je, sa verovatnoćom 1, dakle, izvesno da će se desiti a_1, mera neodređenosti takvog događaja iznosi 0. Znači: $f(A) = 0$.

Pored toga, mera neodređenosti bi morala dostići maksimalnu vrednost onda kada je slučajan događaj A definisan nekim brojem jednako verovatnih ishoda $a_1, ..., a_m$:

$$p(a_i) = \frac{1}{m} \qquad (1 \leq i \leq m)$$

Jasno je i to da bi se sa porastom broja m jednako verovatnih ishoda morala uvećavati i mera neodređenosti slučajnog događaja. Drugim rečima, funkcija f bi morala biti i monotono rastuća.

Ukoliko još usvojimo zahtev da mera neodređenosti slučajnog događaja treba da bude neprekidna, onda se prostor u kojem tražimo takvu funkciju bitno sužava. Naime, imajući u vidu uslov aditivnosti, posmatraćemo funkcionalnu jednačinu:

$$f(xy) = f(x) + f(y)$$

Ova jednačina ima jedinstveno neprekidno rešenje definisano na skupu svih pozitivnih realnih brojeva, i to će biti logaritamska funkcija. Proverimo to: neka je f neprekidna funkcija definisana na skupu $(0, +\infty)$ koja zadovoljava posmatranu funkcionalnu jednačinu. Uvodeći smenu: $x = a^t$ i $g(t) = f(a^t)$, dobijamo: $t = \log_a x$ i $f(x) = g(\log_a x)$. Biće takođe: $g(x+y) = f(a^{x+y}) = f(a^x a^y) = f(a^x) + f(a^y) = g(x) + g(y)$, tj. funkcija g je aditivna. Odavde, dalje, esencijalno koristeći neprekidnost, na čemu se ovde nećemo detaljnije zadržavati, sledi da je funkcija g linearna, odnosno da je $g(x) = kx+n$. Međutim, iz jednakosti $f(1) = f(1 \cdot 1) = f(1) + f(1)$, sledi $f(1) = 0$ i $g(0) = 0$, odnosno $n = 0$. Dakle, $f(x) = k \cdot \log_a x$ za $k \neq 0$, kao i $f(x) = 0$ za $k = 0$. Zanemarujući ovaj drugi trivijalan slučaj, zaključujemo da bi tražena funkcija posredstvom koje bi trebalo izraziti entropiju, morala biti definisana preko logaritamske funkcije.

Prvi značajniji pokušaj egzaktnog definisanja entropije u teoriji informacija potiče od Hartley–a (R. V. L. Hartley) 1928. godine, a kasnije opšte prihvaćenu definiciju, koju ćemo ovde navesti i analizirati, daje Shannon 1948. godine. Nedostatak Hartley–eve definicije u odnosu na Shannon–ovu je što ta definicija ne uzima u obzir verovatnoće pojedinih ishoda već samo njihov ukupan broj. Naime, po Hartley-u, mera neodređenosti nekog slučajnog događaja sa m mogućih ishoda iznosi $\log m$, nezavisno od toga kakva je funkcija raspodele verovatnoća posmatranog slučajnog događaja. Koncept entropije koji potiče od Hartley–a nije, zapravo, ništa originalno u odnosu na entropiju koja potiče iz termodinamike. Jedina novost jeste kontekst teorije informacija u kojem se sada pojavljuje ovaj pojam.

Navedimo sada Shannon–ovu definiciju entropije.

Neka je A slučajan događaj sa mogućim ishodima $a_1, ..., a_m$ čija je raspodela verovatnoća data kao

$$A : \begin{pmatrix} a_1 & ... & a_m \\ p_1 & ... & p_m \end{pmatrix}$$

za

$$(\forall i) p_i \geq 0 \quad \text{i} \quad p_1 + ... + p_m = 1$$

Tada *entropiju* slučajnog događaja A, u oznaci $H(A)$ ili $H(p_1, ..., p_m)$, definišemo sledećom formulom:

$$H(A) = -\sum_{i=1}^{m} p_i \log p_i$$

gde kao osnovicu logaritma koji učestvuje u ovoj definiciji, ali i nadalje, osim ako to eksplicitno nije drugačije naglašeno, po dogovoru, uzimamo da je 2, čime se odlučujemo za BIT kao jedinicu mere entropije. (*Bit* je kovanica od engleske sintagme *binary digit,* što znači *binarna cifra.*)

Napomena. Ukoliko bi se i desilo da je verovatnoća nekog ishoda $p_i = 0$, imajući u vidu da je $\lim_{x \to 0+} x \log x = 0$, opravdano bi bilo i odgovarajući sabirak u gornjoj definiciji izostaviti.

Znajući kako se logaritam jedne osnove može predstaviti posredstvom logaritma druge osnove: $\log_a b = \log_a c \log_c b$, vidimo da dogovor da se za osnovicu logaritma uzme broj 2, nije toliko bitan. Glavni razlozi za ovo su praktične prirode. Naime, kada je A slučajan događaj sa svega dva jednako verovatna ishoda: $p_1 = p_2 = \frac{1}{2}$, onda je njegova entropija jednaka upravo jedinici mere: $H(A) = 1BIT$.

Ukažimo sada na neke od osnovnih osobina entropije koje neposredno slede iz njene definicije.

Teorema. *Ako je m broj mogućih ishoda slučajnog događaja A, onda*

$$0 \leq H(A) \leq \log m$$

Dokaz. Nije teško proveriti da funkcija $f(x) = \ln x - x + 1$ dostiže svoj maksimum za $x = 1$. Kako je $f(1) = 0$, to je i $f(x) \leq 0$, tj. $\ln x \leq x - 1$. Dakle, važi: $\log x \leq (x-1) \log e$.

Koristeći ovu nejednakost i elementarne algebarske transformacije, izvodimo:

$$\begin{aligned}
H(A) - \log m &= -\sum_{i=1}^{m} p_i \log p_i - \log m \\
&= \sum_{i=1}^{m} p_i \log \frac{1}{p_i} - \log m \sum_{i=1}^{m} p_i \\
&= \sum_{i=1}^{m} p_i \log \frac{1}{p_i m} \\
&\leq \sum_{i=1}^{m} p_i (\frac{1}{p_i m} - 1) \log e \\
&= \sum_{i=1}^{m} (\frac{1}{m} - p_i) \log e \\
&= \log e (\sum_{i=1}^{m} \frac{1}{m} - \sum_{i=1}^{m} p_i) = 0
\end{aligned}$$

Prema tome: $H(A) - \log m \leq 0$, odnosno: $H(A) \leq \log m$. Drugi deo nejednakosti: $H(A) \geq 0$, neposredno sledi iz pretpostavke da $0 \leq p_i \leq 1$ $(1 \leq i \leq m)$, osnovnih svojtava logaritamske funkcije i definicije entropije.$\square$

Primetimo, nakon ovog dokaza, da važi:

$$H(A) = \log m \quad \textit{akko} \quad (\forall i)(1 \leq i \leq m \rightarrow p_i = \frac{1}{m})$$

Drugim rečima, entropija $H(A)$ dostiže maksimalnu vrednost onda kada su svi ishodi a_i slučajnog događaja A jednako verovatni, što se, zapravo, i moralo očekivati.

Ovo je prilika da pomenemo i pojam *propusne informacione moći alfabeta.* Naime, količina informacija prosečno sadržana u jednom simbolu nekog alfabeta dostiže svoju najveću vrednost ukoliko je pojavljivanje svakog simbola jednako verovatno. Mera te količine informacija će biti logaritam ukupnog broja simbola posmatranog alfabeta, što će predstavljati njegovu propusnu informacionu moć.

Uz neposredno korišćenje definicije entropije i opštih osobina nezavisnih slučajnih događaja, moguće je dokazati sledeće karakteristično tvrđenje:

Teorema. *Ako su A i B nezavisni slučajni događaji, onda*

$$H(AB) = H(A) + H(B)$$

Ukratko, prethodne dve teoreme nam tvrde da entropija definisana ovako ima osobine mere, nenegativna je i aditivna, a najveću vrednost dostiže u slučaju kada je i neizvesnost realizacije proizvoljnog ishoda najveća, što će biti slučaj sa jednakoverovatnim ishodima.

Od ostalih značajnih osobina entropije ističemo sledeće:

a) entropija je *neprekidna funkcija*;

b) entropija je *simetrična funkcija* svojih argumenata:

$$H(p_1, ..., p_m) = H(p_{i_1}, ..., p_{i_m})$$

gde je $(i_1, ..., i_m)$ bilo koja permutacija uređene m–torke $(1, ..., m)$, što bi značilo da vrednost ove funkcije ne zavisi od redosleda mogućih ishoda;

c) važi:

$$H(p_1, ..., p_m) = H(p_1 + p_2, p_3, ..., p_m) + (p_1 + p_2)H\left(\frac{p_1}{p_1 + p_2}, \frac{p_2}{p_1 + p_2}\right)$$

Upravo navedene osobine *a)–c)* su izuzetno interesantne za teorijska razmatranja. Fadejev[23] je pokazao (v. D. K. Fadejev (1956)) da je jedina funkcija koja ima ove osobine upravo funkcija oblika:

$$H(p_1, ..., p_m) = -C\sum_{i=1}^{m} p_i \log p_i$$

gde je C neka konstanta. Ovo znači da je entropiju moguće uvesti aksiomatski, polazeći od pomenutih osobina.

Po analogiji sa slučajnim događajem sa konačno mnogo ishoda, definiše se entropija slučajnog događaja A sa prebrojivo mnogo mogućih ishoda čije su odgovarajuće verovatnoće: $p_1, p_2, ...$, kao:

$$H(A) = -\sum_{n=1}^{\infty} p_n \log p_n$$

uz pretpostavku da ovaj red konvergira.

