Temeljna načela prediktivne statističke mehanike kao osnova za teoriju ireverzibilnosti

Kuić, Domagoj

UNIVERSITY OF ZAGREB
FACULTY OF SCIENCE
DEPARTMENT OF PHYSICS

Domagoj Kuić

FOUNDATIONAL PRINCIPLES OF PREDICTIVE STATISTICAL MECHANICS AS THE BASIS FOR THEORY OF IRREVERSIBILITY

DOCTORAL THESIS

Zagreb, 2013.

SVEUČILIŠTE U ZAGREBU
PRIRODOSLOVNO-MATEMATIČKI FAKULTET
FIZIČKI ODSJEK

DOMAGOJ KUIĆ

TEMELJNA NAČELA PREDIKTIVNE STATISTIČKE MEHANIKE KAO OSNOVA ZA TEORIJU IREVERZIBILNOSTI

DOKTORSKI RAD

Mentor:
Prof. dr. sc. Paško Županović

Zagreb, 2013.

UNIVERSITY OF ZAGREB
FACULTY OF SCIENCE
DEPARTMENT OF PHYSICS

Domagoj Kuić

FOUNDATIONAL PRINCIPLES OF PREDICTIVE STATISTICAL MECHANICS AS THE BASIS FOR THEORY OF IRREVERSIBILITY

DOCTORAL THESIS

Supervisor:
Prof. dr. sc. Paško Županović

Zagreb, 2013.

INFORMACIJE O MENTORU

Prof. dr. sc. Paško Županović

Životopis

Rođen 23. siječnja 1954. godine u Splitu.
Diplomirao elektrotehniku 1977. godine na Fakultetu elektrotehnike strojarstva i brodogradnje Sveučilišta u Splitu. Iste godine upisao magistarski studij fizike kondenzirane tvari. Magistrirao 1985. na Prirodoslovno-matematičkom fakultetu Sveučilišta u Zagrebu. Na istoj Ustanovi doktorirao 1988. godine, također na fizici kondenzirane tvari.
Početno područje znanstvenog interesa bila je fizika kondenzirane tvari ili preciznije fizika lančastih vodiča. Osim fizike kondenzirane tvari bavi se fizikom neravnotežnih procesa u fizici i biofizici te edukacijskom fizikom.
Publicirao 25 znanstvenih radova u časopisima koje registrira Current Contents te 5 radova u časopisima s međunarodnom recenzijom.
Vodio projekt "Dielektrična i fotoemisijska svojstva kristala s više elektronskih vrpci" MZOŠ-a.
Suradnik na projektu MZOŠ-a "Razvoj i primjene principa maksimalne proizvodnje entropije" voditelja Davora Juretića s PMF-a u Splitu.
Bio suradnik na slovensko-hrvatskom bilateralnom projektu "Primjena načela najveće proizvodnje entropije u biološkim sustavima" koji je trajao od 2006-2010. god.
Komentor dr. sc. Željani Bonačić Lošić kod izrade njene magistarske i doktorske radnje.
Pod njegovim vodstvom magistrirao prof. Miro Plavšić.
Trenutno mentor dipl. ing. Domagoju Kuiću pri izradi njegove doktorske radnje.
Vodio 18 diplomskih radnji i 8 završnih radova.
Od 1977-1980. radio na Kemijsko-tehnološkom fakultetu u Splitu a od 1980. do danas na Prirodoslovno-matematičkom fakultetu u Splitu.
Predaje kolegije: "Opću fiziku IV", "Mehaniku neprekidnih sredina" i "Termodinamiku nepovratnih procesa".

SAŽETAK

TEMELJNA NAČELA PREDIKTIVNE STATISTIČKE MEHANIKE KAO OSNOVA ZA TEORIJU IREVERZIBILNOSTI

Prediktivna statistička mehanika je oblik zaključivanja iz dostupnih podataka, bez dodatnih pretpostavki, koje nastoji predvidjeti reproducibilne pojave. Primjenom prediktivne statističke mehanike na sustave s Hamiltonovom dinamikom razmotren je problem predviđanja makroskopske vremenske evolucije sustava u slučaju nepotpune informacije o mikroskopskoj dinamici. Cilj istraživanja je analizom temeljnih načela prediktivne statističke mehanike produbiti razumijevanje njene opće primjenjivosti u relaciji s drugim teorijama koje nastoje objasniti pojavu ireverzibilnosti. Osnovna hipoteza je da se to može postići detaljnim razmatranjem uloge koju informacija o sustavu ima u razjašnjenju problema ireverzibilnosti. Hipoteza je provjerena kroz analizu i daljnu generalizaciju rezultata osnovnog teorijskog modela. U modelu zatvorenog Hamiltonovog sustava koji uz Liouvilleovu jednadžbu koristi pojmove teorije informacije analiziran je gubitak korelacije između početnih putanja u faznom prostoru i konačnih mikrostanja, i s tim povezan gubitak informacije o stanju sustava. Primjena načela najveće informacijske entropije maksimizacijom uvjetne informacijske entropije, uz ograničenje koje je dano Liouvilleovom jednadžbom usrednjenom po faznom prostoru, omogućila je definiciju brzine promjene entropije bez dodatnih pretpostavki. Početni model je generaliziran te je uvođenjem dodatnih ograničenja koja su ekvivalentna hidrodinamičkim jednadžbama kontinuiteta doveden u izravnu vezu s poznatim rezultatima iz neravnotežne statističke mehanike i termodinamike ireverzibilnih procesa. Dobiveni rezultati upućuju na općenitu primjenjivost načela prediktivne statističke mehanike i njihovu važnost za teoriju ireverzibilnosti.

ABSTRACT

FOUNDATIONAL PRINCIPLES OF PREDICTIVE STATISTICAL MECHANICS AS THE BASIS FOR THEORY OF IRREVERSIBILITY

Predictive statistical mechanics is a form of inference from available data, without additional assumptions, for predicting reproducible phenomena. By applying predictive statistical mechanics to systems with Hamiltonian dynamics, a problem of predicting the macroscopic time evolution of the system in the case of incomplete information about the microscopic dynamics was considered. The goal of the research is to deepen the understanding, through the analysis of the fundamental principles of predictive statistical mechanics, of its general applicability in relation to other theories that seek to explain the phenomenon of irreversibility. Basic hypothesis is that this can be achieved with the detailed consideration of the role that information about the system has in the clarification of the problem of irreversibility. The hypothesis was tested through the analysis and further generalization of the results of the basic theoretical model. In a model of a closed Hamiltonian system that with the Liouville equation uses the concepts of information theory, analysis was conducted of the loss of correlation between the initial phase space paths and final microstates, and of the related loss of information about the state of the system. Applying the principle of maximum information entropy by maximizing the conditional information entropy, subject to the constraint given by the Liouville equation averaged over the phase space, allowed a definition of the rate of change of entropy without additional assumptions. Initial model was generalized, and with the introduction of the additional constraints which are equivalent to the hydrodynamic continuity equations, brought into direct connection with the known results from the nonequilibrium statistical mechanics and thermodynamics of irreversible processes. The results obtained in this work suggest the general applicability of the principles of predictive statistical mechanics and their importance for the theory of irreversibility.

KLJUČNE RIJEČI

Ključne riječi: načelo najveće informacijske entropije, statistička mehanika, neravnotežna teorija, Hamiltonova dinamika, termodinamika ireverzibilnih procesa

Keywords: maximum entropy principle, statistical mechanics, nonequilibrium theory, Hamiltonian dynamics, thermodynamics of irreversible processes

Sadržaj

Poglavlje 1

UVOD

MaxEnt formalizam, nazvan alternativno i MaxEnt algoritam, formulirao je E. T. Jaynes u poznatim radovima [1, 2] namijenjenim primjeni u statističkoj mehanici. U tim radovima Jaynes je dao puni razvoj rezultata ravnotežne statističke mehanike i Gibbsovog formalizma [3] kao oblika statističkog zaključivanja zasnovanog na Shannonovom konceptu mjere informacije [4]. Shannonova mjera informacije se naziva još informacijska entropija, i u Jaynesovoj interpretaciji predstavlja ispravnu mjeru "iznosa nesigurnosti" reprezentirane raspodjelom vjerojatnosti [5]. Maksimizacija informacijske entropije uz zadana ograničenja je osnovni koncept u Jaynesovom pristupu, koji se naziva *načelo najveće informacijske entropije*. Primjena načela najveće informacijske entropije omogućuje konstrukciju raspodjele vjerojatnosti kojom se u raspodjelu uključuje jedino informacija reprezentirana zadanim ograničenjima, bez ikakvih dodatnih pretpostavki. Jaynesov pristup se temelji na Gibbsovom formalizmu statističke mehanike, za koji je Jaynes smatrao da predstavlja općenitu metodu statističkog zaključivanja u različitim problemima kada dostupna informacija nije potpuna [6]. To uključuje ravnotežnu statističku mehaniku [1], i formulaciju teorije ireverzibilnosti [2], koju je Jaynes pokušao dati u kasnijim radovima [5, 6, 7, 8, 9].

Predikcije i proračuni za različite ireverzibilne procese obično obuhvaćaju tri posebne faze [6]:

(1) Postavljanje "ansambla", tj. odabir početne matrice gustoće, ili u našem slučaju raspodjele N čestica, koja treba opisati naše početno znanje o sustavu koji nas zanima;

(2) Rješavanje problema dinamike; tj. izvođenje vremenske evolucije sustava iz mikroskopskih jednadžbi gibanja;

(3) Izvođenje fizikalnih predikcija iz vremenski evoluiranog ansambla.

Jaynes je prepoznao da dostupnost općenitog rješenja faze (1) pojednostavljuje komplicirano fazu (2). Problem uključuje također i jednako važnu fazu (0) koja se sastoji od neke vrste mjerenja ili opažanja kojom se definiraju zajedno sustav i problem [10]. U izravnim matematičkim pokušajima koji vode do teorije ireverzibilnosti, Liouvilleov teorem sa sačuvanjem volumena faznog prostora koje je inherentno Hamiltonovoj dinamici, često se predstavlja kao jedna od glavnih poteškoća. Relacija Liouvilleove jednadžbe i ireverzibilnog makroskopskog ponašanja sustava jedan je od centralnih problema u statističkoj mehanici. Iz tog razloga ova iznimno komplicirana jednadžba reducira se na ireverzibilnu jednadžbu nazvanu Boltzmannova jednadžba, rate jednadžba ili master jednadžba. S druge strane, Jaynes je smatrao da Liouvilleova jednadžba i s njom povezana konstantnost u vremenu Gibbsove H funkcije ne generiraju poteškoće, nego predstavljaju upravo ono dinamičko svojstvo koje potrebno za razrješenje ovog problema, smatrajući da je problem više konceptualne nego matematičke prirode [5, 7].

Matematička jasnoća ovog stajališta ima bazu u graničnom teoremu koji je dao Shannon [4], poznatom kao asimptotski ekviparticijski teorem teorije informacije. Primjena tog teorema povezuje u određenim slučajevima, u granici velikog broja čestica, Boltzmannovu formulu za entropiju makrostanja i Gibbsovu H funkciju [6, 7, 9]. Matematička povezanost s Boltzmannovom interpretacijom entropije kao logaritma broja načina (ili mikrostanja) kojim makroskopsko stanje može biti realizirano, tada daje jednostavnu fizikalnu interpretaciju Gibbsovom formalizmu, i njegovoj generalizaciji u obliku MaxEnt formalizma. Maksimizacijom informacijske entropije uz zadana ograničenja predviđa se makroskopsko ponašanje koje se može dogoditi na najveći broj načina kompatibilno sa informacijom koja je reprezentirana zadanim ograničenjima. U primjeni na vremenski ovisne procese, Jaynes je to nazvao načelom najvećeg kalibra [8, 9]. Jaynes je jasno istaknuo da prediktivna statistička mehanika ne predstavlja fizikalnu teoriju koja objašnjava ponašanje različitih sustava deduktivnim zaključivanjem iz prvih principa, nego oblik statističkog zaključivanja koje daje predikcije opservabilnih pojava iz nepotpune informacije [8]. Iz tog razloga prediktivna statistička mehanika ne može pružiti sigurnost za svoje predikcije na način na koji to može deduktivna teorija. To ne znači da prediktivna statistička mehanika ignorira zakone mikrofizike; ona koristi sve što je poznato o strukturi mikrostanja i sve podatke o makroskopskim veličinama, bez bilo kakvih dodatnih fizikalnih pretpostavki izvan onoga što je dano dostupnom informacijom. Važno je primjetiti da su oštre, određene predikcije makroskopskog ponašanja moguće samo kada je neko ponašanje karakteristično za golemu većinu mikrostanja kompatibnih sa ograničenjima. Iz istog razloga, to je upravo ponašanje koje se eksperimentalno reproducira pod tim ograničenjima; to je u osnovi poznato kao načelo makroskopske uniformnosti [1, 2], ili kao *načelo makroskopske reproducibilnosti* [9]. U nešto drugačijem kontekstu ovo svojstvo je prepoznato kao koncept makroskopskog determinizma, čija precizna definicija uključuje neku vrstu termodinamičke granice [11].

Sa Jaynesovog stanovišta, dinamička invarijantnost Gibbsove H funkcije daje jednostavan dokaz drugog zakona, koji je tada poseban slučaj općenitog uvjeta za eksperimentalnu reproducibilnost bilo kojeg makroskopskog procesa [7]. Polazeći od jednostavne demonstracije zasnovane na Liouvilleovom teoremu, to omogućuje Jaynesu da generalizira drugi zakon izvan ograničenja početnih i konačnih ravnotežnih stanja, razmatrajući ga kao poseban slučaj općenitog ograničenja na smjer bilo kojeg reproducibilnog procesa [7, 12]. Prema Jaynesu [7] pravi razlog za drugi zakon, budući da je volumen u faznom prostoru sačuvan u dinamičkoj evoluciji, je temeljni zahtjev na bilo koji reproducibilni proces da volumen W' faznog prostora, kompatibilan s konačnim (makroskopskim) stanjem, ne može biti manji od volumena W_0 faznog prostora, koji opisuje našu sposobnost da reproduciramo početno stanje. Volumen W_0 faznog prostora mjeri stupanj nepoznavanja stvarnog ali nepoznatog mikrostanja kada su poznate jedino vrijednosti makroskopskih parametara, ali također mjeri i stupanj kontrole koji eksperimentator tada ima nad mikroskopskim stanjem sustava [7]. Jaynes [7] i Grandy [13, str. 148-150] su proveli ovu analizu za reproducibilan adijabatski proces s ravnotežnim početnim i konačnim stanjem. U toj analizi W_0 i W' predstavljaju volumen područja velike vjerojatnosti u faznom prostoru s mikrostanjima kompatibilnim s početnim makroskopskim stanjem, i volumen područja velike vjerojatnosti s mikrostanjima kompatibilnim s konačnim makroskopskim stanjem, redom (za precizniju definiciju područja velike vjerojatnosti vidi reference [7], [13, str. 45], ili fusnotu 1). Ako ta područja označimo sa M_0 i M', nužan uvjet za reproducibilnost procesa je $M_t \subseteq M'$, gdje je M_t područje s volumenom W_t u koje se M_0 bijektivno preslika Hamiltonovim gibanjem u faznom prostoru $M_0 \to M_t$. Zbog Liouvilleovog teorema, logička posljedica ovih tvrdnji je $W_0 = W_t \leq W'$. Dakle, u ovoj analizi nejednakost $W_0 \leq W'$ je nužan uvjet da bi makroskopski proces bio reproducibilan za bilo koje nepoznato mikrostanje u W_0.

Za sustave s velikim brojem čestica, Boltzmannova formula za entropiju makrostanja $S = \log W$ i Gibbsova H funkcija su povezane[1], pa ova analiza vodi do drugog zakona

[1]Pri tom se misli na povezanost u "termodinamičkoj granici" velikog broja čestica $N \to \infty$ uz konstantnu gustoću, ako relativne fluktuacije u energiji teže ka nuli kao $N^{-1/2}$. U toj granici H/N je konačan i vrijedi $\lim_{N\to\infty} \{[H + \log W(\epsilon)]/N\} = 0$. Pri tome je $W(\epsilon)$ volumen područja velike vjerojatnosti u faznom prostoru koje je definirano tako da se sastoji od svih točaka gdje je vrijednost N-čestične raspodjele vjerojatnosti $f(q,p) \geq C$, gdje je C konstanta odabrana tako da je ukupna vjerojatnost tog područja jednaka $1 - \epsilon$, za $0 < \epsilon < 1$. Glavno svojstvo ovog rezultata je neovisnost o ϵ. To znači da u granici velikog sustava $\log W(\epsilon)/N$ postaje jednak H/N, neovisno o ϵ [7, 9], [13, str. 45]. Dakle, za velike sustave, Gibbsova H funkcija mjeri logaritam volumena područja velike vjerojatnosti u faznom prostoru, ali na način koji nije ovisan o preciznoj definiciji velike vjerojatnosti. U određenim slučajevima ovaj rezultat je posljedica teorema koji je dao Shannon [4, odjeljak 21], poznatog kao asimptotski ekviparticijski teorem

termodinamike u obliku u kojem taj zakon vrijedi za reproducibilan adijabatski proces s ravnotežnim početnim i konačnim stanjem [7]. Pri tome je važno napomenuti da se u tim razmatranjima radi o makrostanjima koja su eksperimentalno reproducibilna kontrolom određenog skupa makrovarijabli koje definiraju odgovarajući termodinamički sustav. Dani fizikalni sustav odgovara različitim termodinamičkim sustavima ovisno o tome koji skup makrovarijabli opažamo i/ili kontroliramo, pri čemu se vrijednosti termodinamičke entropije mogu razlikovati [7, 12], [13]. Grandy je istaknuo da drugi zakon termodinamike predstavlja [13, str. 150] tvrdnju o razlikama u eksperimentalnim entropijama između dva ravnotežna stanja istog termodinamičkog sustava. Uvođenje dodatnih makroskopskih varijabli izvan danog skupa makrovarijabli znači promjenu termodinamičkog sustava a time i problema koji se razmatra [7], [13].

Argumenti koji su korišteni u ovoj demonstraciji impliciraju također pitanje koja neravnotežna ili ravnotežna stanja su dostupna iz drugih, što nije moguće precizno odrediti bez informacije o dinamici, konstantama gibanja, nametnutim ograničenjima, itd. Jaynes je isticao da drugi zakon termodinamike predviđa jedino da će promjena makroskopskog stanja ići u općenitom smjeru veće konačne entropije, ali ne i kojom brzinom, ili duž koje putanje [8, 9, 12]. Očito je da su bolje predikcije moguće jedino uz uvođenje dodatnih podataka. Makroskopska stanja veće entropije mogu biti realizirana na značajno veći broj načina, što je osnovni razlog visoke pouzdanosti Gibbsovih ravnotežnih predikcija [9]. U ovom kontekstu Jaynes je također spekulirao da je slučajan uspjeh u obratu ireverzibilnog procesa eksponencijalno male vjerojatnosti [12].

Jaynesova interpretacija ireverzibilnosti i drugog zakona odražava stajalište eksperimentatora. Zurek [14] je predložio definiciju fizikalne entropije kao zbroja nedostajuće informacije o mikroskopskom stanju, dane Shannonovom informacijskom entropijom, i algoritamskog sadržaja informacije prisutne u dostupnim podacima o sustavu. U granici Zurekovog pristupa u kojoj je mjerenje potpuno i mikrostanje je poznato, fizikalna entropija sustava dana je algoritamskim sadržajem informacije o mikroskopskom stanju u kojem se sustav nalazi [14]. Zurekova interpretacija fizikalne entropije i termodinamike dana je na razini opažača koji mjerenjem prikupljaju podatke i obrađuju ih u skladu s osnovnim zakonima računanja na način koji je analogan Turingovim strojevima. Jaynes je zastupao stajalište prema kojem mjerenja [2] uvijek reprezentiraju daleko manje od maksimalnog opažanja koje bi nam omogućilo da odredimo čisto stanje (tj. mikroskopsko stanje sustava). Prema Jaynesovom stajalištu to je razlog zbog kojeg [2] moramo pribjeći MaxEnt metodi da bi reprezentirali naš stupanj znanja o sustavu na način koji je slobodan od proizvoljnih pretpostavki u odnosu na informaciju koja nedostaje.

MaxEnt algoritam je općenita metoda konstrukcije raspodjele vjerojatnosti primje-

teorije informacije.

nom načela najveće informacijske entropije u slučajevima kada raspodjela nije jedinstveno određena dostupnom informacijom. Proizvoljne pretpostavke mogu se izbjeći odabirom raspodjele vjerojatnosti koja je kompatibilna s dostupnom informacijom, a karakterizira je najveća nesigurnost povezana s informacijom koja nedostaje. Zaključci koji se izvode iz takve raspodjele vjerojatnosti ovise jedino o stvarnom stupnju znanja [1, 2]. Raspodjela vjerojatnosti koja maksimizira informacijsku entropiju (nesigurnost) uz ograničenja zadana dostupnim makroskopskim podacima, u prediktivnoj statističkoj mehanici reprezentira stvarnu nesigurnost povezanu s nedostatkom informacije o točnom mikrostanju sustava.

U ovom radu se razmatra primjena prediktivne statističke mehanike na problem predikcije makroskopske vremenske evolucije sustava s Hamiltonovom dinamikom, u slučaju kad informacija o mikroskopskoj dinamici sustava nije potpuna. Cilj istraživanja je analizom temeljnih načela prediktivne statističke mehanike produbiti razumijevanje njene opće primjenjivosti u relaciji s drugim teorijama koje nastoje objasniti pojavu ireverzibilnosti. Osnovna hipoteza je da se to može postići kroz primjenu temeljnih načela prediktivne statističke mehanike na problem koji se razmatra u ovom radu, ako se u razmatranje uključi uloga koju informacija o sustavu ima u razjašnjenju problema ireverzibilnosti.

Temeljna načela prediktivne statističke mehanike su razmotrena u poglavlju 2. Argumenti na kojima se temelji Shannonova [4] definicija informacijske entropije kao mjere nesigurnosti reprezentirane raspodjelom vjerojatnosti, i Shannonov dokaz jedinstvenosti te definicije, dani su u odjeljku 2.1. Aksiomatski okvir u kojem je dana općenita definicija vjerojatnosti prikazan je u odjeljku 2.2. Jaynesova formulacija [1, 2] načela najveće informacijske entropije kao općenitog kriterija za konstrukciju raspodjele vjerojatnosti kada dostupna informacija nije dovoljna za njeno jedinstveno određivanje, argumentirana je u odjeljku 2.3. Pokazano je da se konstrukcija raspodjele vjerojatnosti primjenom načela najveće informacijske entropije temelji na interpretaciji vjerojatnosti kao reprezentacije stupnja znanja, i na Shannonovoj definiciji mjere nesigurnosti reprezentirane raspodjelom vjerojatnosti. U odjeljku 2.3 su također dani matematički dokazi tvrdnji koje su važne za jednoznačnost i ispravnost rezultata koji slijede iz primjene MaxEnt algoritma i MaxEnt formalizma. Jaynesova formulacija načela makroskopske reproducibilnosti [1, 2, 9] prethodno je već prikazana u ovom uvodu. U kontekstu primjene načela najveće informacijske entropije u statističkoj fizici, interpretacija MaxEnt formalizma i veza s Gibbsovim formalizmom uspostavlja se u odjeljku 2.4.

U cilju istraživanja temeljnih načela prediktivne statističke mehanike kroz primjenu na problem koji se razmatra u radu, razvijen je osnovni teorijski model makroskopske vremenske evolucije zatvorenih Hamiltonovih sustava. Koncepti Hamiltonove mehanike i raspodjela vjerojatnosti u faznom prostoru koji se koriste u modelu definirani su u poglavlju 3. Model je postavljen i razmotren u poglavlju 4. Rezultati ovog istraživanja prethodno su bili objavljeni u članku [15]. Osnovna hipoteza ovog rada provjerava se u poglavlju 5 analizom rezultata, daljnom generalizacijom osnovnog teorijskog modela i usporedbom rezultata koji proizlaze iz generalizacije pristupa s poznatim rezultatima neravnotežne teorije.

U okviru osnovnog teorijskog modela razmotrene su dvije različite primjene MaxEnt algoritma na predikciju makroskopske vremenske evolucije zatvorenih sustava sa Hamiltonovom dinamikom. U odjeljcima 3.1 i 3.2 definiran je koncept *putanja u faznom prostoru* određenih rješenjima Hamiltonovih jednadžbi i koncept *raspodjele vjerojatnosti putanja*. Za mikrostanja u faznom prostoru je definirana *uvjetna raspodjela vjerojatnosti* uz uvjet specificirane putanje. Ovim raspodjelama odgovaraju *informacijska entropija putanja* i *uvjetna informacijska entropija* koje su definirane u odjeljku 4.1, u korespondenciji s definicijama odgovarajućih veličina u Shannonovoj teoriji informacije [4].

Prvi pristup osnovnom teorijskom modelu razmotren je u odjeljcima 4.1 i 4.2. U prvom pristupu je Liouvilleova jednadžba za uvjetnu raspodjelu vjerojatnosti uvedena kao strogo *mikroskopsko ograničenje* na vremensku evoluciju u faznom prostoru. Vremenska evolucija uvjetne raspodjele vjerojatnosti je tada potpuno određena ovim ograničenjem i početnim vrijednostima. Maksimizacijom uvjetne informacijske entropije uz to ograničenje predviđa se makroskopsko ponašanje koje se može realizirati na najveći broj načina, kompatibilno s informacijom o mikroskopskoj dinamici. Informacija o mikroskopskoj dinamici je reprezentirana Hamiltonovim jednadžbama i skupom mogućih putanja u faznom prostoru. Ako se vjerojatnosti interpretiraju isključivo u objektivnom smislu, kao svojstvo sustava a ne kao reprezentacija našeg stupnja znanja, puno opravdanje ovog pristupa je moguće jedino ako je naše znanje o mikroskopskoj dinamici potpuno.

Prvi pristup osnovnom teorijskom modelu omogućio je strogu definiciju koncepata koji su baza za drugi pristup koji je također razmotren u odjeljcima 4.1 i 4.2. Razlika u odnosu na prvi pristup je u uvođenju Liouvilleove jednadžbe za uvjetnu raspodjelu vjerojatnosti kao *makroskopskog ograničenja*. Ovo ograničenje na vremensku evoluciju u faznom prostoru uzeto je uz usrednjavanje, koje je dano integralom Liouvilleove jednadžbe za uvjetnu raspodjelu vjerojatnosti po dostupnom faznom prostoru. U tom smislu slično je ograničenjima koja su dana podacima o makroskopskim veličinama. U Jaynesovoj prediktivnoj statističkoj mehanici više objektivnosti se pripisuje eksperimentalno mjerenim veličinama nego raspodjelama vjerojatnosti. Subjektivno gledište koje ovdje postaje bitno sastoji se od toga da su vjerojatnosti pridružene zbog nepotpunog znanja, tj. zbog parcijalne informacije, i zato reprezentiraju naš stupanj znanja o sustavu. Ako informacija o dinamici nije dovoljno iscrpna da se detaljno odredi vremenska evolucija, uzima se prosjek po svim slučajevima koje smatramo mogućim na osnovi parcijalne informacije. U drugom pristupu osnovnom modelu pokazuje se da su tada elementi ireverzibilnog makroskopskog ponašanja u zatvorenim sustavima s Hamiltonovom dinamikom posljedica postupnog gubitka informacije o mogućim mikrostanjima sustava.

Ovu ideju je razvio Jaynes u formalizmu matrice gustoće [2]. Grandy [16, 17] je u okviru MaxEnt formalizma razvio detaljan model vremenski ovisnih vjerojatnosti i matrice gustoće za makroskopske sustave s ograničenjima koja se mijenjaju u vremenu, i primijenio ga na tipične procese u neravnotežnoj termodinamici i hidrodinamici [18]. U kontekstu uzajamnog djelovanja makroskopskih ograničenja na sustav i njegove mikroskopske dinamike, zanimljivo je spomenuti da je MaxEnt formalizam također bio razmatran kao metoda približnog rješavanja parcijalnih diferencijalnih jednadžbi koje opisuju vremensku evoluciju raspodjela vjerojatnosti. Zbog potpunije daljne reference, ovdje jedino spominjemo da je ova metoda, uz ostale primjere, bila primijenjena na Liouville–von Neumannovu jednadžbu [19], familiju dinamičkih sustava sa bezdivergentnim tokovima u faznom prostoru koja uključuje Hamiltonove sustave [20], generaliziranu Liouvilleovu jednadžbu i jednadžbe kontinuiteta [21]. Univerzalnost ovog pristupa je uspostavljena za općenitu klasu evolucijskih jednadžbi koje zadovoljavaju osnovne zahtjeve linearnosti i sačuvanja normalizacije [22]. Ova metoda je također bila razmatrana za klasične evolucijske jednadžbe sa članom izvora unutar konteksta u kojem normalizacija nije sačuvana [23].

U Jaynesovoj [2] interpretaciji ireverzibilnost fizikalnih procesa odražava jedino našu nemogućnost praćenja stanja sustava tijekom procesa. S tim je povezan i postupni gubitak informacije o stanju sustava. U odjeljcima 4.1 i 4.2 je pokazano da takva interpretacija ima jasnu matematičku formulaciju u konceptima maksimizacije uvjetne informacijske entropije i njenoj relaciji s informacijskom entropijom. Vrijednosti uvjetne informacijske entropije i informacijske entropije uvijek zadovoljavaju relaciju nejednakosti iz Shannonove teorije informacije [4], po kojoj je informacijska entropija gornja granica za uvjetnu informacijsku entropiju. Maksimizacijom uvjetne informacijske entropije uz ograničenje koje je zadano Liouvilleovom jednadžbom usrednjenom po faznom prostoru, ta relacija informacijskih entropija postaje jednakost. Jednakost uvjetne informacijske entropije i informacijske entropije je ekvivalentna statističkoj nezavisnosti početnih putanja u faznom prostoru i konačnih mikrostanja. Logička posljedica statističke nezavisnosti je *potpuni gubitak korelacije* između početnih putanja i konačnih mikrostanja. Logički ispravna interpretacija gubitka korelacije zahtijeva i definiciju karakterističnog vremena koje je potrebno za gubitak korelacije.

Ključan element pristupa kojim je u teorijski model uključen opisani gubitak korelacije je uvođenje Liouvilleove jednadžbe za uvjetnu raspodjelu vjerojatnosti kao makroskopskog ograničenja. Na taj način je nepotpunost informacije o mikroskopskoj dinamici uključena, konzistentno s temeljnim načelima prediktivne statističke mehanike, u problem predikcije makroskopske vremenske evolucije sustava. Maksimizacija uvjetne informacijske entropije uz ograničenje zadano Liouvilleovom jednadžbom usrednjenom po faznom prostoru, rezultira potpunim gubitkom korelacije između početnih putanja u faznom prostoru i konačnih mikrostanja. Ovaj gubitak korelacije je povezan s gubitkom informacije o mogućim mikrostanjima sustava. Maksimum uvjetne informacijske entropije i odgovarajuća vrijednost informacijske entropije jednaki su Boltzmann-Gibbsovoj entropiji sustava, koja je dana logaritmom volumena dostupnog faznog prostora. U osnovnom teorijskom modelu je bez dodatnih pretpostavki na taj način proizišla definicija promjene entropije i brzine promjene entropije za zatvoren Hamiltonov sustav. Integral po vremenu Lagrangeova multiplikatora uzet preko vremenskog intervala koji je dulji od karakterističnog vremena gubitka korelacije daje promjenu vrijednosti uvjetne informacijske entropije u odnosu na njenu vrijednost u početnom trenutku. Daljnje promjene informacijskih entropija su jednake promjeni Boltzmann-Gibbsove entropije sustava. Na taj način, za sve slijedeće vremenske intervale integral Lagrangeova multiplikatora po tim intervalima i Lagrangeov multiplikator mogu se identificirati s promjenom entropije i *brzinom promjene entropije*, redom. Ovaj rezultat u skladu je s interpretacijom ireverzibilnosti kao posljedice gubitka informacije o mogućim mikrostanjima sustava.