Slično, ako se radi o slučajnoj promenljivoj neprekidnog tipa, entropija se definiše posredstvom sledećeg nesvojstvenog integrala

$$-\int_{-\infty}^{+\infty} f(x) \log f(x) dx$$

uz pretpostavku da isti postoji. Za funkciju f, koja se javlja pod integralom, uzima se gustina (v. D. S. Jones, (1979)) ili funkcija raspodele posmatrane slučajne promenljive.

Sada ćemo definisati pojam uslovne entropije.

Pretpostavimo da su A i B slučajni događaji

$$A : \begin{pmatrix} a_1 & ... & a_m \\ p(a_1) & ... & p(a_m) \end{pmatrix} \text{ i } B : \begin{pmatrix} b_1 & ... & b_n \\ p(b_1) & ... & p(b_n) \end{pmatrix}$$

i sa $p(b_j|a_i)$ označimo uslovnu verovatnoću slučajnog ishoda b_j slučajnog događaja B pod uslovom da se realizuje ishod a_i događaja A. *Uslovnu entropiju događaja B uz uslov A*, u oznaci $H(B|A)$, definišemo formulom:

$$H(B|A) = \sum_{i=1}^{m} p(a_i)H(B|a_i)$$

[23] D. K. Fadejev (1907–1989), sovjetski matematičar.

gde je

$$H(B|a_i) = -\sum_{j=1}^{n} p(b_j|a_i) \log p(b_j|a_i) \quad (1 \leq i \leq m)$$

što je tzv. *entropija događaja B uz uslov a_i.*

Navedimo i najznačajnije tvrđenje koje se odnosi na uslovnu entropiju.

Teorema. (C. E. Shannon) $0 \leq H(B|A) \leq H(B)$

Dokaz. Očigledno je da će $H(B|A)$ biti nenegativna vrednost. Drugi deo nejednakosti obrazlažemo uz pomoć *Jensenove*[24] *nejednakosti:* ako je funkcija f konkavna na nekom realnom intervalu I, onda za svaki niz $x_1, ..., x_n \in I$ i svaki niz pozitivnih brojeva $q_1, ..., q_n$, takav da je $q_1 + ... + q_n = 1$, važi:

$$\sum_{i=1}^{n} q_i f(x_i) \leq f(\sum_{i=1}^{n} q_i x_i)$$

Da podsetimo, za neprekidnu realnu funkciju f kažemo da je *konkavna* na intervalu I ako

$$(\forall x, y \in I)(\forall p, q > 0)(p + q = 1 \to pf(x) + qf(y) \leq f(px + qy))$$

Jensenovu nejednakost dokazujemo primenjujući princip matematičke indukcije. Bazu indukcije ($n = 2$) opravdava sama definicija konkavnosti. Razmotrimo induktivni korak od n do $n + 1$. Uvedimo oznake:

$$q'_n = q_n + q_{n+1} \quad \text{i} \quad x'_n = \frac{q_n}{q'_n} x_n + \frac{q_{n+1}}{q'_n} x_{n+1}$$

Tada imamo i:

$$q_n f(x_n) + q_{n+1} f(x_{n+1}) = q'_n (\frac{q_n}{q'_n} f(x_n) + \frac{q_{n+1}}{q'_n} f(x_{n+1})) \leq q'_n f(x'_n)$$

gde je, očigledno, nejednakost koju smo dobili, direktna posledica pretpostavke o konkavnosti funkcije f. Koristeći ovu nejednakost, izvodimo:

$$\begin{aligned}\sum_{i=1}^{n-1} q_i f(x_i) + q_n f(x_n) + q_{n+1} f(x_{n+1}) &\leq \sum_{i=1}^{n-1} q_i f(x_i) + q'_n f(x'_n) \\ &\leq f(\sum_{i=1}^{n-1} q_i x_i + q'_n x'_n) \\ &= f(\sum_{i=1}^{n+1} q_i x_i)\end{aligned}$$

[24] J. L. W. V. Jensen (1859–1925), danski matematičar.

Drugu nejednakost u gornjem izvođenju opravdava indukcijska hipoteza. Na ovaj način smo opravdali i indukcijski korak, čime je kompletiran dokaz Jensenove jednakosti.

Podsetimo još da je za dva puta diferencijabilnu funkciju, da bi bila konkavna, dovoljno da njen izvod ne bude pozitivan. Ovako se, na primer, može ustanoviti da je funkcija $f(x) = -x \log x$ konkavna. Primenjujući Jensenovu nejednakost na ovu funkciju, stavljajući da je $q_i = p(a_i)$ i $x_i = p(b_j|a_i)$ za $1 \leq i \leq m$ i $1 \leq j \leq n$, dobijamo da, za svaki j $(1 \leq j \leq n)$, važi:

$$-\sum_{i=1}^{m} p(a_i)p(b_j|a_i) \log p(b_j|a_i) \leq$$

$$\leq -(\sum_{i=1}^{m} p(a_i)p(b_j|a_i)) \log(\sum_{i=1}^{m} p(a_i)p(b_j|a_i))$$

Međutim, znajući da je, prema teoremi potpune verovatnoće

$$\sum_{i=1}^{m} p(a_i)p(b_j|a_i) = p(b_j)$$

imamo, takođe za svaki j $(1 \leq j \leq n)$:

$$-\sum_{i=1}^{m} p(a_i)p(b_j|a_i) \log p(b_j|a_i) \leq -p(b_j) \log p(b_j)$$

odakle, sabirajući ovih n nejednakosti, dobijamo upravo

$$H(B|A) \leq H(B)$$

čime je završen dokaz ove teoreme.$\square$

Napomenimo da važi:

$$H(B|A) = H(B) \quad \textit{akko} \quad A \text{ i } B \text{ su nezavisni slučajni događaji.}$$

Teorema. *Za proizvoljne slučajne događaje A i B važi:*

$$H(AB) = H(A) + H(B|A) = H(B) + H(A|B)$$

Dokaz. Znajući da za proizvoljne slučajne ishode važi

$$p(ab) = p(a)p(b|a) = p(b)p(a|b)$$

i, za svaki i $(1 \leq i \leq m)$:

$$\sum_{j=1}^{n} p(b_j|a_i) = 1$$

imamo:

$$
\begin{aligned}
H(AB) &= -\sum_{i=1}^{m}\sum_{j=1}^{n} p(a_i b_j)\log p(a_i b_j) \\
&= -\sum_{i=1}^{m}\sum_{j=1}^{n} p(a_i)p(b_j|a_i)\log(p(a_i)p(b_j|a_i)) \\
&= -\sum_{i=1}^{m} p(a_i)\log p(a_i)\sum_{j=1}^{n} p(b_j|a_i) - \sum_{i=1}^{m} p(a_i)\sum_{j=l}^{n} p(b_j|a_i)\log p(b_j|a_i) \\
&= H(A) + H(B|A). \square
\end{aligned}
$$

Kao neposredna posledica ovog tvrđenja se može izvesti zaključak da kada su slučajni događaji A i B nezavisni, onda $H(AB) = H(A) + H(B)$.

Uvedeni pojmovi su dovoljni da se definiše i razvije koncept informacije.

Ovde ćemo egzaktno definisati pojam *mere međusobne zavisnosti* slučajnih događaja, odnosno *mere količine informacija.*

Ako sa $I(A, B)$ označimo količinu podataka o događaju A sadržanu u događaju B, ili obrnuto, onda se najznačajnijim rezultatom ovakvog razmatranja može smatrati potpuna karakterizacija međusobne nezavisnosti slučajnih događaja A i B iskazana ovakvim tvrđenjem:

$$I(A, B) = 0 \quad \textit{akko} \quad A \text{ i } B \text{ su nezavisni slučajni događaji.}$$

Inače, funkcija I se definiše kao: $I(A, B) = H(B) - H(B|A)$ i naziva *informacijom.*

Kao neposredne posledice ovakve definicije informacije I i prethodno datih osobina entropije, možemo dobiti sledeći niz tvrđenja:

Teorema. *Za proizvoljne slučajne događaje A i B, $I(A, B) = I(B, A)$.*

Dokaz. Prema jednoj od prethodnih teorema imali smo $H(A) + H(B|A) = H(B) + H(A|B)$ odakle je

$$I(A, B) = H(B) - H(B|A) = H(A) - H(A|B) = I(B, A). \square$$

Teorema. *Za svaka dva slučajna događaja A i B, $I(A, B) = H(A) + H(B) - H(AB)$.*

Teorema. *Za proizvoljne slučajne događaje A i B, $I(A, B) = 0$ akko A i B su nezavisni slučajni događaji.*

Teorema. *Za proizvoljne slučajne događaje A i B, $0 \leq I(A, B) \leq H(A)$.*

Teorema. *Za svaki slučajan događaj A, $I(A, A) = H(A)$.*

U kontekstu ergodičke teorije ova tvrđenja dobijaju i nove interpretacije kojima ćemo se ovde baviti.

§3.2. Entropija u teoriji mere

Pretpostavimo da je $(X, \mathcal{B}, \mu)$ prostor verovatnoća, $T : X \to X$ transformacija koja čuva meru, $\alpha = \{A_1, \ldots, A_n\}$ particija skupa X. Tada *entropiju particije* α definišemo kao

$$H(\alpha) = -\sum_{i=1}^{n} \mu(A_i) \log \mu(A_i).$$

Ako su $\alpha = \{A_1, \ldots, A_n\}$ i $\beta = \{B_1, \ldots, B_m\}$ merljive particije skupa X, tada sa $\alpha \vee \beta$ označavamo sledeću particiju: $\alpha \vee \beta = \{A_i \cap B_j | 1 \leq i \leq n \wedge 1 \leq j \leq n\}$.