Poglavlje 5 posvećeno je generalizaciji pristupa koji je razvijen u okviru osnovnog teorijskog modela. U odjeljku 5.1 provedena je analiza rezultata osnovnog teorijskog modela s ciljem daljne generalizacije pristupa. Zaključeno je da za generalizaciju pristupa treba u osnovni teorijski model uključiti podatke o onim makroskopskim veličinama koje su relevantne za predikcije na specificiranim vremenskim skalama. Uvjeti takve generalizacije razmotreni su u odjeljku 5.2 kroz prikaz reduciranog opisa neravnotežnih sustava, gdje se dijelom slijedi pristup iz reference [24]. U odjeljku 5.2 je za razrijeđen klasični plin identičnih čestica dana hijerarhija vremenskih skala, uz specifikaciju skupa makroskopskih veličina koje su relevantne za opis neravnotežnog sustava na pojedinim vremenskim skalama. Taj primjer je preuzet iz reference [24]. Za prvi korak u generalizaciji pristupa je odabrana hidrodinamička vremenska skala, na kojoj je za opis neravnotežnog sustava potrebna manje detaljna informacija o mikroskopskoj dinamici u usporedbi s drugim vremenskim skalama.

U odjeljku 5.3 su za klasični fluid identičnih čestica, uz pretpostavku lokalne ravnoteže, dane relevantne lokalne dinamičke varijable i odgovarajuće lokalne makroskopske veličine. Budući da su te lokalne dinamičke varijable gustoće sačuvanih veličina, njihove jednadžbe gibanja mogu se zapisati u obliku lokalnih mikroskopskih zakona sačuvanja. Odgovarajuće lokalne makroskopske veličine zadovoljavaju lokalne makroskopske zakone sačuvanja. U kontekstu pristupa koji je prikazan u ovom radu, razmotren je slučaj kada informacija o mikroskopskoj dinamici fluida nije potpuna, uz pretpostavku da tada Liouvilleova jednadžba vrijedi u obliku o kojem nemamo potpunu informaciju. Uz ovu pretpostavku specificiran je uvjet uz koji lokalni mikroskopski zakoni sačuvanja vrijede u obliku koji ne ovisi o informaciji o mikroskopskoj dinamici koja nedostaje. Uz istu pretpostavku specificiran je uvjet uz koji lokalni makroskopski zakoni sačuvanja vrijede u obliku koji ne ovisi o informaciji o mikroskopskoj dinamici koja nedostaje. Prvom specifikacijom je obuhvaćen slučaj kada informacija o mikroskopskoj dinamici koja nedostaje nije relevantna za opis vremenske evolucije lokalnih dinamičkih varijabli. Drugom specifikacijom je obuhvaćen slučaj kada informacija o mikroskopskoj dinamici koja nedostaje nije relevantna za opis vremenske evolucije lokalnih makroskopskih veličina. Lokalni makroskopski zakoni sačuvanja za fluide poznati su kao hidrodinamičke jednadžbe kontinuiteta, i osnova su za izvod jednadžbi hidrodinamike. S obzirom da su hidrodinamičke jednadžbe kontinuiteta osnovni element reduciranog opisa na hidrodinamičkoj vremenskoj skali, u analizi na kraju odjeljka 5.3 one se uzimaju za relevantnu dodatnu informaciju koju kao dodatna ograničenja treba uključiti u početni model. U odjeljku 5.3 je također izveden ekvivalentan oblik tih jednadžbi koji je pogodan za korištenje u varijacijskom računu.

Predikcije koje slijede iz maksimizacije uvjetne informacijske entropije, uz ograničenje zadano Liouvilleovom jednadžbom usrednjenom po dostupnom faznom prostoru, i dodatna ograničenja ekvivalentna hidrodinamičkim jednadžbama kontinuiteta, izvedene su i razmotrene u odjeljku 5.4. Maksimizacija uvjetne informacijske entropije uz dodatna ograničenja uvedena u generaliziranom pristupu, također rezultira potpunim gubitkom korelacije između početnih putanja u faznom prostoru i konačnih mikrostanja. Dobivena raspodjela vjerojatnosti mikrostanja po obliku je identična relevantnoj raspodjeli za klasični fluid u lokalnoj ravnoteži koja je poznata iz literature [24]. Uz pretpostavku lokalne ravnoteže, Lagrangeovi multiplikatori u raspodjeli vjerojatnosti mikrostanja, usporedbom te raspodjele s relevantnom raspodjelom, dovedeni su u vezu s lokalnim parametrima koji imaju jednostavnu termodinamičku interpretaciju. Pretpostavka lokalne ravnoteže omogućila je precizniju definiciju karakterističnog vremena potrebnog za gubitak korelacije. Ono je na taj način dovedeno uz vezu s vremenom koje je potrebno za uspostavljanje stanja lokalne ravnoteže fluida s odgovarajućom relevantnom raspodjelom. Kao rezultat generaliziranog pristupa, dobiven je izraz za brzinu promjene entropije koji je u skladu s poznatim izrazom za brzinu promjene entropije zatvorenog sustava iz termodinamike ireverzibilnih procesa. Taj rezultat je omogućio definiciju *gustoće produkcije entropije* za klasični fluid identičnih čestica u skladu s temeljnim postulatima termodinamike ireverzibilnih procesa. Time je pokazano da uvođenje dodatnih ograničenja koja su relevantna za hidrodinamičku vremensku skalu, preciznije određuje brzinu promjene entropije sustava, čija je definicija proizišla u osnovnom teorijskom modelu.

U poglavlju 6 je dana kratka diskusija otvorenih pitanja uz njihov slobodan prikaz. U poglavlju 7 dan je zaključak koji slijedi iz rezultata ovog rada.

Poglavlje 2

OSNOVNI ELEMENTI PREDIKTIVNE STATISTIČKE MEHANIKE

2.1 Entropija teorije informacije kao svojstvo raspodjele vjerojatnosti

Shannonovim radom [4] je na dosljedan način dan odgovor na osnovno pitanje teorije informacije, a to je kako definirati iznos informacije u komunikacijskom sustavu. U razvoju teorije informacije i u primjenama u tehničkim znanostima Shannonov je doprinos imao utjecaj koji nije potrebno posebno naglašavati. Utjecaj tog otkrića je značajan i u drugim područjima, gdje je iz različitih razloga također potrebna precizna definicija mjere informacije. Informacija je općenit pojam koji ima i značenja koja vjerojatno nisu matematički definirana niti su obuhvaćena pojmovima koji se koriste u teoriji informacije. Općenito svojstvo informacije je da se razlikuje od svoje reprezentacije. S druge strane, moguće je relativno dobro definirati iznos informacije koja je povezana s mjerenjima i opažanjima fizikalnih procesa. Informacija je tako pojam koji je zanimljiv i za fiziku.

Shannonova teorija komunikacije imala je brojne prethodnike, poput Nyquista, Hartleya i Wienera. Mjera informacije je istraživana i prije Shannonovog rada (više detalja može se pronaći u referencama [4], [13], [25]). Polazeći od razumijevanja da problem konstrukcije komunikacijskog sustava ovisi o statističkoj strukturi informacije koju se želi njime komunicirati, Shannon je dao najopćenitiju definiciju mjere informacije do tada. Ako simboli $A_1, A_2, \ldots, A_n$ nekog alfabeta (događaji) imaju vjerojatnosti $p_1, p_2, \ldots, p_n$, mjera iznosa informacije je definirana izrazom[1]

$$H(p_1, \ldots, p_n) = -K \sum_{i=1}^{n} p_i \log p_i \ . \tag{2.1}$$

Nizovi simbola, odnosno "slova", mogu činiti skup "riječi" određene duljine, a iznos informacije može se odrediti na analogan način. Pozitivna konstanta K u izrazu (2.1) ovisi o izboru jedinice kojom se izražava iznos informacije. Ako se informacija kodira nizovima binarnih znamenki, prikladno je u izrazu (2.1) koristiti logaritam po bazi 2, i na taj način uz $K = 1$ izraziti H u bitovima.

Poznato je da za iznos informacije po simbolu (događaju), dan izrazom (2.1), vrijedi nejednakost

$$H(1/n, \ldots, 1/n) \geq H(p_1, \ldots, p_n) \ . \tag{2.2}$$

Za konkavne funkcije realne varijable vrijedi Jensenova nejednakost [26],

$$f\left(\sum_{i=1}^{n} \lambda_i x_i\right) \geq \sum_{i=1}^{n} \lambda_i f(x_i) \ , \tag{2.3}$$

gdje je $\sum_{i=1}^{n} \lambda_i = 1$, $\lambda_i \geq 0$ za $i = 1, 2, \ldots, n$. Relacija (2.2) slijedi iz nejednakosti (2.3) uvrštavanjem: $f(p_i) = -p_i \log p_i$, $\lambda_i = 1/n$ za $i = 1, 2, \ldots, n$, i korištenjem uvjeta koji vrijedi za vjerojatnosti,

$$\sum_{i=1}^{n} p_i = 1 \ . \tag{2.4}$$

Uz uvjet (2.4) izraz (2.1) ima maksimum kad je svih n vjerojatnosti jednako, $p_i = 1/n$, $i = 1, 2, \ldots, n$. Dokaz se može naći u referenci [27, str. 14–16].

S obzirom na način na koji vjerojatnosti p_i ulaze u izraz (2.1), nije teško razumjeti zašto se u literaturi [28] može još susresti i naziv srednji iznos informacije, pri čemu se $\log(1/p_i)$ interpretira kao iznos informacije simbola, odnosno događaja s oznakom i. U stvarnim primjenama radi se o duljinama kodnih riječi, npr. nizova binarnih znamenki, a izraz (2.1) predstavlja očekivani broj bitova nužan za kodiranje poruka (po simbolu). Dakle, ova interpretacija iznosa informacije se odnosi na prethodno kodirane i naknadno primljene poruke, ali ona nije neophodna da se opravda osnovna definicija iznosa informacije (2.1). Za vjerojatnosti $p_i < 1/n$, funkcija $\log(1/p_i)$ ima veću vrijednost od maksimuma iznosa informacije (2.2), ali je srednji iznos informacije tada manji od maksimuma. U granici $p_i \to 0$ je doprinos srednjem iznosu informacije jednak nula. Međutim, nejasnoće nastaju kada se u općenitijem kontekstu od ovog prethodno navedenog, $\log(1/p_i)$ interpretira

[1] U ovom odjeljku se privremeno za mjeru informacije koristi oznaka H koju je koristio Shannon. Kasnije će se oznaka H koristiti za Hamiltonovu funkciju. Na mjestima na kojima nije specificirana baza logaritma, $\log x$ predstavlja logaritam po bazi koja je tada određena kontekstom. Najčešće predstavlja logaritam po bazi e (prirodni logaritam), koji je uobičajeno označavati s $\ln x$.

kao iznos informacije koju *donosi* pojava pojedinačnog događaja. Alternativno, $\log(1/p_i)$ naziva se još nesigurnost ili iznenađenje.

Na primjeru dva najjednostavnija slučaja je lako vidjeti da je također moguća i drugačija interpretacija. Ako svaki od n mogućih događaja ima jednaku vjerojatnost, tj. $p_i = 1/n$ za $i = 1, 2, \dots, n$, tada je iznos informacije (2.1) maksimalan i jednak $H(1/n, \dots, 1/n) = K \log n$. S druge strane, ako je raspodjela vjerojatnosti dana s $p_i = \delta_{ij}$ (tj. događaj označen s j je siguran), tada je $H(0, \dots, p_j = 1, \dots, 0) = 0$. S obzirom da je veličina (2.1) definirana kao mjera iznosa informacije, iz razmatranja ova dva slučaja može, alternativno, proizići i slijedeća interpretacija. Prije pojave jednog od n jednako vjerojatnih događaja, nesigurnost, odnosno iznos *informacije koja nam nedostaje* za određivanje događaja, jednaka je $H = K \log n$. Nakon pojave jednog od n događaja, može se također reći da $K \log n$ predstavlja iznos informacije dobivene njegovim određivanjem. Međutim, za razliku od prethodne interpretacije koja je zasnovana na $\log(1/p_i)$ kao iznosu informacije događaja, ova interpretacija slijedi iz funkcije (2.1) koja ovisi o vjerojatnostima $p_1, p_2, \dots, p_n$. Dakle, ne možemo biti sigurni da je jedna od specifičnih interpretacija iznosa informacije koju donosi pojava događaja uvijek ispravna u svim kontekstima. Nazivi srednji iznos informacije, iznos informacije koja nam nedostaje i iznos informacije dobivene pojavom nekog događaja, koriste se u različitim kontekstima iz kojih onda proizlaze i njihove točne interpretacije.

Shannonova interpretacija međutim nije ovisna o bilo kojem od tih specifičnih konteksta. Funkciji (2.1) dao je interpretaciju mjere "izbora" uključenog u odabir događaja. Preciznije, Shannon je funkciju (2.1) definirao kao mjeru *naše nesigurnosti* vezane uz pojavu mogućih događaja, odnosno mjeru nesigurnosti *reprezentirane raspodjelom vjerojatnosti* $p_1, p_2, \dots, p_n$. Jaynes je smatrao da je u kontekstu teorije komunikacije Shannonova mjera informacije i nesigurnosti povezana s prethodnim znanjem inženjera o porukama koje će se slati komunikacijskim sustavom koji konstruira i njihovim vjerojatnostima [25].

Shannon je funkciju (2.1) nazvao *entropija* skupa vjerojatnosti, odnosno entropija raspodjele vjerojatnosti $p_1, p_2, \dots, p_n$. Odabir tog naziva uzrokovao je čestu konfuziju s termodinamičkom entropijom. Entropija je u termodinamici svojstvo termodinamičkog sustava koji je definiran eksperimentalno mjerljivim makroskopskim veličinama; ona je funkcija drugih termodinamičkih veličina. S druge strane, funkcija (2.1) predstavlja svojstvo raspodjele vjerojatnosti. Prema literaturi, Shannon je izvorno namjeravao funkciju (2.1) nazvati nesigurnost, jer se može intepretirati kao mjera nesigurnosti reprezentirane raspodjelom vjerojatnosti.[2] Nadalje, definicija (2.1) nije dana u fizikalnom kontekstu[3], a

[2]Prema brojnim navodima iz literature (jedan od primjera je i referenca [13]), Shannon je funkciju (2.1) nazvao *entropija* na nagovor poznatog matematičara von Neumanna.

[3]Isto tako niti interpretacije koje smo prethodno izložili nisu dane u fizikalnom kontekstu.

jedina povezanost s fizikom je oblik funkcije (2.1), koja je istog oblika kao entropija čija definicija proizlazi iz statističke mehanike. Shannon je istaknuo tu prepoznatljivu sličnost navodeći kao primjer funkciju H iz Boltzmannovog H teorema [4].

Zbog potpunosti je važno iznijeti argumente na kojima se temelji Shannova definicija funkcije (2.1) kao mjere nesigurnosti reprezentirane raspodjelom vjerojatnosti $p_1, p_2, \ldots, p_n$. Polazna točka je pitanje koja su svojstva uvjet za konzistentnu definiciju neke funkcije $H(p_1, \ldots, p_n)$ kao mjere nesigurnosti raspodjele vjerojatnosti. Shannon je smatrao da je smisleno zahtijevati slijedeća svojstva [4]:[4]

(1) H je neprekidna funkcija p_i.

(2) Ako su svi p_i jednaki, $p_i = 1/n$, veličina $A(n) = H(1/n, \ldots, 1/n)$ je monotono rastuća funkcija n.

(3) (Zakon slaganja) Ako grupiramo prvih k dogadjaja u događaj vjerojatnosti $w_1 = p_1 + \cdots + p_k$, slijedećih m događaja u događaj vjerojatnosti $w_2 = p_{k+1} + \cdots + p_{k+m}$, itd., i tako grupiranjem svih n događaja dobijemo r složenih događaja čije su vjerojatnosti $w_1, \ldots, w_r$, nesigurnost povezana sa r složenih događaja je $H(w_1, \ldots, w_r)$. Uvjetne vjerojatnosti prvih k događaja uz uvjet da se dogodio prvi složeni događaj su $p_1/w_1, \ldots, p_k/w_1$, itd. Dodatna nesigurnost koja se pojavljuje s vjerojatnošću prvog složenog događaja w_1 jednaka je $H(p_1/w_1, \ldots, p_k/w_1)$, itd. Tada vrijedi

$$\begin{aligned} H(p_1, \ldots, p_n) = H(w_1, \ldots, w_r) \quad &+ \quad w_1 H\left(\frac{p_1}{w_1}, \ldots, \frac{p_k}{w_1}\right) \\ &+ \quad w_2 H\left(\frac{p_{k+1}}{w_2}, \ldots, \frac{p_{k+m}}{w_2}\right) + \ldots \end{aligned}$$

Shannon je potom dokazao slijedeću tvrdnju [4, Teorem 2]:

> Jedina funkcija $H(p_1, \ldots, p_n)$ koja zadovoljava ove tri pretpostavke je funkcija oblika (2.1), gdje je K pozitivna konstanta.

Dokaz teorema prikazat ćemo uz neznatne modifikacije u izvornoj verziji koju je dao Shannon [4, Appendix 2]. Neke detalje preuzet ćemo iz Jaynesove verzije dokaza [25, str. 346–350].

Zbog uvjeta (1) kojim se zahtjeva neprekidnost, funkciju $H(p_1, \ldots, p_n)$ je dovoljno odrediti za sve sumjerljive (racionalne) vjerojatnosti

$$p_i = \frac{d_i}{\sum_{i=1}^{n} d_i} \ , \tag{2.5}$$

[4]Zbog boljeg razumijevanja, Shannonove uvjete dajemo u preciznijem obliku u kojem ih je formulirao Jaynes [1, Appendix A], [25, str. 347-348]

gdje su d_i nenegativni cijeli brojevi. Ako su p_i iracionalni oni se mogu aproksimirati racionalnim brojevima, a isti izraz za $H(p_1, \ldots, p_n)$ mora vrijediti zbog uvjeta neprekidnosti (1).

Za niz m uzastopnih izbora koji se u svakom koraku sastoje od s jednako vjerojatnih mogućnosti, uvjet (3) povezuje nesigurnosti svih mogućih izbora u nizu i nesigurnost izbora od s^m jednako vjerojatnih konačnih događaja. Iz uvjeta (3) u tom slučaju slijedi

$$A(s^m) = mA(s) \ . \tag{2.6}$$

Slično tome vrijedi

$$A(t^n) = nA(t) \ . \tag{2.7}$$

Neka su s i t bilo koja dva cijela broja koji nisu manji od 2. Možemo odabrati n (proizvoljno velik) i naći m takav da vrijedi

$$s^m \leq t^n < s^{m+1} \ . \tag{2.8}$$

Logaritmiranjem i dijeljenjem sa $n \log s$ dobije se

$$\frac{m}{n} \leq \frac{\log t}{\log s} \leq \frac{m+1}{n} \ . \tag{2.9}$$

Oduzimanjem m/n se dobije da za apsolutne iznose vrijedi

$$\left| \frac{m}{n} - \frac{\log t}{\log s} \right| \leq \frac{1}{n} \ , \tag{2.10}$$

Korištenjem uvjeta (2), relacija (2.6), (2.7) i (2.8), dobije se

$$mA(s) \leq nA(t) \leq (m+1)A(s) \ . \tag{2.11}$$

Nadalje, iz uvjeta (3) slijedi da je $A(1) = 0$. Zbog uvjeta (2), tada za svaki cijeli broj $s \geq 2$ mora vrijediti da je $A(s) > 0$. Dijeljenjem (2.11) sa $nA(s)$ slijedi

$$\frac{m}{n} \leq \frac{A(t)}{A(s)} \leq \frac{m+1}{n} \ , \tag{2.12}$$

odnosno

$$\left| \frac{m}{n} - \frac{A(t)}{A(s)} \right| \leq \frac{1}{n} \ . \tag{2.13}$$

Korištenjem nejednakosti $|x_1 + x_2| \leq |x_1| + |x_2|$ zajedno s nejednakostima (2.10) i (2.13), slijedi da je

$$\left| \frac{A(t)}{A(s)} - \frac{\log t}{\log s} \right| \leq \frac{2}{n} \ , \tag{2.14}$$

odnosno

$$\left| \frac{A(t)}{\log t} - \frac{A(s)}{\log s} \right| \leq \frac{2A(s)}{n \log t} \ . \tag{2.15}$$

S obzirom da je n proizvoljno velik,

$$\left| \frac{A(t)}{\log t} - \frac{A(s)}{\log s} \right| \leq \epsilon \ , \tag{2.16}$$

gdje je

$$\epsilon = \frac{2A(s)}{n \log t} \ , \tag{2.17}$$

proizvoljno mali. Za bilo koja dva cijela broja $s, t \geq 2$ vrijedi relacija (2.16), iz koje se dobije

$$A(t) = K \log t \ , \tag{2.18}$$

gdje je K konstanta, koja mora biti pozitivna zbog uvjeta (2). Nadalje, $A(1) = 0$ tako da (2.18) vrijedi za svaki cijeli broj $t \geq 1$.

Odredimo sada $H(p_1, \ldots, p_n)$ za slučaj izbora od n mogućnosti sa sumjerljivim vjerojatnostima koje su dane izrazom (2.5). Promotrimo najprije zaseban slučaj izbora od ukupno $\sum_{i=1}^{n} d_i$ mogućnosti jednakih vjerojatnosti. Taj izbor možemo rastaviti (grupirati) na izbor od n mogućnosti s vjerojatnostima (2.5), i ako je tada izabran i, izbor od d_i mogućnosti s međusobno jednakim vjerojatnostima. Iz uvjeta (3) i jednadžbe (2.18) slijedi da je za ukupni izbor od $\sum_{i=1}^{n} d_i$ mogućnosti jednakih vjerojatnosti nesigurnost jednaka

$$K \log \left(\sum_{i=1}^{n} d_i \right) = H(p_1, \ldots, p_n) + K \sum_{i=1}^{n} p_i \log d_i \ . \tag{2.19}$$

Iz jednadžbe (2.19) slijedi

$$\begin{aligned} H(p_1, \ldots, p_n) &= -K \sum_{i=1}^{n} p_i \log \left(\frac{d_i}{\sum_{j=1}^{n} d_j} \right) \\ &= -K \sum_{i=1}^{n} p_i \log p_i \ . \end{aligned} \tag{2.20}$$

Funkcija $H(p_1, \ldots, p_n)$ je određena za sve sumjerljive vjerojatnosti (2.5). Zbog uvjeta neprekidnosti (1), isti izraz za $H(p_1, \ldots, p_n)$ mora vrijediti i za iracionalne p_i.

Kao osvrt je možda najbolje navesti Shannonov [4] komentar teorema: “Ovaj teorem, i pretpostavke koje su potrebne za njegov dokaz, nisu ni na koji način nužne za postojeću teoriju. On je prevenstveno dan da pruži određenu uvjerljivost nekim od naših kasnijih definicija. Stvarno opravdanje ovih definicija, međutim, nalazit će se u njihovim implikacijama.”

Shannonovim teoremom specificirana su određena svojstva koja su dovoljna za jednoznačno određivanje funkcije (2.1). To su svojstva za koja intuitivno očekujemo da ih funkcija koju koristimo kao mjeru nesigurnosti reprezentirane raspodjelom vjerojatnosti mora zadovoljavati. Ta svojstva nismo dodatno argumentirali; Shannon ih je nazvao smislenim, razložnim, ali ih, osim njihovog preciznog formuliranja, nije posebno opravdao, vjerojatno jer je smatrao da je neophodnost tih svojstava evidentna. Shannonov teorem argument je za primjenu funkcije (2.1) u teoriji informacije i u brojnim drugim područjima, ali nije dokaz da je naša definicija mjere nesigurnosti, ili mjere informacije konzistentna. Ipak, u implikacijama koje proizlaze takvom definicijom, te u brojnim rezultatima primjene u različitim područjima povezanim teorijom vjerojatnosti, nisu pronađene nikakve nekonzistentnosti, što također predstavlja važan argument [25].

Funkcija (2.1) ima dodatna, također važna svojstva koja su detaljno raspravljena u referencama [4], [27], [28]. Definicija funkcije (2.1) kao mjere nesigurnosti reprezentirane raspodjelom vjerojatnosti je konzistentna s tim svojstvima koja očekujemo od mjere nesigurnosti. Naprimjer, funkcija (2.1) ima maksimum kad su sve vjerojatnosti jednake (relacija (2.2)). Funkcija (2.1) kao mjera nesigurnosti tako zadovoljava intuitivno očekivanje da je nesigurnost najveća kad su svi događaji jednako vjerojatni. Dolje navodimo neka od ostalih važnih svojstava fukcije (2.1) u obliku u kojem su dana u Shannonovom radu [4] (dokazi se mogu pronaći i u drugim prethodno navedenim referencama).

(1) $H = 0$ ako i samo ako su sve vjerojatnosti p_i jednake nuli, osim jedne koja je jednaka jedinici. Nesigurnost je jednaka nuli kad je jedan od događaja siguran.

(2) Za bilo koji dani n, H je maksimalan i jednak je $K \log n$ kad je svih n vjerojatnosti jednako, $p_i = 1/n$, $i = 1, 2, \ldots, n$.

(3) Za diskretne varijable x i y, koje mogu proprimiti vrijednosti $x = i$ $(i = 1, 2, \ldots, m)$ i $y = j$ $(j = 1, 2, \ldots, n)$, vjerojatnost zajedničkog pojavljivanja $x = i$ i $y = j$ je $p(i, j)$. Nesigurnost vezana uz pojavu zajedničkog događaja je

$$H(x, y) = -K \sum_{i=1}^{m} \sum_{j=1}^{n} p(i, j) \log p(i, j) \ ,$$

dok su pojedinačne nesigurnosti

$$H(x) = -K \sum_{i=1}^{m} \sum_{j=1}^{n} p(i, j) \log \left(\sum_{j=1}^{n} p(i, j) \right) \ ,$$

$$H(y) = -K \sum_{i=1}^{m} \sum_{j=1}^{n} p(i, j) \log \left(\sum_{i=1}^{m} p(i, j) \right) \ ,$$

Tada vrijedi

$$H(x, y) \leq H(x) + H(y) \ , \tag{2.21}$$

s jednakošću ako i samo ako su događaji nezavisni, $p(i, j) = p(i)p(j)$. Nesigurnost zajedničkog događaja je manja od ili je jednaka zbroju pojednačnih nesigurnosti.

(4) Ako se vjerojatnosti p_i $(i = 1, 2, \ldots, n)$ usrednje u obliku

$$p_i' = \sum_{j=1}^{n} a_{ij} p_j \ ,$$

gdje je $\sum_{i=1}^{n} a_{ij} = \sum_{j=1}^{n} a_{ij} = 1$ i $a_{ij} \geq 0$ za sve $i, j = 1, 2 \ldots, n$, tada H raste (osim u posebnom slučaju kada ta transformacija odgovara permutaciji vjerojatnosti p_i; H tada ostaje isti).

(5) Za diskretne varijable x i y, koje mogu proprimiti vrijednosti $x = i$ $(i = 1, 2, \ldots, m)$ i $y = j$ $(j = 1, 2, \ldots, n)$, uvjetna vjerojatnost za $y = j$ ako je poznato da je $x = i$ je jednaka

$$p_i(j) = \frac{p(i, j)}{\sum_{j=1}^{n} p(i, j)} \ .$$

Uvjetna nesigurnost y je prosjek nesigurnosti y uz poznate x, uzet po svim vjerojatnostima x,

$$H_x(y) = -K \sum_{i=1}^{m} \sum_{j=1}^{n} p(i, j) \log p_i(j) \ .$$

Tada vrijedi

$$H(x, y) = H(x) + H_x(y) \ ,$$

Nesigurnost zajedničkog događaja je jednaka zbroju nesigurnosti prvog i nesigurnosti drugog događaja kad je prvi poznat.

(6) Iz (3) i (5) slijedi

$$H(x) + H(y) \geq H(x, y) = H(x) + H_x(y) \ ,$$

odnosno

$$H(y) \geq H_x(y) \ ,$$

s jednakošću ako i samo ako su događaji nezavisni. Ako je poznat prvi događaj, nesigurnost drugog događaja smanji se ili ostane ista.

2.2 Aksiomatski okvir za konstrukciju modela u teoriji vjerojatnosti

Prvi korak konstrukcije matematičkog modela u teoriji vjerojatnosti je uvođenje nepraznog skupa Ω (prostor elementarnih događaja) i familije $\mathcal{A}$ podskupova od Ω takve da vrijedi:

(1) $\emptyset \in \mathcal{A}$,

(2) $A \in \mathcal{A} \Longrightarrow A^c = \Omega \setminus A \in \mathcal{A}$,

(3) $A_n \in \mathcal{A}, n \in \mathbb{N} \Longrightarrow \bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$.

Tada se $\mathcal{A}$ naziva σ-algebra na Ω. Uređen par $(\Omega, \mathcal{A})$ poznat je kao izmjeriv prostor. Elemente od $\mathcal{A}$ nazivamo izmjerivim skupovima. Najmanja σ-algebra na Ω je $\mathcal{A} = \{\emptyset, \Omega\}$, a najveća je partitivni skup 2^Ω (skup svih podskupova od Ω). Iz definicije σ-algebre slijedi tvrdnja da je familija skupova koji su zajednički za dvije σ-algebre na Ω također σ-algebra na Ω. Dakle, presjek dvije σ-algebre na Ω je σ-algebra na Ω. Daljne proširenje je tvrdnja da je presjek proizvoljne familije σ-algebri $\bigcap_{i\in I} \mathcal{A}_i$ također σ-algebra. Za bilo koju familiju podskupova $\mathcal{A}_0 \subset 2^\Omega$ postoji jedinstvena "najmanja" σ-algebra $\mathcal{S} \supseteq \mathcal{A}_0$. To je presjek svih σ-algebri na Ω koje sadrže $\mathcal{A}_0$, a naziva se σ-algebra generirana sa $\mathcal{A}_0$.

Slijedeći korak je uvođenje vjerojatnosne mjere na izmjerivom prostoru $(\Omega, \mathcal{A})$. Mjera μ na izmjerivom prostoru $(\Omega, \mathcal{A})$ je funkcija $\mu : \mathcal{A} \to \overline{\mathbb{R}}$ takva da vrijedi:

(1) $\mu(A) \geq 0$ za sve $A \in \mathcal{A}$, i $\mu(\emptyset) = 0$.

(2) Ako je $(A_n)_{n\in\mathbb{N}}$ niz međusobno disjunktnih skupova $A_n \in \mathcal{A}$, tada je

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n) \ .$$

Prostor mjere je trojka $(\Omega, \mathcal{A}, \mu)$ koja se sastoji od izmjerivog prostora $(\Omega, \mathcal{A})$ i mjere μ. Svojstvo (2) funkcije μ naziva se σ-aditivnost ili prebrojiva aditivnost.

Prostor mjere $(\Omega, \mathcal{A}, \mathrm{P})$ takav da vrijedi $\mathrm{P}(\Omega) = 1$ naziva se vjerojatnosni prostor, a mjera P naziva se vjerojatnosna mjera ili vjerojatnost na $(\Omega, \mathcal{A})$. Vjerojatnost na $(\Omega, \mathcal{A})$ je funkcija P definirana na $\mathcal{A}$ koja svakom $A \in \mathcal{A}$ pridružuje realan broj $\mathrm{P}(A)$ tako da vrijedi:

(1) $0 \leq \mathrm{P}(A) \leq 1$ za svaki $A \in \mathcal{A}$.

(2) $\mathrm{P}(\Omega) = 1$.