Entropiju h transformacije T u odnosu na α, u oznaci $h(\alpha, T)$, definišemo kao:

$$h(\alpha, T) = h_\mu(\alpha, T) = \lim_{n \to \infty} \frac{1}{n} H(\alpha \vee T^{-1}\alpha \vee \cdots \vee T^{-n+1}\alpha).$$

Uslovnu entropiju particije α *u odnosu na particiju* β definišemo kao

$$H(\alpha|\beta) = -\sum_{j \in J} \mu(B_j) \sum_{i \in I} \mu(A_i|B_j) \log \mu(A_i|B_j)$$

gde je *uslovna mera* definisana kao:

$$\mu(A|B) = \frac{\mu(A \cap B)}{\mu(B)}.$$

Ako je $\beta = \{X\}$, onda je $H(\alpha|\beta) = H(\alpha)$. Entropiju $H(\alpha|\beta)$ možemo zapisati u obliku $H(\alpha|\beta) = \int_X I_{\alpha|\beta} d\mu$, gde je $I_{\alpha|\beta}$ *uslovna informacija* koju definišemo kao $I_{\alpha|\beta}(x) = -\log \mu(A(x)|B(x))$, a $A(x)$ i $B(x)$ redom elementi particija α i β koji sadrže x.

Za particije α i β kažemo da su *jednake*, u oznaci $\alpha = \beta \pmod 0$, ako za proizvoljni element strogo pozitivne mere $C \in \alpha$, postoji element $D \in \beta$ takav da je $\mu(C \triangle D) = \mu((C \setminus D) \cup (D \setminus C)) = 0$. Za particiju β kažemo da je *finija* od particije α, u oznaci $\alpha \leq \beta$, ako za svaki $D \in \beta$ postoji $C \in \alpha$ takav da je $D \subset C$. Za dve merljive particije α i β kažemo da su *nezavisne*, u oznaci $\alpha \perp \beta$, ako je $\mu(A \cap B) = \mu(A)\mu(B)$, za sve $A \in \alpha$ i $B \in \beta$.

Teorema. *Neka je* $(X, \mathcal{B}, \mu)$ *prostor verovatnoća,* $\alpha = \{A_i | i \in I\}$, $\beta = \{B_j | j \in J\}$ *i* $\gamma = \{C_k | k \in K\}$ *najviše prebrojive particije od* X *i* $\delta = \{X\}$. *Tada:*

(1) $0 < -\log(\sup_{i \in I} \mu(A_i)) \leq H(\alpha)) \leq \log(card(\alpha))$; *ukoliko je particija* α *konačna, onda jednakost* $H(\alpha) = \log(card(\alpha))$ *važi ako i samo ako svi elementi particije* α *imaju međusobno jednaku meru.*

(2) $0 \leq H(\alpha|\beta) \leq H(\alpha)$; $H(\alpha|\beta) = H(\alpha)$ *važi ako i samo ako su particije* α *i* β *nezavisne;* $H(\alpha|\beta) = 0$ *ako i samo ako je* $\alpha \leq \beta$; *ako je* $\gamma \geq \beta$, *onda je* $H(\alpha|\gamma) \leq H(\alpha|\beta)$;

(3) $H(\alpha \vee \beta|\gamma) = H(\alpha|\gamma) + H(\beta|\alpha \vee \gamma)$; *u slučaju* $\gamma = \delta$ *imamo* $H(\alpha \vee \beta) = H(\alpha) + H(\beta|\alpha)$;

(4) $H(\alpha \vee \beta|\gamma) \leq H(\alpha|\gamma) + H(\beta|\gamma)$ *i* $H(\alpha \vee \beta) \leq H(\alpha) + H(\beta)$;

(5) $H(\alpha|\beta) + H(\beta|\gamma) \geq H(\alpha|\gamma)$;

(6) Ako je λ druga mera na prostoru X, onda za svaku particiju α koja je merljiva u odnosu na obe mere, μ i λ, i za svaki $p \in [0,1]$ važi $pH_\mu(\alpha)+(1-p)H_\lambda(\alpha) \leq H_{p\mu+(1-p)\lambda}(\alpha)$.

Posledica. *Neka su α i β merljive particije sa konačnim entropijama i*

$$d_R(\alpha,\beta) = H(\alpha|\beta) + H(\beta|\alpha).$$

Tada je d_R metrika na skupu merljivih particija sa konačnom entropijom.

Metrika d_R se naziva Rockhlin–ova metrika.

Dokaz posledice. Iz (2) sledi $d_R(\alpha,\beta) \geq 0$. Ako je $d_R(\alpha,\beta) = 0$, onda je $H(\alpha|\beta) = H(\beta|\alpha) = 0$, što po (2) znači $\alpha \leq \beta$ i $\beta \leq \alpha$, odnosno $\alpha = \beta$ skoro svuda. Simetričnost d_R proizilazi direktno iz definicije. Konačno, iz (5) imamo

$$\begin{aligned} d_R(\alpha,\beta) = H(\alpha|\beta) + H(\beta|\alpha) &\leq H(\alpha|\gamma) + H(\gamma|\beta) + H(\beta|\gamma) + H(\gamma|\alpha) \\ &= d_R(\alpha,\gamma) + d_R(\gamma,\beta). \square \end{aligned}$$

Dokaz teoreme. (1) Iz definicije entropije imamo da je $H(\alpha) \geq 0$. Ukoliko particija α sadrži bar dva skupa strogo pozitivne mere, onda je $H(\alpha) > 0$. Dakle, ako je $H(\alpha) = 0$, onda je $\alpha = \delta$ skoro svuda. Kako je $-\log(\sup_{i\in I}\mu(A_i)) = \inf(I(\alpha))$, imamo $H(\alpha) \geq -\log(\sup_{i\in I}\mu(A_i))$.

Da bismo dokazali da je $H(\alpha)) \leq \log(card(\alpha))$, pretpostavićemo da je α konačna particija. Razmotrimo funkciju

$$\phi(x) = \begin{cases} x\log x, & x \geq 0 \\ 0, & x = 0 \end{cases}$$

Kako je funkcija $\phi(x)$ strogo konveksna, jer je $\phi''(x) = \frac{1}{x} > 0$, važi $\phi(\sum_{i=1}^{\infty} a_i x_i) \leq \sum_{i=1}^{\infty} a_i\phi(x_i)$, za nenegativne a_i, $i = 1,2,\ldots$ za koje je $\sum_{i=1}^{\infty} a_i = 1$. Zatim, neka je $\alpha = \{A_1, A_2, \ldots, A_k\}$, $a_i = \frac{1}{k}$ i $x_i = \mu(A_i)$ za $i = 1,2,\ldots,k$. Iz konveksnosti funkcije ϕ sledi

$$-\frac{1}{k}\log k = \phi(\frac{1}{k}) = \phi\left(\frac{1}{k}\sum_{i=1}^{k}\mu(A_i)\right) \leq \sum_{i=1}^{k}\frac{1}{k}\phi(\mu(A_i)) = -\frac{1}{k}H(\alpha),$$

odkle sledi

$$H(\alpha) \leq \log k.$$

Drugi deo tvrđenja sledi iz stroge konveksnosti funkcije ϕ.

(2) Ova nejednakost takođe sledi iz konveksnosti funkcije ϕ:

$$\begin{aligned} 0 \leq H(\alpha|\beta) &= -\sum_{j\in J}\mu(B_j)\sum_{i\in I}\phi(\mu(A_i|B_j)) \\ &= -\sum_{i\in I}\sum_{j\in J}\mu(B_j)\phi(\mu(A_i|B_j)) \\ &\leq -\sum_{i\in I}\phi\left(\sum_{j\in J}\mu(B_j)\mu(A_i|B_j)\right) \\ &= -\sum_{i\in I}\phi(\mu(A_i)) \\ &= H(\alpha). \end{aligned}$$

Kako je $\phi(x) < 0$, za $x \in (0,1)$, imamo da ako je $H(\alpha|\beta) = 0$, onda za svaki $j \in J$ za koji je $\mu(B_j) > 0$ važi $\phi(\mu(A_i|B_j)) = 0$ za sve $i \in I$, pa i $\alpha \leq \beta$. Ako je $H(\alpha|\beta) = H(\alpha)$, onda u prethodnom izrazu svuda važi jednakost, odnosno imamo

$$\phi(\mu(A_i)) = \phi\left(\sum_{j\in J,\mu(B_j)>0} \mu(B_j)\mu(A_i|B_j)\right) = \sum_{j\in J,\mu(B_j)>0} \mu(B_j)\phi(\mu(A_i|B_j)).$$

Iz stroge konveksnosti funkcije ϕ sledi da ako je $\mu(A_i) > 0$ i $\mu(B_j) > 0$, onda je $\mu(A_i|B_j) = \mu(A_i)$, što znači da je $\mu(A_i \cap B_j) = \mu(A_i)\mu(B_j)$.

(3) Važi:

$$\begin{aligned}
H(\alpha \vee \beta|\gamma) &= -\sum_{(i,j,k)\in I\times J\times K} \mu(A_i \cap B_j \cap C_k) \log \frac{\mu(A_i \cap B_j \cap C_k)}{\mu(C_k)} \\
&= -\sum_{(i,j,k)} \mu(A_i \cap B_j \cap C_k) \log \frac{\mu(A_i \cap C_k)}{\mu(C_k)} \\
&\quad - \sum_{(i,j,k)} \mu(A_i \cap B_j \cap C_k) \log \frac{\mu(B_j \cap C_k)}{\mu(C_k)} \\
&= -\sum_{(i,j,k)} \mu(A_i \cap B_j \cap C_k) \log \frac{\mu(A_i \cap C_k)}{\mu(C_k)} + H(\beta|\alpha \vee \gamma) \\
&= H(\alpha|\gamma) + H(\beta|\alpha \vee \gamma)
\end{aligned}$$

(4) Ova nejednakost sledi iz (3) i nejednakosti $H(\beta|\alpha \vee \gamma) \leq H(\beta|\gamma)$, koja sledi iz (2) s obzirom da je $\alpha \vee \gamma \geq \gamma$.