(3) Ako su $A_i \in \mathcal{A}$, $i \in \mathbb{N}$, međusobno disjunktni skupovi, tada je

$$\mathrm{P}\left(\bigcup_{i\in\mathbb{N}} A_i\right) = \sum_{i\in\mathbb{N}} \mathrm{P}(A_i) \ .$$

Ako je $(\Omega, \mathcal{A}, \mathrm{P})$ vjerojatnosni prostor i $A \in \mathcal{A}$, tada se skup A naziva događaj, a broj $\mathrm{P}(A)$ naziva se vjerojatnost događaja A. Svi događaji $A \in \mathcal{A}$ imaju vjerojatnosti u intervalu $0 \leq \mathrm{P}(A) \leq 1$, $\mathrm{P}(\emptyset) = 0$. Svojstva vjerojatnosti P formulirana su u okviru teorije mjere i uzimaju se za aksiome teorije vjerojatnosti. Ostala svojstva vjerojatnosti su posljedice tih aksioma.

Vjerojatnosni prostor $(\Omega, \mathcal{A}, \mathrm{P})$ naziva se potpun vjerojatnosni prostor ako $B \subseteq A \in \mathcal{A}$ i $\mathrm{P}(A) = 0$ implicira $B \in \mathcal{A}$ i $\mathrm{P}(B) = 0$. U slučaju kada je prostor elementarnih događaja Ω konačan ili prebrojiv skup, za familiju događaja uzima se najveća σ-algebra $\mathcal{A} = 2^{\Omega}$, a vjerojatnosni prostor $(\Omega, 2^{\Omega}, \mathrm{P})$ se tada naziva diskretni vjerojatnosni prostor. Na konačnom ili prebrojivom prostoru elementarnih događaja ($\Omega = \{\omega_1, \ldots, \omega_n\}$ ili $\Omega = \{\omega_i : i \in \mathbb{N}\}$) vjerojatnosti elementarnih događaja $\omega_i \in \Omega$ zadane su brojevima $\mathrm{P}(\{\omega_i\}) = p_i$ takvim da vrijedi

$$0 \leq p_i \leq 1, \qquad \sum_i p_i = 1 \ . \tag{2.22}$$

Tada je za $A \subseteq \Omega$ vjerojatnost $\mathrm{P}(A)$ jednaka

$$\mathrm{P}(A) = \sum_{\{i:\ \omega_i \in A\}} p_i \ . \tag{2.23}$$

S obzirom da $B \subseteq \Omega$ implicira da je $B \in 2^{\Omega}$, diskretni vjerojatnosni prostor $(\Omega, 2^{\Omega}, \mathrm{P})$ je potpun vjerojatnosni prostor. Potpunost je posljedica izbora najveće σ-algebre na Ω, partitivnog skupa 2^{Ω}, kao familije događaja.

Moguće je uvesti konačnu ili prebrojivu familiju $\{E_i : i \in I\}$ nepraznih, međusobno disjunktnih podskupova od Ω, takvu da vrijedi

$$\Omega = \bigcup_{i \in I} E_i \ . \tag{2.24}$$

Tada se familija $\{E_i : i \in I\}$ naziva particija od Ω ili potpun sustav događaja na Ω. Tom particijom od Ω je generirana najmanja σ-algebra $\mathcal{E}$ na Ω koja je sadrži, ali to ne implicira da je $(\Omega, \mathcal{E}, \mathrm{P})$ potpun vjerojatnosni prostor. Je li vjerojatnosni prostor $(\Omega, \mathcal{E}, \mathrm{P})$ potpun ili nije, ovisi o σ-algebri $\mathcal{E}$ i o vjerojatnosnoj mjeri P. Ako je prostor elementarnih događaja Ω konačan ili prebrojiv skup, tada se ovaj problem može vrlo jednostavno riješiti uvođenjem najveće σ-algebre na Ω, skupa 2^{Ω} svih podskupova od Ω kao familije događaja i zadavanjem vjerojatnosti za sve elementarne događaje $\omega_i \in \Omega$.

Općenito, u teoriji mjere, problemi slične vrste se rješavaju uvođenjem vanjske mjere m^*. Funkcija $m^* : 2^{\Omega} \to [0, +\infty]$ se naziva vanjska mjera na 2^{Ω} ako vrijedi:

(1) $m^*(\emptyset) = 0$.

(2) $A \subseteq B \subseteq \Omega \Longrightarrow m^*(A) \leq m^*(B)$.

(3) $A_n \subseteq \Omega, n \in \mathbb{N} \Longrightarrow m^*\left(\bigcup_{n \in \mathbb{N}} A_n\right) \leq \sum_{n \in \mathbb{N}} m^*(A_n)$.

Umjesto prebrojive aditivnosti funkcije mjere za međusobno disjunktne skupove iz $\mathcal{A}$, vanjska mjera m^* na 2^Ω ima svojstvo prebrojive subaditivnosti (3) za sve podskupove od Ω, disjunktne i nedisjunktne. Podskup $B \subseteq \Omega$ naziva se m^* izmjeriv ako je za svaki $A \subseteq \Omega$,

$$m^*(A) = m^*(A \cap B) + m^*(A \cap B^c) \ . \tag{2.25}$$

Tada se mogu dokazati slijedeće tvrdnje. Familija $\mathcal{M}_{m^*}$ svih m^* izmjerivih podskupova od Ω je σ-algebra na Ω, vanjska mjera $m^*|_{\mathcal{M}_{m^*}}$ je mjera na $(\Omega, \mathcal{M}_{m^*})$, a $(\Omega, \mathcal{M}_{m^*}, m^*|_{\mathcal{M}_{m^*}})$ je prostor mjere [29][Teorem 2].

Neka je $(\Omega, \mathcal{A}, \mu)$ prostor mjere. Definirajmo funkciju $\mu^* : 2^\Omega \to [0, +\infty]$:

$$\mu^*(B) = \inf \left\{ \sum_{i \in \mathbb{N}} \mu(A_i) : \{A_i\}_{i \in \mathbb{N}} \subset \mathcal{A}, B \subseteq \bigcup_{i \in \mathbb{N}} A_i \right\} \ . \tag{2.26}$$

Funkcija μ^* je vanjska mjera na 2^Ω jer zadovoljava svojstva (1), (2) i (3). To je posljedica općenitijeg teorema [29][Teorem 6]. σ-algebra $\mathcal{A}$ i svi podskupovi $B \subseteq A \in \mathcal{A}$ takvi da vrijedi $\mu(A) = 0$, generiraju najmanju σ-algebru $\mathcal{A}^*$ koja ih sadrži. Vanjska mjera μ^* određena iz (2.26) je ekstenzija mjere μ na $\mathcal{A}^*$ sa slijedećim svojstvima:

(1) $\mu^*|_{\mathcal{A}} = \mu$.

(2) $\mu^*(B) = 0$, za sve $B \subseteq A \in \mathcal{A}$ takve da vrijedi $\mu(A) = 0$.

(3) $\mu^*(C \cup B) = \mu(C)$, za sve $C \in \mathcal{A}$ i $B \subseteq A \in \mathcal{A}$ takve da je $\mu(A) = 0$.

S obzirom da je $\mathcal{A}^* \subseteq \mathcal{M}_{\mu^*}$ ([29][Teorem 6]), tada je prema gore navedenoj tvrdnji vanjska mjera μ^* mjera i na $(\Omega, \mathcal{A}^*)$. Cijeli postupak se naziva upotpunjenje jer je $(\Omega, \mathcal{A}^*, \mu^*|_{\mathcal{A}^*})$ potpun prostor mjere. To implicira da je i $(\Omega, \mathcal{M}_{\mu^*}, \mu^*|_{\mathcal{M}_{\mu^*}})$ potpun prostor mjere. Važan primjer je Lebesgueova vanjska mjera, koja omogućuje ekstenziju Lebesgueove mjere sa σ-algebre generirane n-dimenzionalnim intervalima na sve ostale Lebesgue izmjerive skupove u $\mathbb{R}^n$ (Lebesgueova σ-algebra na $\mathbb{R}^n$).

2.3 Načelo najveće informacijske entropije

Interpretacija funkcije (2.1) kao iznosa informacije, ili srednjeg iznosa informacije u komunikacijskom sustavu, proizišla je u kontekstu teorije komunikacije. Funkcija (2.1) ima i drugačije interpretacije, koje su važnije za druga područja. U odjeljku 2.1 iznesene su interpretacije koje su kompatibilne s općenitom definicijom funkcije (2.1) kao mjere nesigurnosti reprezentirane raspodjelom vjerojatnosti. Pokazano je da se ta općenita definicija može u posebnim kontekstima na različite načine interpretirati kao mjera informacije koja nam nedostaje za određivanje događaja. Tu je važno napomenuti da riječ "događaj" označava pojam koji ima različite interpretacije, iako u teoriji vjerojatnosti, kao što je prikazano u odjeljku 2.2, taj pojam ima vrlo preciznu matematičku definiciju. Događaji su u teoriji vjerojatnosti podskupovi prostora (skupa) elementarnih događaja. Na taj način možemo događaje matematički prikazivati kao skupove elementarnih događaja.

Aksiomatski okvir teorije vjerojatnosti ne sadrži nikakvo dodatno objašnjenje apstraktne matematičke definicije događaja, niti objašnjava pojam vjerojatnosti u potpunosti. U tom formalnom okviru dana je općenita definicija vjerojatnosti na konačnim, prebrojivim i neprebrojivim prostorima elementarnih događaja. Svojstva koja takvu definiciju vjerojatnosti omogućuju postulirana su u obliku aksioma, ali time nije specificiran odgovor na pitanje kako odrediti, ili na koji način konstruirati vjerojatnosti. Jednom kada su vjerojatnosti zadane u skladu s aksiomima, posljedice tih aksioma omogućuju izračun vjerojatnosti složenijih događaja. S druge strane, određivanje inicijalnih vjerojatnosti ovisi o interpretaciji vjerojatnosti u pojedinačnim slučajevima koji nas zanimaju.

Iz tog razloga je razumljivo da definicija funkcije vjerojatnosti (2.1) kao mjere nesigurnosti ne objašnjava u potpunosti značenje pojma nesigurnost. Preciznije, moglo bi se reći da je razlog u tome što interpretacije te definicije ne generiraju kontradikcije, nego ovisno o kontekstu, upotpunjuju značenje samog pojma. Neovisnost o kontekstu je argument za valjanost općenite definicije. To je važno i zato što je jasno da interpretacija funkcije vjerojatnosti (2.1) u velikoj mjeri ovisi o interpretaciji samih vjerojatnosti. Osim toga, možemo se upitati, koje su posljedice definicije mjere nesigurnosti za teoriju vjerojatnosti.

U odjeljku 2.2 je pokazano da su, pri konstrukciji diskretne raspodjele vjerojatnosti, uvjeti (2.22) jedini kojih se moramo pridržavati da bi raspodjela bila u skladu s aksiomima teorije vjerojatnosti. Naravno, uvjeti (2.22) sami za sebe nisu dovoljni za jedinstveno određivanje raspodjele vjerojatnosti. Jaynes [1, 2] je tražeći rješenje problema konstrukcije raspodjele vjerojatnosti formulirao *načelo najveće informacijske entropije* kao općeniti kriterij za odabir raspodjele vjerojatnosti kada dostupna informacija nije dovoljna za njeno jedinstveno određivanje.[5]

Načelo najveće informacijske entropije zasniva se na općenitoj pretpostavci da vjerojatnosti reprezentiraju stupanj znanja o problemu na kojeg se odnose [1, 5, 6, 25]. Ako prihvatimo ovu pretpostavku, tada je prirodno postaviti pitanje kako konstruirati raspodjelu koja ovisi isključivo o dostupnoj informaciji, jer jedino takva raspodjela reprezentira stvarni stupanj znanja, odnosno poznavanja danog problema. U idealnom slučaju to ne podrazumijeva naše subjektivno znanje o problemu; u problemima koji se proučavaju u fizici pojam znanje podrazumijeva ukupno teorijsko i eksperimentalno znanje o nekom

[5] Umjesto naziva entropija koji se koristi u teoriji informacije, ovdje se za Shannonovu mjeru nesigurnosti (2.1) dosljedno upotrebljava naziv *informacijska entropija*, koji se najčešće koristi u relaciji s fizikom, zbog potrebe razlikovanja s termodinamičkom entropijom. Vrlo često se u literaturi informacijska entropija naziva još i mjerom informacije koja nedostaje.

problemu.

Definicija mjere nesigurnosti (2.1), uz spomenutu pretpostavku o vjerojatnostima, vodi nas do općenitog rješenja koje je dao Jaynes [1, 2]. Izbor raspodjele vjerojatnosti je ograničen informacijom koja je dostupna. Najčešće su to podaci u obliku očekivanih vrijednosti. Od svih raspodjela koje zadovoljavaju zadana ograničenja, odabiremo raspodjelu vjerojatnosti s najvećom nesigurnošću povezanu s informacijom koja nedostaje. Odabirom te raspodjele vjerojatnosti, uzimamo kao sigurnu jedino informaciju koja je reprezentirana zadanim ograničenjima. Na taj način ne prepostavljamo nikakvu dodatnu informaciju koja nam u stvarnosti nije dostupna, pa odabrana raspodjela vjerojatnosti *jedina reprezentira* stvarni stupanj znanja. U idealnom slučaju to znanje je relevantno za problem koji nas zanima.

Opisana metoda konstrukcije raspodjele vjerojatnosti primjenom načela najveće informacijske entropije često se naziva MaxEnt algoritam. Ovdje je prikazujemo na jednostavnom primjeru diskretne raspodjele vjerojatnosti.

Neka je x varijabla koja može poprimiti n različitih vrijednosti $\{x_1, \ldots, x_n\}$, a $f_k(x)$ su funkcije od x, gdje je $k = 1, 2, \ldots, m < n$. Želimo naći odgovarajuću raspodjelu vjerojatnosti $\{p_1, \ldots, p_n\}$ za koju je, uz ograničenja zadana očekivanim vrijednostima

$$F_k = \langle f_k(x) \rangle = \sum_{i=1}^{n} p_i f_k(x_i) \ , \qquad k = 1, 2, \ldots, m < n \ , \tag{2.27}$$

i uvjetom

$$\sum_{i=1}^{n} p_i = 1 \ , \qquad\qquad (p_i \geq 0, \quad i = 1, 2, \ldots, n) \ , \tag{2.28}$$

nesigurnost, odnosno informacijska entropija S_I,

$$S_I = -\sum_{i=1}^{n} p_i \log p_i \ , \tag{2.29}$$

maksimalna. Ovdje pozitivnu konstantu K izostavljamo, jer njen iznos ovisi jedino o izboru jedinice mjere u izrazu (2.1).

Pretpostavimo da je sustav koji je zadan sa $m = n - 1$ jednadžbi (2.27) i jednadžbom (2.28) kompatibilan.[6] Tada zadani sustav jednadžbi (2.27) i (2.28) ima jedinstveno rješenje $\{p_1, \ldots, p_n\}$ ako i samo ako su sve jednadžbe sustava linearno nezavisne. Tim rješenjem određena je vrijednost informacijske entropije (2.29). Ako je ukupan broj jednadžbi (2.27) takav da vrijedi $m < n - 1$, tada sustav jednadžbi (2.27) i (2.28) nema jedinstveno rješenje. U golemoj većini slučajeva koji nas u stvarnosti zanimaju, dostupna

[6]Sustav linearnih jednadžbi se naziva kompatibilan ako ima barem jedno rješenje. Takav sustav se još naziva rješivim, mogućim, konzistentnim i sl.

informacija ukupno je daleko manja od dovoljne za jedinstveno određivanje vjerojatnosti, tj. gotovo uvijek je $m \ll n-1$. Tada vjerojatnosti $\{p_1, \ldots, p_n\}$ određujemo primjenom načela najveće informacijske entropije.

Raspodjelu vjerojatnosti $\{p_1, \ldots, p_n\}$ za koju je informacijska entropija (2.29) maksimalna uz ograničenja (2.27), najlakše je naći primjenom metode Lagrangeovih multiplikatora. Da bi našli maksimum funkcije (2.29) uz uvjete (2.27) i (2.28), formiramo funkciju

$$I = -\sum_{i=1}^{n} p_i \log p_i - (\lambda_0 - 1)\left(\sum_{i=1}^{n} p_i - 1\right) - \sum_{k=1}^{m} \lambda_k \left(\sum_{i=1}^{n} p_i f_k(x_i) - F_k\right) , \tag{2.30}$$

gdje su $\lambda_0 - 1$, λ_k, $k = 1, 2, \ldots, m$, Lagrangeovi multiplikatori, i zatim tražimo običan maksimum ove funkcije I. Rješavanjem jednažbi

$$\frac{\partial I}{\partial p_i} = -\log p_i - \lambda_0 - \sum_{k=1}^{m} \lambda_k f_k(x_i) = 0 , \qquad i = 1, 2, \ldots, n , \tag{2.31}$$

dobije se

$$p_i = \exp\left\{-\lambda_0 - \sum_{k=1}^{m} \lambda_k f_k(x_i)\right\} , \qquad i = 1, 2, \ldots, n . \tag{2.32}$$

Tada iz uvjeta (2.27) i (2.28) možemo odrediti Lagrangeove multiplikatore λ_0, λ_k, $k = 1, 2, \ldots, m$. Iz jednadžbi (2.28) i (2.32) slijedi

$$\lambda_0 = \log\left\{\sum_{i=1}^{n} \exp\left[-\sum_{k=1}^{m} \lambda_k f_k(x_i)\right]\right\} . \tag{2.33}$$

Normalizacijski faktor e^{λ_0} raspodjele vjerojatnosti (2.32) jednak je

$$Z \equiv Z(\lambda_1, \ldots, \lambda_m) = \sum_{i=1}^{n} \exp\left\{-\sum_{k=1}^{m} \lambda_k f_k(x_i)\right\} . \tag{2.34}$$

U MaxEnt formalizmu normalizacijski faktor (2.34) se naziva *particijska funkcija*. Korištenjem jednadžbi (2.32), (2.33) i (2.34) dobije se da je MaxEnt raspodjela vjerojatnosti jednaka

$$p_i = \frac{1}{Z} \exp\left\{-\sum_{k=1}^{m} \lambda_k f_k(x_i)\right\} , \qquad i = 1, 2, \ldots, n . \tag{2.35}$$

Očekivane vrijednosti funkcija $\langle f_k(x)\rangle = F_k$, $k = 1, 2, \ldots, m$, dane su uvjetima (2.27), ili ekvivalentno, jednadžbama

$$F_k = \langle f_k(x)\rangle = -\frac{\partial \log Z(\lambda_1, \ldots, \lambda_m)}{\partial \lambda_k} , \qquad k = 1, 2, \ldots, m . \tag{2.36}$$

Rješavanjem jednadžbi (2.36) mogu se odrediti Lagrangeovi multiplikatori λ_k, $k = 1, 2, \ldots, m$, kao funkcije očekivanih vrijednosti.

Uvrštavanjem MaxEnt raspodjele vjerojatnosti (2.35) u izraz (2.29) dobije se maksimum informacijske entropije S_I uz uvjete (2.27) i (2.28),

$$(S_I)_{\max} = \log Z(\lambda_1, \dots, \lambda_m) + \sum_{k=1}^{m} \lambda_k \langle f_k(x) \rangle \ . \tag{2.37}$$

Jednostavno je pokazati da je (2.37) najveća vrijednost S_I uz uvjete (2.27) i (2.28). Dokaz koji ovdje dajemo preuzet je iz referenci [5, 13, 25] uz manje modifikacije.

Neka su $\{p_1, \dots, p_n\}$ i $\{q_1, \dots, q_n\}$ dvije raspodjele vjerojatnosti koje zadovoljavaju uvjete (2.27) i (2.28). Za $x \geq 0$ vrijedi relacija

$$\log x \leq x - 1 \ , \tag{2.38}$$

s jednakošću ako i samo ako je $x = 1$. Primjenom relacije (2.38) se dobije

$$\sum_{i=1}^{n} q_i \log \left(\frac{p_i}{q_i} \right) \leq \sum_{i=1}^{n} q_i \left(\frac{p_i}{q_i} - 1 \right) = 0 \ , \tag{2.39}$$

s jednakošću ako i samo ako je $q_i = p_i$, za sve $i = 1, 2, \dots, n$. Iz (2.39) dalje slijedi

$$-\sum_{i=1}^{n} q_i \log q_i \leq -\sum_{i=1}^{n} q_i \log p_i \ . \tag{2.40}$$

Ako se za raspodjelu vjerojatnosti $\{p_1, \dots, p_n\}$ odabere raspodjela (2.35), relacija (2.40) daje

$$-\sum_{i=1}^{n} q_i \log q_i \leq \sum_{i=1}^{n} q_i \left[\log Z + \sum_{k=1}^{m} \lambda_k f_k(x_i) \right] \ . \tag{2.41}$$

Raspodjela vjerojatnosti $\{q_1, \dots, q_n\}$ zadovoljava iste uvjete (2.27) i (2.28) kao i raspodjela (2.35). Uzimajući to u obzir, korištenjem relacija (2.29), (2.37) i (2.41) dobije se

$$S_I(q_1, \dots, q_n) \leq \log Z + \sum_{k=1}^{m} \lambda_k \langle f_k(x) \rangle = (S_I)_{\max} \ , \tag{2.42}$$

s jednakošću ako i samo ako je raspodjela $\{q_1, \dots, q_n\}$ jednaka raspodjeli (2.35).

Dakle, dokazana je tvrdnja da na skupu svih raspodjela koje zadovoljavaju uvjete (2.27) i (2.28), informacijska entropija S_I poprima svoju najveću vrijednost jedino za MaxEnt raspodjelu vjerojatnosti (2.35). Time je dokazano da je primjena bilo koje druge raspodjele u suprotnosti s načelom najveće informacijske entropije. MaxEnt raspodjela (2.35) nađena je standarnim postupkom uvjetne maksimizacije, ali dokaz ovih tvrdnji je dan nezavisnim postupkom.

Posljedica ili korolar ovih tvrdnji je to da su Lagrangeovi multiplikatori λ_0, λ_k, $k = 1, 2, \dots, m$, jedinstveno određeni uvjetima (2.27) i (2.28). Dokaz se u skiciranom obliku može pronaći u referenci [13, str. 29-30]. Ovdje taj dokaz malo detaljnije prikazujemo; razlog za to je važna uloga Lagrangeovih multiplikatora u MaxEnt formalizmu.

Pretpostavimo da postoje neke druge vrijednosti λ'_0, λ'_k, $k = 1, 2, \ldots, m$, za koje su jednadžbe (2.33) i (2.36) također zadovoljene. Tim vrijednostima odgovara raspodjela vjerojatnosti

$$q_i = \frac{1}{Z(\lambda'_1, \ldots, \lambda'_m)} \exp\left\{ -\sum_{k=1}^{m} \lambda'_k f_k(x_i) \right\} , \qquad i = 1, 2, \ldots, n \ . \tag{2.43}$$

koja mora zadovoljavati nejednakost (2.42). Zamjenom mjesta raspodjela (2.35) i (2.43) u relaciji (2.40), a onda i u relacijama (2.41) i (2.42), dobije se

$$S_I(p_1, \ldots, p_n) \leq \log Z(\lambda'_1, \ldots, \lambda'_m) + \sum_{k=1}^{m} \lambda'_k \langle f_k(x) \rangle \ . \tag{2.44}$$

Relacije (2.42) i (2.44) impliciraju da za raspodjele (2.35) i (2.43) vrijedi $S_I(q_1, \ldots, q_n) = S_I(p_1, \ldots, p_n) = (S_I)_{\max}$.

Jednakosti u relacijama (2.42) i (2.44) vrijede ako i samo ako su raspodjele jednake, $q_i = p_i$, za sve $i = 1, 2, \ldots, n$. Iz jednakosti raspodjela (2.35) i (2.43) dalje slijedi da vrijede jednadžbe

$$\sum_{k=1}^{m} (\lambda_k - \lambda'_k) f_k(x_i) = \sum_{k=1}^{m} (\lambda_k - \lambda'_k) f_k(x_l) \ , \qquad i, l = 1, 2, \ldots, n \ . \tag{2.45}$$

Pretpostavimo da je sustav koji se sastoji od m jednadžbi (2.27) i jednadžbe (2.28) kompatibilan i da su jednadžbe sustava linearno nezavisne. Od koeficijenata uz nepoznanice u jednadžbama (2.27) formiramo $n \times m$ matricu $\mathbf{A}$ sa elementima $(\mathbf{A})_{ik} = f_k(x_i)$. Iz početne pretpostavke o kompatibilnosti i linearnoj nezavisnosti jednadžbi (2.27) i (2.28) slijedi da je $m \leq n - 1$, odnosno $m < n$. Rang matrice definiran je kao maksimalan broj linearno nezavisnih stupaca matrice. Maksimalan broj linearno nezavisnih redaka matrice jednak je rangu matrice. Prethodne tvrdnje impliciraju da je matrica $\mathbf{A}$ ranga m. To dalje implicira da je broj linearno nezavisnih redaka matrice $\mathbf{A}$ jednak m.

Odabirom m linearno nezavisnih redaka matrice $\mathbf{A}$ formiramo $m \times m$ matricu $\mathbf{B}$ na slijedeći način. Elementi matrice $\mathbf{B}$ su dani sa $(\mathbf{B})_{jk} = f_k(x_{i(j)}) - f_k(x_{i(m+1)})$, gdje je funkcijom $i(j)$, $j = 1, \ldots, m$, označeno m odabranih linearno nezavisnih redaka matrice $\mathbf{A}$. Redak matrice $\mathbf{A}$ koji sadrži elemente $(\mathbf{A})_{i(m+1),k} = f_k(x_{i(m+1)})$ može se zapisati kao linearna kombinacija m prethodno odabranih linearno nezavisnih redaka.

Jednadžbe (2.45) se za $i = i(j)$, $j = 1, 2, \ldots, m$, i $l = i(m+1)$ mogu pomoću matrice $\mathbf{B}$ zapisati u ekvivalentnom obliku

$$\sum_{k=1}^{m} (\mathbf{B})_{jk} (\lambda_k - \lambda'_k) = 0 \ , \qquad j = 1, 2, \ldots, m \ . \tag{2.46}$$

Jednostavno je pokazati da je $\det \mathbf{B} \neq 0$. Tvrdnja da je $\det \mathbf{B} \neq 0$ je ekvivalentna tvrdnji da su svi redci (i svi stupci) $m \times m$ matrice $\mathbf{B}$ linearno nezavisni. Ekvivalentno ovim tvrdnjama, matrične jednadžbe (2.46) imaju samo trivijalno rješenje. To dalje implicira da je $\lambda'_k = \lambda_k$, za sve $k = 1, 2, \ldots, m$. Jednadžba (2.33) tada daje $\lambda'_0 = \lambda_0$.

Time je dokazana tvrdnja da su Lagrangeovi multiplikatori λ_k, $k = 1, 2, \ldots, m$, jedinstveno određeni uvjetima (2.27) i (2.28). Jednadžbom (2.34) određen je e^{λ_0} kao funkcija koja je faktor normalizacije u MaxEnt raspodjeli vjerojatnosti. Važno je napomenuti da se u dokazu koristi pretpostavka o kompatibilnosti i linearnoj nezavisnosti jednadžbi sustava zadanog sa (2.27) i (2.28). Ako su funkcije $f_k(x)$ i njihove očekivane vrijednosti F_k takve da sustav jednadžbi (2.27) i (2.28) zadovoljava ove pretpostavke, tada za svaki takav skup $m < n-1$ funkcija i njihovih očekivanih vrijednosti $\{(f_1(x), F_1), \ldots, (f_m(x), F_m)\}$ postoji *jedna i samo jedna* m-torka $(\lambda_1, \ldots, \lambda_m)$. Tada je dozvoljeno reći da su za takav skup funkcija $\{f_1(x), \ldots, f_m(x)\}$ Lagrangeovi multiplikatori λ_k, $k = 1, 2, \ldots, m$, *jednoznačne funkcije* $\lambda_k(F_1, \ldots, F_m)$ očekivanih vrijednosti $F = (F_1, \ldots, F_m)$. Na taj način je definirana jednoznačna funkcija

$$\lambda(F) \equiv (\lambda_1(F), \ldots, \lambda_m(F)) \ , \tag{2.47}$$

za sve $F = (F_1, \ldots, F_m)$ iz domene $\mathcal{D}(\lambda) \subseteq \mathbb{R}^m$ određene pretpostavkom o kompatibilnosti.

Naravno, sam problem određivanja domene $\mathcal{D}(\lambda)$ funkcije λ kao i njene slike $\mathcal{R}(\lambda) \subseteq \mathbb{R}^m$ nije time rješen. Općenito, za funkciju $\varphi : X \to Y$ postoji inverzna funkcija $\varphi^{-1} : Y \to X$ ako i samo ako je φ bijekcija. Pretpostavimo da postoji podskup $\mathcal{S} \subseteq \mathcal{D}(\lambda)$ takav da je restrikcija funkcije λ na $\mathcal{S}$, koju ćemo označiti sa $\lambda|_{\mathcal{S}} : \mathcal{S} \to \mathcal{R}(\lambda)$, injektivna funkcija. Ako za kodomenu uzmemo njenu sliku $\mathcal{R}(\lambda|_{\mathcal{S}}) \subseteq \mathcal{R}(\lambda)$, tada je funkcija $\lambda|_{\mathcal{S}} : \mathcal{S} \to \mathcal{R}(\lambda|_{\mathcal{S}})$ bijektivna funkcija, pa ekvivalentno tome ima inverznu funkciju $\lambda|_{\mathcal{S}}^{-1} : \mathcal{R}(\lambda|_{\mathcal{S}}) \to \mathcal{S}$.

Jednadžbom (2.36) je definirana funkcija

$$F(\lambda) \equiv (F_1(\lambda), \ldots, F_m(\lambda)) \ , \tag{2.48}$$

čije su varijable $\lambda = (\lambda_1, \ldots, \lambda_m)$. U idealiziranoj teoriji slijedeći korak bio bi nalaženje jednoznačne restrikcije funkcije F na skup $\mathcal{R}(\lambda|_{\mathcal{S}})$, takve da je time definirano preslikavanje $F|_{\mathcal{R}(\lambda|_{\mathcal{S}})} : \mathcal{R}(\lambda|_{\mathcal{S}}) \to \mathcal{S}$. S obzirom da pri tome vrijedi

$$F|_{\mathcal{R}(\lambda|_{\mathcal{S}})}(\lambda) = \lambda|_{\mathcal{S}}^{-1}(\lambda) \ , \tag{2.49}$$

za svaki $\lambda = (\lambda_1, \ldots, \lambda_m)$ iz skupa $\mathcal{R}(\lambda|_{\mathcal{S}})$, funkcija $\lambda|_{\mathcal{S}}^{-1} : \mathcal{R}(\lambda|_{\mathcal{S}}) \to \mathcal{S}$ je time određena. Jednakost (2.49) nam dozvoljava da rješavanjem jednadžbi (2.36) za sve $F = (F_1, \ldots, F_m)$ iz skupa $\mathcal{S}$ odredimo vrijednosti funkcije $\lambda|_{\mathcal{S}}(F)$. Funkcija $\lambda|_{\mathcal{S}} : \mathcal{S} \to \mathcal{R}(\lambda|_{\mathcal{S}})$ je restrikcija funkcije λ na skup $\mathcal{S} \subseteq \mathcal{D}(\lambda)$, bijektivna je i ima inverznu funkciju $\lambda|_{\mathcal{S}}^{-1}$.

Pretpostavimo zatim dodatno da postoji konačna familija $\{\mathcal{S}_1, \ldots, \mathcal{S}_n\}$ nepraznih, međusobno disjunktnih podskupova od $\mathcal{D}(\lambda)$, takvih da su funkcije $\lambda|_{\mathcal{S}_i} : \mathcal{S}_i \to \mathcal{R}(\lambda|_{\mathcal{S}_i})$

bijekcije, za svaki $i = 1, 2, \ldots, n$, i da je $\mathcal{D}(\lambda) = \bigcup_{i=1}^{n} \mathcal{S}_i$. Iz definicije restrikcije funkcije slijedi da je $\lambda|_{\mathcal{S}_i}(F) = \lambda(F)$ za sve $F = (F_1, \ldots, F_m) \in \mathcal{S}_i$, gdje je $\mathcal{S}_i \in \{\mathcal{S}_1, \ldots, \mathcal{S}_n\}$. To implicira da se konačnim brojem prethodno opisanih postupaka mogu odrediti vrijednosti funkcije λ na cijeloj njezinoj domeni definicije $\mathcal{D}(\lambda)$.