(5) Iz (3) i (4) imamo $H(\gamma|\alpha \vee \beta) = H(\alpha \vee \gamma|\beta) - H(\alpha|\beta) \leq H(\gamma|\beta)$. Korišćenjem (3) nekoliko puta dobijamo:

$$\begin{aligned}
H(\alpha|\beta) + H(\beta|\gamma) &= H(\alpha \vee \beta) + H(\beta \vee \gamma) - H(\beta) - H(\gamma) \\
&= H(\alpha \vee \beta) + H(\gamma|\beta) - H(\gamma) \\
&= H(\alpha \vee \beta \vee \gamma) - H(\gamma|\alpha \vee \beta) + H(\gamma|\beta) - H(\gamma) \\
&\geq H(\alpha \vee \beta \vee \gamma) - H(\gamma) \\
&\geq H(\alpha \vee \gamma) - H(\gamma) \\
&= H(\alpha|\gamma).
\end{aligned}$$

(6) Ovaj deo tvrđenja sledi direktno iz konveksnosti funkcije ϕ:

$$\begin{aligned}
pH_\mu(\alpha) + (1-p)H_\lambda(\alpha) &= -p\sum_{i\in I} \phi(\mu(A_i)) - (1-p)\sum_{i\in I} \phi(\lambda(A_i)) \\
&\leq -\sum_{i\in I} \phi((p\mu + (1-p)\lambda)(A_i)) = H_{p\mu+(1-p)\lambda}(\alpha). \square
\end{aligned}$$

Bez dokaza, navodimo još neke osobine entropije.

Teorema. *Neka je $T : (X, \mu) \to (X, \mu)$ transformacija koja čuva meru na prostoru verovatnoća (X, μ) i neka su α i β merljive particije sa konačnim entropijama. Tada važi;*
(1) $0 \le \limsup_{n\to\infty}(-\frac{1}{n}\log(\sup_{C\in\alpha\vee\cdots\vee T^{-n+1}(\alpha)}\mu(C)) \le h(\alpha, T) \le H(\alpha)$;
(2) $h(\alpha \vee \beta, T) \le h(\alpha, T) + h(\beta, T)$;
(3) $h(\beta, T) \le h(\alpha, T) + H(\beta|\alpha)$; *ukoliko je* $\alpha \le \beta$, *onda je* $h(\alpha, T) \le h(\beta, T)$;
(4) $|h(\alpha, T) - h(\beta, T)| \le H(\alpha|\beta) + H(\beta|\alpha)$ *(Rokhlin–ova nejednakost);*
(5) $h(T^{-1}(\alpha), T) = h(\alpha, T)$;
(6) $h(\alpha, T) = h(\vee_{i=0}^{k} T^{-i}(\alpha), T)$ *za* $k \in N$;
(7) Ako je ν *druga mera i* $p \in [0, 1]$, *onda važi*

$$ph_\mu(\alpha, T) + (1-p)h_\nu(\alpha, T) \le h_{p\mu+(1-p)\nu}(\alpha, T).$$

§3.3. Entropija u teoriji dinamičkih sistema

Sada ćemo pojmove entropije i uslovne entropije razmotriti u jednom savremenom kontekstu ergodičke teorije.

Neka je $(X, \mathcal{B}, \mu)$ prostor verovatnoća, $T : X \to X$ transformacija koja čuva meru, $\alpha = \{A_1, \ldots, A_n\}$ particija skupa X. Entropija particije α, koju smo definisali u prethodnom delu, jeste

$$H(\alpha) = -\sum_{i=1}^{n} \mu(A_i)\log\mu(A_i).$$

Sada navodimo definiciju entropije h transformacije T u odnosu na α, u oznaci $h(\alpha, T)$, takođe iz prethodnog dela:

$$h(\alpha, T) = h_\mu(\alpha, T) = \lim_{n\to\infty}\frac{1}{n}H(\alpha \vee T^{-1}\alpha \vee \cdots \vee T^{-n+1}\alpha)$$

Da je ova definicija korektna, tj. da posmatrana granična vrednost postoji, zaključujemo iz sledećih činjenica:

Entropija je *subaditivna:* $H(\alpha \vee \beta) \le H(\alpha) + H(\beta)$, što sledi iz sledećeg razmatranja:

$$\begin{aligned}
H(\alpha \vee \beta) &= -\sum_{i,j}\mu(A_i \cap B_j)\log\mu(A_i \cap B_j) \\
&= -\sum_j\sum_i \mu(A_i)\frac{\mu(A_i \cap B_j)}{\mu(A_i)}\log\frac{\mu(A_i \cap B_j)}{\mu(A_i)} - \sum_i \mu(A_i)\log\mu(A_i) \\
&\le -\sum_j(\sum_i \mu(A_i \cap B_j))\log(\sum_i \mu(A_i \cap B_j)) + H(\alpha) \\
&\le H(\alpha) + H(\beta)
\end{aligned}$$

gde smo, u prethodna dva koraka ovog izvođenja koristili Jensenovu nejednakost i formulu potpune verovatnoće.

Ako transformacija T čuva meru, onda je entropija *stacionarna:* $H(T^{-1}\alpha) = H(\alpha)$.

Iz subaditivnosti i stacionarnosti imamo:

$$H(\alpha \vee T^{-1}\alpha \vee \cdots \vee T^{-(n+n')+1}\alpha) \leq H(\alpha \vee T^{-1}\alpha \vee \cdots \vee T^{-n+1}\alpha) + H(\alpha \vee T^{-1}\alpha \vee \cdots \vee T^{-n'+1}\alpha)$$

Ako uvedemo sledeću oznaku: $\eta_n = H(\alpha \vee T^{-1}\alpha \vee \cdots \vee T^{-n+1}\alpha)$, imaćemo: $\eta_n \leq \eta_{n+1}$, jer finijoj particiji odgovara veća entropija, pa ćemo iz gornje nejednakosti dobiti i:

$$\eta_{n+n'} \leq \eta_n + \eta_{n'}$$

što, konačno, osigurava egzistenciju posmatrane granične vrednosti.

Neka su $\alpha = \{A_1, \ldots, A_n\}$ i $\beta = \{B_1, \ldots, B_m\}$ merljive particije skupa X. Uslovnu entropiju $H(\alpha|\beta)$ definišemo kao u prethodnom delu:

$$H(\alpha|\beta) = -\sum_{j=1}^{m} \mu(B_j) \sum_{i=1}^{n} \mu(A_i|B_j) \log \mu(A_i|B_j) = \sum_{i,j} \mu(A_i \cap B_j) \log \mu(A_i|B_j)$$

Kao posledicu već dokazanog tvrđenja $H(\alpha \vee \beta) = H(\alpha) + H(\beta|\alpha) = H(\beta) + H(\alpha|\beta)$ dobijamo:

$$\begin{aligned} H(\alpha \vee T^{-1}\alpha \vee \cdots \vee T^{-n+1}\alpha) - H(\alpha \vee T^{-1}\alpha \vee \cdots \vee T^{-n+2}\alpha) = & \\ &= H(T^{-n+1}\alpha|\alpha \vee T^{-1}\alpha \vee \cdots \vee T^{-n+2}\alpha) \\ &\leq H(T^{-n+1}\alpha|T^{-1}\alpha \vee \cdots \vee T^{-n+2}\alpha) \\ &= H(T^{-n+2}\alpha|\alpha \vee T^{-1}\alpha \vee \cdots \vee T^{-n+3}\alpha) \\ &= H(\alpha T^{-1}\alpha \vee \cdots \vee T^{-n+2}\alpha) \\ &- H(\alpha \vee T^{-1}\alpha \vee \cdots \vee T^{-n+3}\alpha) \end{aligned}$$

što znači:

$$\eta_n - \eta_{n-1} \leq \eta_{n-1} - \eta_{n-2}$$

za svaki $n \in \mathbf{N}$. Odavde izvodimo:

$$\lim_{n\to\infty}(\eta_n - \eta_{n-1}) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} (\eta_i - \eta_{i-1}) = \lim_{n\to\infty} \frac{1}{n}(\eta_n - \eta_1) = \lim_{n\to\infty} \frac{\eta_n}{n}$$

odnosno i sledeće tvrđenje:

Teorema. *Za svaku merljivu particiju $\alpha = \{A_1, \ldots, A_n\}$ skupa X i svaku transformaciju T koja čuva meru važi:*

$$h(\alpha, T) = \lim_{n\to\infty}(H(\alpha \vee T^{-1}\alpha \vee \cdots \vee T^{-n+1}\alpha) - H(\alpha \vee T^{-1}\alpha \vee \cdots \vee T^{-n+2}\alpha))$$

Kako entropiju $H(A)$ možemo posmatrati i kao količinu informacija sadržanih u slučajnom događaju A, što sledi prema prethodnom tvrđenju, postavlja se pitanje relevantnosti i korisnosti tih informacija u eventualnom predviđanju. Naime, polazeći od entropije nekog dinamičkog sistema, koja nam je poznta do datog trenutka, možemo pokušati da saznamo nešto o ponašanju tog sistema u bliskoj ili daljoj budućnosti[25]. Neka je, po definiciji, $b_k = i$, ukoliko $T^{-k}(x) \in A_i$. Posmatraćemo vremensku seriju $b_0, b_1, \dots$ i istraživati neodređenost budućih vrednosti članova serije na osnovu poznavanja prethodnih. Neka je $\Lambda(n)$ skup svih nizova oblika $(b_0, b_1, \dots, b_{n-1})$, za $n \in \mathbf{N}$. Kada kardinalnost particije α iznosi m, moguće je postojanje ukupno m^n takvih nizova. Neka je $p(\lambda)$ verovatnoća realizacije događaja $\lambda \in \Lambda(n)$. Pretpostavka o stacionarnosti transformacije T imaće za posledicu da se ove verovatnoće ne menjaju kada se vremenski indeks pomeri za vrednost k, tj. pri posmatranju niza $(b_k, b_{k+1}, \dots, b_{k+n-1})$, odakle, za $n = 1$ dobijamo moguće vrednosti $i = b_k$ koje se pojavljuju sa (bezuslovnim) verovatnoćama $\mu(A_i)$, što bi značilo da se one mogu menjati nakon saznanja prethodne (ili prethodnih) vrednosti b_{k-1}. Ovo se može iskoristiti u razmatranju pitanja kako entropija zavisi od dužine n posmatranog niza. Kako je, dakle,

$$H(T,n) = -\sum_{\lambda=1}^{m^n} p(\lambda) \log p(\lambda) = H(\alpha \vee T^{-1}\alpha \vee \cdots \vee T^{-n+1}\alpha)$$

to je ova informacija sadržana u nizovima $(b_k, b_{k+1}, \dots, b_{k+n-1})$ dužine n. Kada je poznat takav niz, onda je za predviđanje vrednosti b_{k+n} neophodno da znamo vrednost

$$h(T,n) = H(T,n+1) - H(T,n)$$

gde je jasno da uvedena razlika $h(T,n)$ ne bi trebalo da raste sa porastom broja n, već, naprotiv, jer je logično da uz veći broj prethodnih vrednosti imamo pouzdanije predviđanje. Naime, izraz

$$h(T) = \lim_{n \to \infty} h(T,n)$$

treba shvatiti kao entropiju transformacije T u odnosu na particiju α. Ipak, postoje i određena ograničenja pri korišćenju entropje kao mere nepredvidljivosti, o čemu ovde neće biti govora.