Time je pokazano da su uz uvedene pretpostavke Lagrangeovi multiplikatori λ_k, $k = 1, 2, \ldots, m$, jednoznačne funkcije $\lambda_k(F)$ očekivanih vrijednosti $F = (F_1, \ldots, F_m)$. Osim toga, Lagrangeovi multiplikatori su jedini parametri u MaxEnt raspodjeli vjerojatnosti (2.35). Time je njihova uloga jasno definirana. Maksimizacija informacijske entropije, uz ograničenja koja su zadana očekivanim vrijednostima, algoritam je konstrukcije raspodjele vjerojatnosti, takav da se u raspodjelu vjerojatnosti uključuje jedino informacija koja je reprezentirana zadanim ograničenjima. Informacijska entropija takve raspodjele vjerojatnosti je nesigurnost koja je povezana sa informacijom koja nedostaje; preciznije, to je nesigurnost koja preostaje kada se u raspodjelu vjerojatnosti uključi jedino informacija reprezentirana zadanim ograničenjima, i nikakva druga informacija osim nje. Zato se može reći da je informacijska entropija MaxEnt raspodjele vjerojatnosti "objektivna" nesigurnost [25]. Nesigurnost reprezentirana MaxEnt raspodjelom vjerojatnosti može se alternativno interpretirati kao iznos informacije koja nam nedostaje za postizanje sigurnosti [13].

Pri tome treba naglasiti da se atribut "objektivna" u gornjoj rečenici odnosi jedino na nesigurnost. MaxEnt raspodjela vjerojatnosti ovisi jedino o informaciji koja je dostupna, pa s obzirom da ne ovisi o informaciji koja nije dostupna, iz tog razloga ne možemo joj dati oznaku objektivnosti. S druge strane, raspodjela vjerojatnosti koja uključuje jedino dostupnu informaciju, a ne uključuje pretpostavke povezane s informacijom koja nije dostupna, jedina je raspodjela vjerojatnosti koja reprezentira stvarni stupanj znanja. Na taj način je MaxEnt raspodjeli vjerojatnosti ipak dan objektivan značaj, ali se time, kao što je prethodno naglašeno, podrazumijeva interpretacija vjerojatnosti kao reprezentacije stupnja znanja. To svojstvo MaxEnt raspodjele vjerojatnosti je općenito, i ne ovisi o tome je li informacija kojom raspolažemo valjana, i je li relevantna za problem koji nas zanima. Jedina iznimka je slučaj kada je informacija kojom raspolažemo kontradiktorna, odnosno kada sustav jednadžbi (2.27) i (2.28) nije kompatibilan. Tada, kao što je i logično očekivati, maksimizacija informacijske entropije uz ograničenja (2.27) i (2.28) neće dati nikakvo rješenje za Lagrangeove multiplikatore i raspodjelu vjerojatnosti.

Primjena MaxEnt algoritma slijedi jednostavnom interpretacijom prethodnih tvrdnji. Pretpostavimo da na osnovi dostupnih podataka koji se sastoje *jedino* od očekivanih vrijednosti $\langle f_k(x) \rangle = F_k$, $k = 1, 2, \ldots, m$, želimo predvidjeti očekivanu vrijednost neke druge funkcije $g(x)$. Kao što je istaknuo Jaynes [1], korištenje bilo koje druge raspodjele vjerojatnosti osim MaxEnt raspodjele (2.35) svodilo bi se na proizvoljne pretpostavke o informaciji koju po osnovnoj hipotezi nemamo. Ako se naše predikcije veličine $g(x)$ i njene očekivane vrijednosti $\langle g(x)\rangle$ ipak razlikuju od vrijednosti dobivenih eksperimentom, razlozi za to mogu biti slijedeći:

- podaci dani očekivanim vrijednostima $\langle f_k(x)\rangle = F_k$, $1 \leq k \leq m$, nisu u potpunosti valjani. Ta se mogućnost može isključiti ako su očekivane vrijednosti tih veličina parametri koje kontroliramo (mjerimo) u eksperimentu.

- naša prethodna informacija o prostoru elementarnih događaja Ω nije valjana. U kontekstu fizike to ima osobitu važnost, jer prostor elementarnih događaja koji se razlikuje od pretpostavljenog prostora upućuje na postojanje novih zakona koji prethodno nisu bili poznati [1, 2].

- dostupna informacija (podaci + prethodna informacija) nije dovoljna za preciznu predikciju $g(x)$. To znači da relevantna informacija za predikciju te veličine nije (u cijelosti) dostupna.

Ako se naše predikcije veličine $g(x)$ i njene očekivane vrijednosti $\langle g(x)\rangle$ dobro slažu s vrijednostima iz eksperimenta, na osnovi toga možemo zaključiti da je informacija reprezentirana ograničenjima (2.27) dovoljna za precizno predviđanje vrijednosti te veličine. U suprotnom slučaju, postupak maksimizacije informacijske entropije možemo ponoviti, uzimajući sada uz ograničenja koja su zadana očekivanim vrijednostima (2.27) i dodatno ograničenje koje je zadano očekivanom vrijednošću $g(x)$ iz eksperimenta. Na sličan način, postupak se može ponavljati do trenutka kada, u idealnom slučaju, raspolažemo informacijom o veličinama koje su relevantne za predikciju vrijednosti drugih veličina koje nas zanimaju. U svakom od koraka procedure, rezultirajuća MaxEnt raspodjela vjerojatnosti reprezentira stvarni stupanj znanja, a u idealnom slučaju to znanje je relevantno za predikcije koje nas zanimaju. Ova interpretacija je u skladu s Jaynesovom tvrdnjom [25, str. 370] da načelo najveće informacijske entropije predstavlja pravilo induktivnog zaključivanja, koje nam kaže koje su predikcije najsnažnije naznačene našom trenutnom informacijom.[7] U slučaju kad predikcije nisu u skladu s rezultatima eksperimenta, opisanom procedurom definirana je [13, str. 29] metoda uključivanja nove informacije u raspodjelu vjerojatnosti.

Na kraju ovog odjeljka ukratko izvodimo neke od važnih relacija koje također slijede iz općenitog MaxEnt formalizma. Tvrdnje koje su prethodno dokazane u ovom odjeljku, zajedno s jednadžbama (2.37) i (2.47), impliciraju da je maksimum informacijske entropije

[7] Ovdje se pojam induktivno zaključivanje ne odnosi na metodu zasnovanu na principu matematičke indukcije, koja se koristi u deduktivnom matematičkom dokazivanju.

uz ograničenja (2.27), neka funkcija očekivanih vrijednosti $F = (F_1, \ldots, F_m)$,

$$(S_I)_{\max} = \log Z(\lambda_1, \ldots, \lambda_m) + \sum_{k=1}^{m} \lambda_k F_k = S(F_1, \ldots, F_m) \ . \tag{2.50}$$

Uz pretpostavku da funkcije $\lambda_k(F)$, $k = 1, 2, \ldots, m$, imaju neprekidne prve parcijalne derivacije (ili su barem po dijelovima glatke), korištenjem jednadžbi (2.36) i (2.50) slijedi

$$\lambda_k = \frac{\partial S(F_1, \ldots, F_m)}{\partial F_k} \ , \qquad k = 1, 2, \ldots, m \ . \tag{2.51}$$

Na osnovi jednadžbi (2.36), (2.50) i (2.51) očigledno je da su funkcija $\log Z(\lambda_1, \ldots, \lambda_m)$ i funkcija $S(F_1, \ldots, F_m)$ međusobno povezane Legendreovim transformacijama. U slučajevima kad je takva transformacija moguća, funkcije koje su na taj način povezane sadrže istu informaciju izraženu pomoću različitih varijabli.

Funkcije $\log Z(\lambda_1, \ldots, \lambda_m)$ i $S(F_1, \ldots, F_m)$ su povezane s MaxEnt raspodjelom vjerojatnosti (2.35) preko vrijednosti parametara $\lambda = (\lambda_1, \ldots, \lambda_m)$ i $F = (F_1, \ldots, F_m)$. Nadalje, ove funkcije na jednostavan način daju varijance i kovarijance funkcija $f_k(x)$, $k = 1, 2, \ldots, m$. Korištenjem jednadžbi (2.34), (2.35) i (2.36), dobije se

$$\begin{aligned} \frac{\partial^2 \log Z(\lambda_1, \ldots, \lambda_m)}{\partial \lambda_l \partial \lambda_k} &= -\frac{\partial F_k}{\partial \lambda_l} = -\frac{\partial F_l}{\partial \lambda_k} \\ &= \langle f_k(x) f_l(x) \rangle - \langle f_k(x) \rangle \langle f_l(x) \rangle \\ &= -A_{kl} \ , \qquad k, l = 1, 2, \ldots, m \ . \end{aligned} \tag{2.52}$$

Na sličan način, korištenjem jednadžbe (2.51) se dobije

$$\begin{aligned} \frac{\partial^2 S(F_1, \ldots, F_m)}{\partial F_l \partial F_k} &= \frac{\partial \lambda_k}{\partial F_l} = \frac{\partial \lambda_l}{\partial F_k} \\ &= B_{kl} \ , \qquad k, l = 1, 2, \ldots, m \ . \end{aligned} \tag{2.53}$$

Tada, korištenjem jednadžbi (2.52), (2.53) i pravila za lančano deriviranje, slijedi

$$\frac{\partial \lambda_j}{\partial \lambda_l} = \sum_{k=1}^{m} \frac{\partial \lambda_j}{\partial F_k} \frac{\partial F_k}{\partial \lambda_l} = B_{jk} A_{kl} = \delta_{jl} \ , \qquad j, l = 1, 2, \ldots, m \ , \tag{2.54}$$

i slično tome,

$$\frac{\partial F_j}{\partial F_l} = \sum_{k=1}^{m} \frac{\partial F_j}{\partial \lambda_k} \frac{\partial \lambda_k}{\partial F_l} = A_{jk} B_{kl} = \delta_{jl} \ , \qquad j, l = 1, 2, \ldots, m \ . \tag{2.55}$$

Dakle, matrice dane jednadžbama (2.52) i (2.53) su inverzne, $A^{-1} = B$.

Elementi matrice A su druge derivacije funkcije $\log Z(\lambda_1, \ldots, \lambda_m)$ koje predstavljaju mjeru *očekivane* disperzije i međusobne koreliranosti funkcija $f_k(x)$, $k = 1, 2, \ldots, m$. Dijagonalni elementi matrice A nam daju predodžbu o odstupanju vrijednosti varijabli $f_k(x)$

od njihovih očekivanih vrijednosti $\langle f_k(x) \rangle$. Korištenjem jednadžbi (2.34), (2.35) i (2.36) može se pokazati da kovarijanca neke druge funkcije $g(x)$ s funkcijama $f_k(x)$ slijedi izravno iz jednadžbe

$$-\frac{\partial \langle g(x) \rangle}{\partial \lambda_k} = \langle g(x) f_k(x) \rangle - \langle g(x) \rangle \langle f_k(x) \rangle \ , \qquad k = 1, 2, \ldots, m \ . \tag{2.56}$$

2.4 Interpretacija MaxEnt formalizma u statističkoj fizici

U prethodnom odjeljku je pokazano da su Lagrangeovi multiplikatori, uz uvedene pretpostavke, jednoznačne funkcije očekivanih vrijednosti kojima su zadana ograničenja na maksimizaciju informacijske entropije. S obzirom da su to jedini parametri u MaxEnt raspodjeli vjerojatnosti, možemo reći da je posredstvom Lagrangeovih multiplikatora precizirana MaxEnt raspodjela koja je kompatibilna sa zadanim ograničenjima. U kontekstu primjene načela najveće informacijske entropije u statističkoj fizici, interpretacija tih veličina je od posebne važnosti. U ovom odjeljku će biti pokazano da fizikalna interpretacija Lagrangeovih multiplikatora slijedi iz osnovne relacije kojom su opisane promjene očekivanih vrijednosti.

Napomenimo da očekivane vrijednosti $F = (F_1, \ldots, F_m)$ općenito mogu predstavljati bilo koje fizikalne veličine o kojima imamo informaciju. Zato možemo, bez potrebe mijenjanja notacije, nastaviti razmatrati općeniti slučaj iz prethodnog odjeljka, gdje su $F = (F_1, \ldots, F_m)$ bile očekivane vrijednosti nekih funkcija $(f_1(x), \ldots, f_m(x))$. U kontekstu statističke fizike, može se uzeti da vrijednosti $f_k(x_i)$ pridružene vrijednostima x_i predstavljaju svojstvene vrijednosti neke fizikalne veličine, naprimjer svojstvene vrijednosti energije E_i, ili svojstvene vrijednosti neke druge veličine iz skupa kompatibilnih veličina.

Pretpostavimo sada da očekivane vrijednosti $\langle f_k(x) \rangle$ mijenjamo malim promjenama vrijednosti funkcija $f_k(x_i)$ i vjerojatnosti p_i,

$$\delta \langle f_k(x) \rangle = \sum_i p_i \delta f_k(x_i) + \sum_i f_k(x_i) \delta p_i \ , \qquad k = 1, 2, \ldots, m \ . \tag{2.57}$$

Iz relacije (2.57) je očigledno da u općenitom slučaju promjena očekivane vrijednosti $\delta \langle f_k(x) \rangle$ i očekivana vrijednost promjene $\langle \delta f_k(x) \rangle = \sum_i p_i \delta f_k(x_i)$ nisu jednake. Razlika tih dviju veličina ovisi o promjenama δp_i,

$$\delta \langle f_k(x) \rangle - \langle \delta f_k(x) \rangle = \sum_i f_k(x_i) \delta p_i \ , \qquad k = 1, 2, \ldots, m \ . \tag{2.58}$$

Promjena informacijske entropije S_I je jednaka

$$\delta S_I = -\sum_i \delta p_i \log p_i \ . \tag{2.59}$$

U prethodnom razmatranju nismo precizirali vjerojatnosti p_i i njihove promjene δp_i, pa izrazi (2.57), (2.58) i (2.59) vrijede općenito. Uvrštavanjem MaxEnt vjerojatnosti (2.35) za p_i u jednadžbi (2.59) i korištenjem jednadžbi (2.58), dobije se

$$\begin{aligned} \delta S &= \sum_{k=1}^{m} \sum_i \lambda_k f_k(x_i) \delta p_i \\ &= \sum_{k=1}^{m} \lambda_k \left(\delta \langle f_k(x) \rangle - \langle \delta f_k(x) \rangle \right) \ . \end{aligned} \tag{2.60}$$

Ako se pretpostavi i to da je raspodjela koja je dana sa $p_i + \delta p_i$ također MaxEnt raspodjela vjerojatnosti, jednadžba (2.60) tada daje promjenu maksimuma informacijske entropije.

Ako su veličine

$$\delta Q_k = \sum_i f_k(x_i) \delta p_i = \delta \langle f_k(x) \rangle - \langle \delta f_k(x) \rangle \ , \qquad k = 1, 2, \ldots, m \ , \tag{2.61}$$

jednake nuli za sve $k = 1, 2, \ldots, m$, tada je $\delta S = 0$. Smisao jednadžbe (2.60) je lakše razumjeti ako se korištenjem (2.61) napiše u obliku

$$\delta S = \sum_{k=1}^{m} \lambda_k \delta Q_k \ . \tag{2.62}$$

Jednadžba (2.61) sugerira interpretaciju koju je dao Jaynes [1], i posebno Grandy [13, 16]. Očekivana vrijednost $\langle \delta f_k(x) \rangle$ promjene $\delta f_k(x)$ je generalizirani "rad". Preostali dio promjene $\delta \langle f_k(x) \rangle$ očekivane vrijednosti $\langle f_k(x) \rangle$ dolazi od promjene u raspodjeli vjerojatnosti $\{p_i\}$ i predstavlja generaliziranu "toplinu" δQ_k za veličinu $f_k(x)$. Ako je funkcija $f_k(x)$ takva da je $f_k(x_i) = E_i$ za sve i, tada je δQ_k toplina u uobičajenom smislu.