Razlika definisanih razlika

$$\delta h(T,n) = h(T,n-1) - h(T,n)$$

determiniše srednju vrednost po kojoj neodređenost niza b_{k+n} opada sa poznavanjem još jedne nove vrednosti b_k iz prošlosti. U ovom kontekstu je moguće definisati i tzv. *efektivnu meru kompleksnosti transformacije* T (v. P. Grassberger (1986)) kao:

$$h_{EMC}(T) = \sum_{n=0}^{\infty} (h(T,n) - h(T))$$

[25]Ovim problem se bavio Grassberger (v. P. Grassberger (1986)) i mi ovde, delom, sledimo njegov pristup.

gde dati red ne mora da konvergira, ali u slučajevima kada isti divergira, ta divergencija mora biti sublinearna, tj. za $H(T,n) = nH(T) + \sigma(n)$, mora biti:

$$\lim_{n\to\infty} \frac{\sigma(n)}{n} = 0$$

Sa druge strane, ukoliko niz $\sigma(n)$ konvergira nekoj vrednosti σ, kada $n \to \infty$, kako je $h(Tn) = h + \sigma(n+1) - \sigma(n)$, imamo

$$h_{EMC}(T) = \sum_{n=0}^{\infty} (\sigma(n+1) - \sigma(n)) = \sigma - \sigma(1)$$

Ovde se *informacija* može definisati kao

$$I(T, n+n') = H(T,n) + H(T,n') - H(T,n+n') \geq 0$$

što izražava činjenicu da je entropija niza dužine $n + n'$ manja od zbira entropija nizova dužina n i n'. Uz pretpostavku o stacionarnosti imaćemo:

$$I(T, n+n') = \sigma(n) + \sigma(n') - \sigma(n+n')$$

odnosno

$$\lim_{n,n'\to\infty} I(T, n+n') = \sigma$$

u slučaju kada posmatrana granična vrednost postoji. Tako, vrednost σ odgovara i slučaju kada se radi o informaciji vezanoj za nizove beskonačne dužine. Napomenimo da niz $\sigma(n)$ ne mora biti konvergentan.

Neka je $\mathcal{F} \subset \mathcal{B}$ pod$-\sigma-$algebra familije merljivih podskupova skupa X. *Uslovno očekivanje* funkcije $f \in L^1(X)$ u odnosu na $\mathcal{F}$, u oznaci $E(f|\mathcal{F})$, jeste $\mathcal{F}-$merljiva funkcija na X za koju važi $\int_A E(f|\mathcal{F})d\mu = \int_A f d\mu$, za sve $A \in \mathcal{F}$. $E(f|\mathcal{F})(x)$ predstavlja očekivanu vrednost f, ukoliko nam je poznato koji skupovi iz $\mathcal{F}$ sadrže x. *Uslovna verovatnoća skupa* $A \in \mathcal{B}$ u odnosu na $\mathcal{F}$ je $\mu(A|\mathcal{F}) = E(\chi_A|\mathcal{F})$. *Uslovna informacija* se definiše na sledeći način:

$$I_{\alpha|\mathcal{F}}(x) = -\sum_{A\in\alpha} \log \mu(A|\mathcal{F})(x)\chi_A(x),$$

dok se *uslovna entropija particije* α u odnosu na $\mathcal{F}$ definiše kao

$$H(\alpha|\mathcal{F}) = \int_X I_{\alpha|\mathcal{F}}(x) d\mu(x).$$

Navešćemo neke osnovne osobine uslovne entropije u odnosu na σ–algebru kao i uslovne informacije u okviru sledećih teorema.

Teorema. *Važi:*
(1) $I_{\alpha|\mathcal{F}} \geq 0$ *skoro svuda, pa je i* $H(\alpha|\mathcal{F}) \geq 0$;
(2) $I_{\alpha|\{\emptyset,X\}} = I_\alpha$ *skoro svuda i* $H(\alpha|\{\emptyset, X\}) = H(\alpha)$;

(3) $I_{\alpha|\mathcal{B}} = 0$ *skoro svuda i* $H(\alpha|\mathcal{B}) = 0$.

Alternativne definicije uslovne informacije $I_{\alpha|\mathcal{F}}(x)$ su $-\sum_{A\in\alpha} \log \mu(A|\mathcal{F})\mu(A|\mathcal{F})$ i $E(I_\alpha|\mathcal{F})$. Očekivanja ovako definisanih uslovnih informacija su:

$$E\left(-\sum_{A\in\alpha} \log \mu(A|\mathcal{F})\mu(A|\mathcal{F})\right) = H(\alpha|\mathcal{F})$$

$$E(E(I_\alpha|\mathcal{F})) = E(I_\alpha) = H(\alpha).$$

Prva jednakost sledi iz činjenice da važi $E(\log \mu(A|\mathcal{F})\chi_A|\mathcal{F}) = \log \mu(A|\mathcal{F})\mu(A|\mathcal{F})$ za svaki $A \in \alpha$, pa imamo

$$H(\alpha|\mathcal{F}) = E(I_{\alpha|\mathcal{F}}) = E(E(I_{\alpha|\mathcal{F}}|\mathcal{F})) = E(-\sum_{A\in\alpha} \log \mu(A|\mathcal{F})\mu(A|\mathcal{F})).$$

Navodimo još neke osobine uslovne entropije i uslovne informacije.

Teorema. *Ako su* α *i* β *prebrojive merljive particije skupa* X *i* $\mathcal{F}$ *pod*$-\sigma-$*algebra od* $\mathcal{B}$*, onda skoro svuda važi* $I_{\alpha\vee\beta|\mathcal{F}} = I_{\alpha|\mathcal{F}} + I_{\beta|\mathcal{B}(\alpha)\vee\mathcal{F}}$.

Dokaz. Za svaki $E \in \beta$, skoro svuda važi $\mu(E|\mathcal{B}(\alpha) \vee \mathcal{F}) = \sum_{A\in\alpha} \frac{\mu(E\cap A|\mathcal{F})}{\mu(A|\mathcal{F})}\chi_A$, jer za $P \cap Q \in \mathcal{B}(\alpha) \vee \mathcal{F}$ ($P \in \alpha$ i $Q \in \mathcal{F}$) imamo:

$$\begin{aligned}
\int_{P\cap Q} \left(\sum_{A\in\alpha} \frac{\mu(E\cap A|\mathcal{F})}{\mu(A|F)}\chi_A\right) d\mu &= \int_Q \chi_P \frac{\mu(E\cap P|\mathcal{F})}{\mu(P|\mathcal{F})} d\mu \\
&= \int_Q E\left(\chi_P \frac{\mu(E\cap P|\mathcal{F})}{\mu(P|\mathcal{F})}|\mathcal{F}\right) \\
&= \int_Q E(\chi_P|\mathcal{F}) \frac{\mu(E\cap P|\mathcal{F})}{\mu(P|\mathcal{F})} d\mu \\
&= \int_Q \mu(E\cap P|\mathcal{F}) d\mu \\
&= \mu(E \cap P \cap Q) \\
&= \int_{P\cap Q} \chi_E d\mu
\end{aligned}$$

Dakle, imamo da je $\sum_{A\in\alpha} \left(\frac{\log \mu(E\cap A|\mathcal{F})}{\mu(A|\mathcal{F})}\right)\chi_A = \log \mu(E|\mathcal{B}(\alpha) \vee \mathcal{F})$ skoro svuda i

$$\begin{aligned}
I_{\alpha\vee\beta|\mathcal{F}} &= -\sum_{A\in\alpha, B\in\beta} \log \mu(A\cap E|\mathcal{F})\chi_{A\cap E} \\
&= -\sum_{A\in\alpha, E\in\beta} \log \mu(A|\mathcal{F})\chi_A\chi_E - \sum_{E\in\beta} \log \mu(E|\mathcal{B}(\alpha) \vee \mathcal{F})\chi_E \\
&= -\sum_{A\in\alpha} \log \mu(A|\mathcal{F})\chi_A - \sum_{E\in\beta} \log \mu(E|\mathcal{B}(\alpha) \vee \mathcal{F})\chi_E \\
&= I_{\alpha|F} + I_{\beta|\mathcal{B}(\alpha)\vee\mathcal{F}}. \square
\end{aligned}$$

U slučaju $\mathcal{F} = \{\emptyset, X\}$, iz prethodne teoreme kao direktnu posledicu izvodimo sledće tvrđenje:

Posledica. Za sve prebrojive particije α i β skoro svuda važi $I_{\alpha\vee\beta} = I_\alpha + I_{\beta|\alpha}$.

Teorema. *Neka su α i β prebrojive merljive particije skupa X i $\mathcal{F}$, $\mathcal{F}_1$ i $\mathcal{F}_2$ pod–σ–algebre od $\mathcal{B}$. Tada važi:*

(1) $H(\alpha\vee\beta|\mathcal{F}) = H(\alpha|\mathcal{F}) + H(\beta|\mathcal{B}(\alpha)\vee\mathcal{F})$.
(2) Ako je $\mathcal{F}_1 \supset \mathcal{F}_2$, onda je $H(\alpha|\mathcal{F}_1) \leq H(\beta|\mathcal{F}_2)$.
(3) $H(T^{-1}\alpha|T^{-1}\mathcal{F}) = H(\alpha|\mathcal{F})$
(4) Ako je $\alpha \leq \beta$, onda je $H(\alpha|\mathcal{F}) \leq H(\beta|\mathcal{F})$.
(5) $H(\alpha\vee\beta|\mathcal{F}) \leq H(\alpha|\mathcal{F}) + H(\beta|\mathcal{F})$.

Dokaz.
(1) Integracijom već dokazane jednakosti $I_{\alpha\vee\beta|\mathcal{F}} = I_{\alpha|\mathcal{F}} + I_{\beta|\mathcal{B}(\beta)\vee\mathcal{F}}$, dobija se tražena jednakost.

(2) Primenom Jensenove nejednakosti na funkciju $f(x) = -x\log x$, za svaki $A \in \alpha$ važi: $E(f\circ\mu(A|\mathcal{F}_1)|\mathcal{F}_2) \leq f\circ E(\mu(A|\mathcal{F}_1)|\mathcal{F}_2) = f\circ\mu(A|\mathcal{F}_2)$. Integracijom prethodne nejednakosti dobijamo $\int_X f\circ\mu(A|\mathcal{F}_1)d\mu \leq \int_X f\circ\mu(A|\mathcal{F}_2)$. Tražena nejednakost se dobija iz prethodne sumiranjem po $A \in \alpha$.

(3) Lako se proverava da važi $\mu(T^{-1}A|T^{-1}\mathcal{F})(x) = \mu(A|\mathcal{F})(Tx)$ skoro svuda, pa imamo

$$\begin{aligned}
H(T^{-1}\alpha|T^{-1}\mathcal{F}) &= -\int\sum_{A\in\alpha}\log\mu(T^{-1}A|T^{-1}\mathcal{F})(x)\chi_{T^{-1}A}(x)d\mu(x)\\
&= -\int\sum_{A\in\alpha}\log\mu(A|\mathcal{F})(Tx)\chi_A(Tx)d\mu(x)\\
&= -\int\sum_{A\in\alpha}\log\mu(A|\mathcal{F})(x)\chi_A(x)d\mu(x)\\
&= H(\alpha|\mathcal{F}).
\end{aligned}$$

(4) Kako je $H(\beta|B(\alpha)\vee\mathcal{F}) \geq 0$ i $\alpha\vee\beta = \beta$ za $\alpha \leq \beta$, iz (1) imamo

$$H(\beta|F) = H(\alpha\vee\beta|\mathcal{F}) = H(\alpha|\mathcal{F}) + H(\beta|\mathcal{B}(\alpha)\vee\mathcal{F}) \geq H(\alpha|\mathcal{F}).$$

(5) Ovo tvrđenje sledi iz (1), (2) i činjenice $\mathcal{F} \subset \mathcal{B}(\alpha)\vee\mathcal{F}$. $\square$

Specijalan slučaj tvrđenja (2) prethodne teoreme, za $\mathcal{F} = \{\emptyset, X\}$, dobijamo $H(\alpha|\mathcal{F}) \leq H(\alpha)$.

Teorema. *Neka su α i β prebrojive merljive particije sa konačnim entropijama. Tada su sledeća tri tvrđenja međusobno ekvivalentna:*
(1) $\alpha \perp \beta$
(2) $H(\alpha\vee\beta) = H(\alpha) + H(\beta)$
(3) $H(\alpha|\beta) = H(\alpha)$
Dokaz. Ekvivalentnost uslova (2) i (3) sledi direktno iz prethodne teoreme, deo (1).

Dokazujemo da iz (1) sledi (3). Neka je $\alpha \perp \beta$. Tada za sve $A \in \alpha$ važi $\mu(A|\mathcal{B}(\beta)) = \mu(A)$ skoro svuda. Dakle $I_{\alpha|\beta} = I_\alpha$ skoro svuda, pa je i $H(\alpha|\beta) = H(\alpha)$.

Dokazujemo da iz (3) sledi (1). Neka je $H(\alpha|\beta) = H(\alpha)$. Tada važi $\sum_{A\in\alpha}(f(\mu(A)) - \int_X f(\mu(A|\mathcal{B}(\beta)))d\mu) = 0$, gde je $f(x) = -x\log x$. Iz dokaza prethodne teoreme, deo (2), zaključujemo da su svi članovi ove sume nenegativni, što znači da je, za sve $A \in \alpha$,

$$f(\mu(A)) = \int_X f(\mu(A|\mathcal{B}(\beta)))d\mu$$

odnosno

$$-\mu(A)\log\mu(A) = \sum_{B\in\beta} -\frac{\mu(A\cap B)}{\mu(B)}\log\frac{\mu(A\cap B)}{\mu(B)}\mu(B).$$

Pretpostavimo da je $\beta = \{B_1, B_2, \dots\}$, i, za sve $k = 1, 2, \dots$, $\lambda_k = \mu(B_k)$ i $\xi_k = \mu(A|B_k) = \frac{\mu(A\cap B_k)}{\mu(B_k)}$. Tada imamo da je $(\mu(A), f(\mu(A))) = \sum \lambda_k(\xi_k, f(\xi_k))$. Površina ispod grafika funkcije $f(x)$ je konveksna i tačka sa ruba te površi je predstavljena kao konveksna kombinacija tačaka koje su takođe sa ruba. Međutim, to je moguće jedino ako su svi ξ_k međusobno jednaki, što znači da je $\frac{\mu(A\cap B_k)}{\mu(B_k)} = \mu(A)$, za sve k, odnosno $\alpha \perp \beta$.□

Entropiju $H(\alpha)$ particije α, dakle, možemo posmatrati kao srednju vrednost informacije o činjenici kom skupu particije pripada posmatrana tačka. Stoga, za svaku tačku $x \in X$, prirodno je definisati *informaciju o x u odnosu na particiju α* kao:

$$I_\alpha(x) = -\log\mu(A_i(x))$$

gde je $A_i(x)$ skup particije α koji sadrži tačku x.

Iz načina na koji smo definisali entropiju $H(\alpha)$ particije α, pojavljuje se i sledeća mogućnost posmatranja entropije $H(\alpha)$ particije α:

$$H(\alpha) = -\int \log\mu(A_i(x))d\mu(x) = \int I_\alpha(x)d\mu(x)$$

pa bi i odgovarajuća transformacija T zadovoljavala uslov:

$$H(\alpha \vee T^{-1}\alpha \vee \cdots \vee T^{-n+1}\alpha) = \int I_{\alpha\vee T^{-1}\alpha\vee\cdots\vee T^{-n+1}\alpha}(x)d\mu(x)$$

kada bi entropija predstavljala sledeću graničnu vrednost:

$$h(\alpha, T) = h_\mu(\alpha, T) = \lim_{n\to\infty}\frac{H(\alpha \vee T^{-1}\alpha \vee \cdots \vee T^{-n+1}\alpha)}{n}$$

koja se, kada je u pitanju ergodična transformacija T, može dobiti i posredstvom sledećeg tvrđenja:

Shannon–McMillan–Breiman–ova ergodička teorema. (C. E. Shannon; B. McMillan; L. Breiman) *Neka ergodična transformacija $T : X \to X$ čuva meru verovatnosnog prostora $(X, \mathcal{B}, \mu)$ i neka je α najviše prebrojiva particija sa konačnom entropijom $H(\alpha)$. Tada:*

$$\lim_{n\to\infty}\frac{1}{n}I_{\alpha\vee\cdots\vee T^{-n+1}\alpha}(x) = h(\alpha, T)$$

za skoro svaki x u odnosu na meru μ.

Dokaz ove teoreme se može naći u K. Petersen (1983).

Oslanjajući se na Kolmogorovljev rad (A. N. Kolmogorov (1958)) Sinai je definisao novi koncept entropije (Y. Sinai (1959)) koji danas igra centralno mesto u teoriji dinamičkih sistema (R. Frigg (2004, 2006)).

Neka je data particija $\alpha = \{A_1, \dots, A_k\}$ skupa X, $H(\alpha) = -\sum_{i=1}^{k} \mu(A_i)\log(\mu(A_i))$ i $(X, \mathcal{B}, \mu, T)$ diskretan dinamički sistem. Granična vrednost

$$H(\alpha, T) = \lim_{n\to\infty} \frac{H(\alpha \vee T^{-1}\alpha \vee \cdots \vee T^{-n+1}\alpha)}{n}$$

postoji na osnovu Shannon–McMillan–Breiman–ove ergodičke teoreme i *Kolmogorov–Sinai–evu entropiju* definišemo kao

$$h(T) = \sup_{\alpha} H(\alpha, T),$$

pri čemu je supremum po svim konačnim particijama α.

Efikasan i dobar metod za izračunavanje entropije može se bazirati i na sledećem tvrđenju:

Kolmogorov–Sinai–eva teorema. (A. N. Kolmogorov; Y. G. Sinai) *Neka je $T : X \to X$ invertibilna transformacija koja čuva meru verovatnosnog prostora $(X, \mathcal{B}, \mu)$ i neka je $\mathcal{A}$ konačna podalgebra od $\mathcal{B}$ generisana nekom particijom α takva da je*

$$\bigvee_{k=-\infty}^{+\infty} T^k\mathcal{A} = \cdots \vee T^{-n}\mathcal{A} \vee \cdots \vee T^{-1}\mathcal{A} \vee T^0\mathcal{A} \vee T^1\mathcal{A} \vee \cdots \vee T^n\mathcal{A} \vee \cdots = \mathcal{B}$$

do na skupove mere nula. Tada: $h(T) = h(\alpha, T)$.

Pre dokaza ove teoreme, razmotrićemo neka tvrđenja koja se koriste u tom dokazu.

Lema 1. *Za sve prebrojive merljive particije α i β važi $h(\beta, T) \le h(\alpha, T) + H(\beta|\alpha)$.*

Dokaz. Neka je $\beta_0^{m-1} = \beta \vee T^{-1}\beta \vee \cdots \vee T^{-m+1}\beta$ i $\alpha_0^{m-1} = \alpha \vee T^{-1}\alpha \vee \cdots \vee T^{-m+1}\alpha$ za $m = 1, 2, \dots$. Tada za svaki m važi:

$$\begin{aligned} H(\beta_0^{m-1}|\alpha_0^{m-1}) &\le H(\beta|\alpha_0^{m-1}) + H(T^{-1}\beta|\alpha_0^{m-1}) + \cdots + H(T^{-m+1}\beta|\alpha_0^{m-1}) \\ &\le H(\beta|\alpha) + H(T^{-1}\beta|T^{-1}\alpha) + \cdots + H(T^{-m+1}\beta|T^{-m+1}\alpha) \\ &= mH(\beta|\alpha) \end{aligned}$$

Takođe imamo da je

$$H(\beta_0^{m-1}) \le H(\beta_0^{m-1} \vee \alpha_0^{m-1}) = H(\alpha_0^{m-1}) + H(\beta_0^{m-1}|\alpha_0^{m-1}) \le H(\alpha_0^{m-1}) + mH(\beta|\alpha),$$

odakle sledi

$$h(\beta, T) = \lim_{m\to\infty} \frac{1}{m} H(\beta_0^{m-1}) \le \lim_{m\to\infty} \frac{1}{m} H(\alpha_0^{m-1}) + H(\beta|\alpha) = h(\alpha, T) + H(\beta|\alpha). \square$$

Lema 2. *$H(\alpha|\mathcal{F}) = 0$ ako i samo ako $\alpha \subset \mathcal{F}$ do na skupove mere nula.*

Dokaz. Ako je $\alpha \subset \mathcal{F}$, onda je $\mu(A|\mathcal{F}) = (E(\chi_A|\mathcal{F})) = \chi_A$ skoro svuda za sve $A \in \alpha$, pa je $I_{\alpha|\mathcal{F}} = 0$ skoro svuda.

Pretpostavimo da je $H(\alpha|\mathcal{F}) = 0$. Tada je $log_2\mu(A|\mathcal{F})\mu(A|\mathcal{F}) = 0$ skoro svuda, odakle sledi da je $\mu(A|\mathcal{F}) = 0$ ili $\mu(A|\mathcal{F}) = 1$ skoro svuda. Dakle $\mu(A|\mathcal{F}) = \chi_B$, za neki $B \in \mathcal{F}$. Zaklkjučujemo da je $\mu(B) = \int \mu(A|\mathcal{F})d\mu = \mu(A)$. Sa druge strane, $\mu(A \cap B) = \int_B \chi_A d\mu = \int_B \mu(A|\mathcal{F})d\mu = \int_B \chi_B d\mu = \mu(B)$. Dakle $A = B$ do na skup mere nula, pa je i $A \subset \mathcal{F}$ do na skupove mere nula.□

Lema 3. *Ako su $\mathcal{B}_1 \subset \mathcal{B}_2 \subset \ldots$ pod–$\sigma-$algebre $\sigma-$algebre $\mathcal{B}$, $\mathcal{B}_\infty = \vee_{n=1}^{\infty}\mathcal{B}_n$ i α merljiva konačna particija, onda $\lim_{n\to\infty} H(\alpha|\mathcal{B}_n) = H(\alpha|\mathcal{B}_\infty)$.*

Dokaz. Prema teoremi o konvergenciji martingala važi

$$\mu(A|\mathcal{B}_n) \to \mu(A|\mathcal{B}_\infty)$$

skoro svuda, za svaki $A \in \alpha$. Odatle sledi da

$$f(\mu(A|\mathcal{B}_n)) \to f(\mu(A|\mathcal{B}_\infty))$$

važi skoro svuda za sve $A \in \alpha$, gde je $f(x) = -x \log x$. Zatim, prema teoremi o konvergenciji ograničenih funkcija, na koju smo se pozivali, imamo da, za svaki $A \in \alpha$, važi

$$\int_X f(\mu(A|\mathcal{B}_n))d\mu \to \int_X f(\mu(A|\mathcal{B}_\infty))d\mu.$$

Sumiranjem po konačnom broju $A \in \alpha$ prethodnog izraza dobijamo $\lim_{n\to\infty} H(\alpha|\mathcal{B}_n) = H(\alpha|\mathcal{B}_\infty)$.□

Dokaz Kolmogorov–Sinai–eve teoreme. Dovoljno je dokazati da za proizvoljnu konačnu particiju β prostora X važi $h(\beta, T) \leq h(\alpha, T)$. Najpre, uvodimo oznake $\alpha_m^n = \vee_{k=m}^{n} T^{-k}\alpha$ i $\alpha_{-\infty}^{\infty} = \vee_{k=-\infty}^{\infty}\mathcal{B}(T^k\alpha)$. Iz leme 1 i činjenice $h(\vee_{k=m}^{n} T^{-k}\alpha, T) = h(\alpha, T)$ imamo

$$h(\beta, T) \leq h(\alpha_{-n}^{n}, T) + H(\beta|\alpha_{-n}^{n}) = h(\alpha, T) + H(\beta|\alpha_{-n}^{n}),$$

za $n = 0, 1, 2, \ldots$ Kako skoro svuda važi $\beta \subset \alpha_{-\infty}^{\infty}$, iz lema 2 i 3 zaključujemo da važi:

$$0 = H(\beta|\alpha_{-\infty}^{\infty}) = \lim_{n\to\infty} H(\beta|\alpha_{-n}^{n}).$$

Tražena nejednakost se dobija kada pustimo da n teži beskonačnosti u prvoj formuli.□

Napomenimo da, ako je

$$\bigvee_{k=0}^{+\infty} T^{-k}\mathcal{A} = T^0\mathcal{A} \vee T^{-1}\mathcal{A} \vee \cdots \vee T^{-n}\mathcal{A} \vee \cdots = \mathcal{B}$$

do na skupove mere nula, onda važi $h(T) = 0$.

Vrednost $h(\alpha, T)$ se može posmatrati kao mera srednje neodređenosti po jedinici vremena koju imamo kada istražujemo koji će element particije α sadržti tačku x u sledćem momentu, imajući u vidu njegovo ponašanje u prethodnom periodu. To bi, naravno, značilo da ako bismo particiju α izabrali proizvoljno, onda bi se moglo desiti da nemamo visok stepen neodređenosti, tj. odgovor na pitanje kojim ćelijama particije α pripadaju tačke $T^j x$ ne bi nam morao doneti više informacija od onih koje bismo dobili nagađanjem. Stoga je entropija transformacije T definisana kao maksimalna neodređenost po svim konačnim particijama α.

Vrednost $h(T)$ će zato predstavljati srednju vrednost neodređenosti preslikavanja tačaka iz X transformacijom T. Dakle, veličina $h(T)$ odražava 'slučajnost' transformacije T, odnosno stepen dezorganizacije prostora pod uticajem transformacije T. Očigledno, $h(T)$ jeste invarijanta izomorfizma T.

U slučaju neprekidnog dinamičkog sistema Kolmogorov–Sinai–eva entropija se definiše kao $h(T_t) = |t| h(T_1),\ t \in R$.

Particiji α možemo dati i drugačiju interpretaciju, posmatrajući je kao skup mogućih ishoda nekog eksperimenta. Vrednost $H(\alpha)$ je mera očekivane neodređenosti ishoda eksperimenta, ili količine informacija koje dobijamo izvodeći eksperiment. U ovom slučaju, particija $\alpha \vee \beta$ se može interpretirati kao složeni ekspeiment koji se sastoji iz simultane realizacije eksperimenata α i β. Tako se $\frac{1}{k} H(\alpha \vee T^{-1}\alpha \vee \cdots \vee T^{-k+1}\alpha)$ može shvatiti kao srednja količina informacija pri ponavljanju eksperimenta α k puta, dok bi $h(\alpha, T)$ predstavljala srednju vrednost količine informacija dobijenih realizacijom eksperimenta α. Konačno, $h(T)$ jeste maksimum informacija koji se može dobiti ponavljanjem bilo kog eksperimenta, primenjujući transformaciju T.

Izračunavanje vrednosti $h(T)$ po definiciji nije uvek jednostavno. Međutim, ukoliko postoji particija α za koju važi jednakost $h(T, \alpha) = h(T)$, naravno $h(T)$ se može lako naći i to je slučaj kada T ima particiju koja generiše σ–algebru $\mathcal{B}$.

U ovom delu rada se takođe uglavnom bavimo konačnim particijama prostora X. Za proizvoljnu particiju α koristimo oznaku $\alpha_m^n = \vee_{k=m}^n T^{-k}\alpha$. U slučaju $m = \infty$ ili $n = \infty$, koristimo sledeće oznake za σ–algebre generisane particijama koje sadrže: $\alpha_1^\infty = \vee_{k=1}^\infty \mathcal{B}(T^{-k}\alpha)$, $\alpha_{-\infty}^{-1} = \vee_{k=1}^\infty \mathcal{B}(T^k\alpha)$ i $\alpha_{-\infty}^\infty = \vee_{k=-\infty}^\infty \mathcal{B}(T^k\alpha)$. Ove σ–algebre možemo posmatrati kao prošlost, budućnost ili celu istoriju eksperimenta α.

Konačnu particiju α nazivamo *generatorom* u odnosu na transformaciju T ako je $\alpha_{-\infty}^\infty = \mathcal{B}$ skoro svuda.

U neposrednoj vezi sa razmatranom materijom je i sledeća teorema koju navodimo bez dokaza:

Krieger–ova teorema o generatorima. *Ako je T ergodična transformacija koja čuva meru na Lebesgue-ovom prostoru i $h(T) < \infty$, onda T ima konačni generator.*

Na kraju, razmotrićemo i dva konkretna primera izračunavanja entropije.

Primer. (Entropija Bernoulli–jeve sheme) Neka je $B(p_1, \ldots, p_n)$ Bernoulli–jeva shema na alfabetu $\{a_1, \ldots, a_n\}$. Neka su $A_i = \{x | x_0 = a_i\}$, $i = 1, 2, \ldots, n$ skupovi iz merljive particije α. Kako je $\sigma^{-m} A_i = \{x | x_m = a_i\}$, $\vee_{-\infty}^\infty \mathcal{B}(\sigma^{-m}\alpha)$ sadrži sve skupove A_i, pa je samim tim jednak $\mathcal{B}$ do na skupove mere nula. Dakle, α je generator i važi

$$h(\sigma) = h(\alpha, \sigma) = \lim_{m \to \infty} \frac{1}{m} H(\alpha \vee \sigma^{-1}\alpha \vee \cdots \vee \sigma^{-m+1}\alpha).$$

Elementi particije $\alpha \vee \sigma^{-1}\alpha \vee \sigma^{-m+1}\alpha$ su oblika $A_{i1} \cap \sigma^{-1}A_{i2} \cdots \cap \sigma^{-m+1}A_{im}$. Kako su particije $\alpha, \sigma^{-1}\alpha, \ldots, \sigma^{-m+1}\alpha$ nezavisne, iz već dokazane teoreme koja se odnosi na nezavisne particije sledi

$$H(\alpha \vee \sigma^{-1}\alpha \vee \cdots \vee \sigma^{-m+1}\alpha) = mH(\alpha) = -m\sum_{i=1}^{n} p_i \log p_i,$$

pa je srednja vrednost neodređenosti u smislu šta će se sledeće pojaviti u nizu $h(\sigma) = -\sum_{i=1}^{n} p_i \log p_i$.

Dakle, sistem $B(\frac{1}{2}, \frac{1}{2})$ ima entropiju $\log 2$, a entropija sistema $B(\frac{1}{3}, \frac{1}{3})$ iznosi $log_2 3$, pa ova dva sistema ne mogu biti izomorfna.$\triangle$

Primer. (Entropija Markovljevog lanca) Neka je Markovljev lanac definisan matricom $A = (a_{ij})$ formata $n \times n$, tj. $\sum_{j=0}^{n} a_{ij} = 1$, za svaki i $(1 \leq i \leq n)$, $p = p(p_0, \ldots, p_n)$, takav da je $p_i > 0$, za svaki i $(1 \leq i \leq n)$, $\sum_{i=0}^{n} p_i = 1$ i $pA = p$. Neka je α particija iz $\{x | x_0 = a_i\}$, $(1 \leq i \leq n)$. Proizvoljni član particije $\alpha \vee \sigma^{-1}\alpha \vee \cdots \vee \sigma^{-m}\alpha$ je oblika $\{x | x_0 = i_0, \ldots, x_m = i_m\}$ i ima meru $p_{i0}a_{i_0 i_1} \ldots a_{i_{m-1}i_m}$, pa važi

$$\begin{aligned}
&H(\alpha \vee \sigma^{-1}\alpha \vee \cdots \vee \sigma^{-m}\alpha) = \\
&= -\sum_{i_0,\ldots,i_m} f(p_{i0}a_{i_0 i_1} \ldots a_{i_{m-1}i_m}) \\
&= -\sum_{i_0,\ldots,i_m} p_{i0}a_{i_0 i_1} \ldots a_{i_{m-1}i_m}(a_{i_{m-1}i_m} \log p_{i0}a_{i_0 i_1} \ldots a_{i_{m-2}i_{m-1}} + a_{i_{m-1}i_m} \log a_{i_{m-1}i_m}) \\
&= -\sum_{i_0,\ldots,i_{m-1}} \left(\sum_{i_m} a_{i_{m-1}i_m}\right) f(p_{i0}a_{i_0 i_1} \ldots a_{i_{m-2}i_{m-1}}) - \\
&\qquad - \sum_{i_{m-1},i_m} \left(\sum_{i_0,\ldots,i_{m-2}} p_{i0}a_{i_0 i_1} \ldots a_{i_{m-2}i_{m-1}}\right) f(a_{i_{m-2}i_{m-1}}) \\
&= -\sum_{i_0,\ldots,i_{m-1}} f(p_{i0}a_{i_0 i_1} \ldots a_{i_{m-2}i_{m-1}}) - \sum_{i_{m-1},i_m} p_{i_{m-1}} f(a_{i_{m-1}i_m}) \\
&= \cdots = -\sum_i p_i \log p_i - m\sum_{i,j} p_i a_{ij} \log a_{ij}.
\end{aligned}$$

Dakle, entropija Markovljevog lanca definisanog matricom $A = (a_{ij})$ jeste

$$h(\sigma) = -\sum_{i,j} p_i a_{ij} \log a_{ij}.\triangle$$

L i t e r a t u r a

N. Abramson, **Information Theory and Coding,** McGraw–Hill, New York, 1963.

S. Aljančić, **Uvod u realnu i funkcionalnu analizu,** Građevinska knjiga, Beograd, 1974.

M. Arsenović, M. Dostanić, D. Jocić, **Teorija mere, funkcionalna analiza, teorija operatora,** Matematički fakultet, Beograd, 1998.

R. B. Ash, **Information Theory,** John Wiley, New York, 1965.

V. Benci, G. Menconi, *Some remarks on the definition of Boltzmann, Shanonn and Kolmogorov entropy,* neobjavljeno, pp. 1–21.

J. Berkovitz, R. Frigg, F. Kronz, *The ergodic hierarchy, randomness and Hamiltonian chaos,* **Studies in History and Philosophy of Modern Physics 37** (2006), pp. 661-691.

R. E. Blahut, **Theory and Practice of Error Control Codes,** Addison–Wesley, Reading, 1984.

L. Brillouin, **Science and Information Theory,** Academic Press, New York, 1956.

I. Cornfeld, S. Fomin, Y. Sinai, **Ergodic Theory,** Springer, Berlin, 1982.

T. Downarowicz, **Entropy in Dynamical Systems,** Cambridge Univ. Press, Cambridge, 2011.

J.-P. Eckmann, D. Ruelle, *Ergodic theory of chaos and strange attractors,* **Reviews of Modern Physics 57** (1985), pp. 617-654.

D. K. Faddeev, *On the concept of entropy of a finite probabilistic scheme,* (Russian) **Uspehi Mat. Nauk (N.S.) 11** (1956), no. 1(67), 227–231.

R. M. Fano, **Transmission of Information,** The MIT Press, Massachusetts, 1961.

R. Frigg, *In what sense is the Kolmogorov–Sinai entropy a measure for chaotic behaviour? — Bridging the gap between dynamical systems theory and communication theory,* **The British Journal for the Philosophy of Science 55** (2004), pp. 411–434.

R. Frigg, *Chaos and randomness: An equivalence proof of a generalised version of the Shannon entropy and the Kolmogorov-Sinai entropy for Hamiltonian dynamical systems,* **Chaos, Solitons and Fractals 28** (2006), pp. 26-31.

A. Garrido, *Analyzing the entropy concept,* **Advanced Modeling and Optimization 11** (2009), Number 4, pp. 525–530.

P. Gaspard, *Entropy,* **Encyclopedia of Nonlinear Science,** A. Scott (ed.), Routledge, New York, 2005, pp. 264–266.

H. Grad, *The many faces of entropy,* **Communications in Pure and Applied Mathematics 14** (1961), pp. 223-254.

P. Grassberger, *Toward a quantitative theory of self–generated complexity,* **International Journal of Theoretical Physics 25** (1986), pp. 907–938.

P. R. Halmos, **Lectures on Ergodic Theory,** Chelsea Publishing Company, New York, 1956.

R. V. L. Hartley, *Transmission of information,* **Bell System Technical Journal** 7 (1928), str. 535–563.

D. S. Jones, **Elementary Information Theory,** Clarendon Press, Oxford, 1979.

J. Jost, **Dynamical Systems,** Examples with Complex Behaviour, Springer, Berlin, 2005.

A. Katok, B. Hasselblatt, **Introduction to the Modern Theory of Dynamical Systems,** Cambridge Univ. Press, Cambridge, 1995.

A. N. Kolmogorov, *A new metric invariant of transitive dynamical systems and automorphisms of Lebesgue spaces,* **Dokl. Acad. Nauk SSSR 119** (1958), pp. 861-864.

P. R. Masani, **Norbert Wiener 1894–1964,** Birkhäuser Verlag, Basel, 1990.

V. Matković, V. Sinković, **Teorija informacije,** Školska knjiga, Zagreb, 1984.

G. Menges, **Information, Inference and Decision,** D. Reidel, Boston, 1974.

Ž. Pauše, **Uvod u teoriju informacije,** Školska knjiga, Zagreb, 1980.

K. Petersen, **Ergodic Theory,** Cambridge Univ. Press, Cambridge, 1983. (Digitalno izdanje 2000.)

C. E. Shannon, *A mathematical theory of communication,* **Bell System Technical Journal** 27 (1948), str. 379-423 i 623-656.

C. E. Shannon, W. Weaver, **The Mathematical Theory of Communication,** University of Illinois Press, Urbana, Chicago, 1949.

Y. Sinai, *On the concept of entropy for dynamical systems,* **Dokl. Acad. Nauk SSSR 124** (1959), pp. 768-771.

P. Walters, **An Introduction to Ergodic Theory,** Springer, New York, 1982.

C. Werndl, *What are the new implications of chaos for unpredictability?,* **The British Journal for the Philosophy of Science 60** (2009), pp. 195-220.

A. M. Yaglom, I. M. Yaglom, **Probability and Information,** D. Reidel, Dordrecht, 1983.