Tri vrste promjena u jednadžbi (2.61), kao što je istaknuo Grandy [13, str. 30], su međusobno povezane, pa specificiranje bilo koje dvije određuje treću. Pri tome treba voditi računa o kompatibilnosti sustava jednadžbi (2.61) sa δp_i kao nepoznanicama. Grandy [13, 16] je jednadžbe (2.61) interpretirao kao općenito pravilo u teoriji vjerojatnosti, čiji poseban slučaj je prvi zakon termodinamike. Za neki makroskopski sustav, ako je $f_k(x_i) = E_i$ za sve i, tada je $\langle E \rangle = U$ unutarnja energija, pa odgovarajuća jednadžba (2.61) ima oblik

$$\delta Q = \delta \langle E \rangle - \langle \delta E \rangle = \delta U - \delta W \ , \tag{2.63}$$

gdje je $\delta W = \langle \delta E \rangle$ rad izvršen na sustavu. Pri tome je [13, 16] toplina δQ energija prenesena preko stupnjeva slobode nad kojima nemamo kontrolu, dok je rad δW energija prenesena preko stupnjeva slobode koje kontroliramo. U takvoj interpretaciji, svaki od generaliziranih δQ_k u jednadžbama (2.61) promjena je odgovarajuće očekivane vrijednosti $\delta \langle f_k(x) \rangle$ povezana s promjenom u raspodjeli vjerojatnosti. Ako su $\{p_i\}$ i $\{p_i + \delta p_i\}$ MaxEnt raspodjele vjerojatnosti, tada jednadžba (2.60), ili ekvivalentna jednadžba (2.62), daju promjenu maksimuma informacijske entropije δS. Jednadžbe (2.61) i (2.62) eksplicitno pokazuju da promjena maksimuma informacijske entropije dolazi od promjene u raspodjeli vjerojatnosti koja je povezana sa δQ_k.

Grandy je taj generalizirani član doveo u vezu s promjenom makroskopskih ograničenja, koja je rezultat vanjskih utjecaja na sustav. Polazeći od tog opažanja Grandy je razvio općeniti pristup [13, 16, 17, 18], koji uz poopćenje Liouville–von Neumannove jednadžbe za matricu gustoće kroz primjenu MaxEnt formalizma vodi do izvoda makroskopskih jednadžbi gibanja. Pristup koji razvijamo u ovom radu zasniva se, kao i Grandyjev općeniti model vremenski ovisnih vjerojatnosti, na maksimizaciji informacijske entropije uz makroskopska ograničenja. Iz tog razloga su interpretacije tih modela kompatibilne.

Razmotrimo sada kvazistatičnu promjenu[8] energije nekog makroskopskog sustava, za kojeg specificiramo jedino da je zatvoren sustav (sustav koji s okolinom može izmjenjivati energiju u obliku rada ili topline, ali ne i čestice). Iz jednadžbi (2.62) i (2.63) slijedi da je tada

$$\delta S = \lambda \delta Q \ , \tag{2.64}$$

i

$$\delta U - \delta W = \delta Q = \frac{1}{\lambda} \delta S \ . \tag{2.65}$$

U termodinamici je jednadžbom istog oblika kao (2.65) dan prvi zakon termodinamike. Ako prvi zakon napišemo u obliku u kojem se eksplicitno pojavljuje termodinamička entropija S_e,

$$dU - \delta W = \delta Q = T dS_e \ , \tag{2.66}$$

Lagrangeov multiplikator λ u analognoj jednadžbi (2.65) možemo identificirati kao

$$\lambda = \frac{1}{kT} \ . \tag{2.67}$$

Promjena maksimuma informacijske entropije δS dana relacijom (2.64) je tada povezana s potpunim diferencijalom termodinamičke entropije dS_e relacijom

$$k dS = dS_e = \frac{\delta Q}{T} \ , \tag{2.68}$$

gdje je T temperatura, a $1/T$ integrativni faktor topline δQ. Izbor jedinice za temperaturu (Kelvin), odnosno za entropiju (Joule Kelvin^{-1}), reflektira se u pojavi Boltzmannove konstante k u prethodnim izrazima.

[8] Proces se naziva kvazistatičnim ako su promjene od kojih se sastoji "dovoljno spore" da se proces može opisati kao uređen slijed stanja makroskopskog sustava koja su "proizvoljno blizu" ravnotežnim stanjima [30].

Potvrdu da je identifikacija dana jednadžbom (2.67) ispravna možemo dobiti uvrštavanjem vrijednosti Lagrangeovog multiplikatora $\lambda = (kT)^{-1}$ u MaxEnt raspodjelu vjerojatnosti koja odgovara slučaju koji ovdje razmatramo. Na taj način se dobije

$$p_i = \frac{1}{Z} \exp\left(-\frac{E_i}{kT}\right) \ , \tag{2.69}$$

što je Gibbsova kanonska raspodjela, kojom je u statističkoj fizici opisan zatvoren sustav poznate temperature u ravnoteži s okolinom. Normalizacijski faktor kanonske raspodjele, particijska funkcija Z, jednaka je

$$Z = \sum_i \exp(-\lambda E_i) = \sum_i \exp\left(-\frac{E_i}{kT}\right) \ . \tag{2.70}$$

Po analogiji s jednadžbom (2.36), očekivana vrijednost energije sustava $\langle E \rangle$ jednaka je derivaciji particijske funkcije (2.70),

$$U = \langle E \rangle = -\frac{\partial \log Z}{\partial \lambda} \ . \tag{2.71}$$

Promjena raspodjele vjerojatnosti (2.69) pri razmatranoj kvazistatičnoj promjeni energije je jednaka

$$\begin{aligned} \delta p_i &= \frac{\partial p_i}{\partial \lambda}\delta\lambda + \sum_j \frac{\partial p_i}{\partial E_j}\delta E_j \\ &= (\langle E \rangle - E_i)p_i\delta\lambda + \lambda(\langle \delta E \rangle - \delta E_i)p_i \ . \end{aligned} \tag{2.72}$$

Korištenjem (2.72) dobije se da je kvazistatična promjena očekivane vrijednosti energije $\langle E \rangle$ jednaka

$$\begin{aligned} \delta\langle E \rangle &= \sum_i E_i \delta p_i + \sum_i p_i \delta E_i \\ &= \sum_i E_i \frac{\partial p_i}{\partial \lambda}\delta\lambda + \sum_i \sum_j E_i \frac{\partial p_i}{\partial E_j}\delta E_j + \sum_i p_i \delta E_i \\ &= (\langle E \rangle^2 - \langle E^2 \rangle)\delta\lambda + \lambda(\langle E \rangle\langle \delta E \rangle - \langle E \delta E \rangle) + \langle \delta E \rangle \end{aligned} \tag{2.73}$$

Relacija (2.73) nije u suprotnosti s tvrdnjom koja je dokazana u odjeljku 2.3, da su Lagrangeovi multiplikatori $\lambda_k(F)$, $k = 1, 2, \ldots, m$, jednoznačne funkcije očekivanih vrijednosti $F = (F_1, \ldots, F_m)$. Važno je istaknuti da je ta tvrdnja dokazana uz pretpostavku da su vrijednosti funkcija $f_k(x_i)$ fiksne za sve i, k, ili eksplicitnije kazano, da se vrijednosti funkcija $f_k(x_i)$ ne mijenjaju pri promjeni očekivanih vrijednosti $F = (F_1, \ldots, F_m)$. U slučaju koji je ovdje razmatran ta prepostavka nije zadovoljena.

Razmotrimo sada poseban slučaj kad je ta pretpostavka zadovoljena, tj. kada su sve svojstvene energije E_i fiksne. Tada se iz općenite relacije (2.73) uz korištenje (2.52) dobije

$$\begin{aligned} \delta U &= \delta\langle E \rangle = \sum_i E_i \delta p_i = \sum_i E_i \frac{\partial p_i}{\partial \lambda}\delta\lambda = (\langle E \rangle^2 - \langle E^2 \rangle)\delta\lambda \\ &= -\frac{d^2 \log Z}{d\lambda^2}\delta\lambda = \frac{dU}{d\lambda}\delta\lambda \end{aligned} \tag{2.74}$$

pa je očito da u ovom slučaju prethodna tvrdnja vrijedi. Očigledno je također da identifikacija (2.67) određuje Langrangeov multiplikator λ kao funkciju koja, osim o očekivanoj vrijednosti energije $\langle E \rangle = U$, ovisi i o drugim parametrima. Promjena tih parametara dovodi do promjena δE_i svojstvenih energija E_i. Među tim parametrima mogu biti mehanički parametri poput volumena sustava V, ili/i neki drugi parametri poput magnetskog polja $\mathbf{H}$, i sl.

Informacijska entropija Gibbsove kanonske raspodjele (2.69) je jednaka

$$S = \log Z + \lambda \langle E \rangle = \log Z + \frac{U}{kT} \ . \tag{2.75}$$

Promjena informacijske entropije pri promjeni Gibbsove kanonske raspodjele (2.69) u općenitom kvazistatičnom slučaju jednaka je

$$\begin{aligned} dS &= \sum_i \frac{\partial \log Z}{\partial E_i} \delta E_i + \lambda \delta \langle E \rangle = -\lambda \sum_i p_i \delta E_i + \lambda \delta \langle E \rangle \\ &= \lambda (\delta \langle E \rangle - \langle \delta E \rangle) \\ &= \lambda \delta Q \\ &= \frac{\delta Q}{kT} \ , \end{aligned} \tag{2.76}$$

gdje posljednja dva reda slijede korištenjem (2.63) i (2.67). Zbog jednostavnosti su u prvom redu izvoda jednadžbe (2.76) napisani samo oni članovi koji se ne pokrate u nulu. Jednadžba (2.76) potvrđuje ispravnost označavanja promjene informacijske entropije Gibbsove raspodjele kao potpunog diferencijala dS u jednadžbi (2.68).

Lagrangeov multiplikator λ, odnosno inverzna temperatura kT, jednak je parcijalnoj derivaciji informacijske entropije Gibbsove raspodjele (2.75) po očekivanoj vrijednosti energije, pri čemu su svi drugi parametri konstantni. To je jednostavno pokazati korištenjem (2.71) i (2.75),

$$\begin{aligned} \left(\frac{\partial S}{\partial U} \right)_{\{\alpha\}} &= \frac{\partial \log Z}{\partial U} + \frac{\partial \lambda}{\partial U} U + \lambda \\ &= \frac{\partial \log Z}{\partial \lambda} \frac{\partial \lambda}{\partial U} + \frac{\partial \lambda}{\partial U} U + \lambda \\ &= -U \frac{\partial \lambda}{\partial U} + \frac{\partial \lambda}{\partial U} U + \lambda \\ &= \lambda = \frac{1}{kT} \ , \end{aligned} \tag{2.77}$$

gdje $\{\alpha\}$ označava da su svi parametri iz skupa $\{\alpha\}$ konstantni pri deriviranju, što se podrazumijeva i u svim ostalim parcijalnim derivacijama u jednadžbi. O vrijednostima parametara iz skupa $\{\alpha\}$ ovise vrijednosti svojstvenih energija E_i.

U statističkoj fizici je standardna interpretacija kanonske raspodjele (2.69) kao raspodjele koja opisuje sustav poznate temperature, zadane temperaturom okoline (toplinski rezervoar) s kojom je promatrani sustav u ravnoteži. Ukupan sustav se sastoji od promatranog sustava i okoline koji po pretpostavci interagiraju slabo, a opisan je mikrokanonskom raspodjelom. U ovom odjeljku je primjenom Jaynesovog pristupa [1, 5] pokazano da MaxEnt raspodjela vjerojatnosti za slučaj kada je poznata jedino očekivana vrijednost energije sustava $\langle E \rangle = U$ odgovara Gibbsovoj kanonskoj raspodjeli. Polazna točka u izvodu je tvrdnja, koja je argumentirana u odjeljku 2.3, da MaxEnt raspodjela vjerojatnosti uključuje jedino dostupnu informaciju o sustavu bez ikakvih dodatnih pretpostavki. Identifikacija Lagrangeovog multiplikatora λ kao inverzne temperature (2.67), slijedi kroz razmatranje kvazistatičke promjene stanja sustava, uz korištenje jednadžbe (2.65) kao prvog zakona termodinamike. Razmatranjem otvorenog sustava (sustava koji može izmjenjivati energiju i čestice s okolinom) na analogan način se pokaže da MaxEnt raspodjela vjerojatnosti za slučaj kada su uz očekivanu vrijednost energije jedino poznate očekivane vrijednosti broja čestica, odgovara Gibbsovoj velekanonskoj raspodjeli [1, 13].

Poglavlje 3

REPREZENTACIJA VREMENSKE EVOLUCIJE HAMILTONOVIH SUSTAVA

3.1 Hamiltonova dinamika i putanje u faznom prostoru

Dinamičko stanje Hamiltonovog sustava sa s stupnjeva slobode opisano je koordinatama $q_1, q_2, \ldots, q_s$ i impulsima $p_1, p_2, \ldots, p_s$. U bilo kojem trenutku vremena t stanje sustava je reprezentirano točkom u $2s$-dimenzionalnom prostoru Γ koji se naziva fazni prostor sustava. Notaciju (q, p) uvodimo za skup $2s$ koordinata i impulsa. Vremenska ovisnost $2s$ dinamičkih varijabli (q, p) određena je Hamiltonovim jednadžbama

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \qquad \dot{p}_i = -\frac{\partial H}{\partial q_i}, \qquad i = 1, 2, \ldots, s, \tag{3.1}$$

gdje je $H \equiv H(q, p)$ Hamiltonova funkcija sustava. Za dane vrijednosti (q_0, p_0) u nekom trenutku t_0, rješenje Hamiltonovih jednadžbi (3.1) jedinstveno određuje vrijednosti dinamičkih varijabli (q, p) u bilo kojem drugom trenutku t,

$$q_i = q_i(t; q_0, p_0), \qquad p_i = p_i(t; q_0, p_0), \qquad i = 1, 2, \ldots, s. \tag{3.2}$$

Bilo koja točka (q, p) u faznom prostoru Γ opisuje krivulju nazvanu *putanjom u faznom prostoru*, koja je jedinstveno određena rješenjem jednadžbi (3.1). U bilo kojem trenutku t, kroz bilo koju točku u Γ prolazi samo jedna putanja, i to je označeno indeksom u $(q, p)_\omega$, gdje je $\omega \in \Omega(\Gamma)$. Skup $\Omega(\Gamma)$ je skup svih putanja u Γ. Brzina točke (q, p) u faznom prostoru Γ u trenutku t je

$$v \equiv |\mathbf{v}| = \sqrt{\sum_{i=1}^{s} \left(\frac{dq_i}{dt}\right)^2 + \sum_{i=1}^{s} \left(\frac{dp_i}{dt}\right)^2} = \sqrt{\sum_{i=1}^{s} \left(\frac{\partial H}{\partial p_i}\right)^2 + \sum_{i=1}^{s} \left(\frac{\partial H}{\partial q_i}\right)^2}. \tag{3.3}$$

Vektor brzine $\mathbf{v}((q,p)_\omega, t)$ je tangencijalan u točki $(q,p)_\omega \in \Gamma$ na putanju ω koja prolazi kroz tu točku u trenutku t. Za sustave koje ovdje razmatramo Hamiltonova funkcija $H(q,p)$ ne ovisi o vremenu, pa je polje brzine $\mathbf{v}(q,p,t)$ svih točaka u faznom prostoru Γ stacionarno, tj. $\mathbf{v}(q,p,t) = \mathbf{v}(q,p)$.

Neka je M_0 bilo koji izmjeriv (u Lebesgueovom smislu) skup točaka u faznom prostoru Γ. Pri Hamiltonovom gibanju skup M_0 transformira se u skup M_t u vremenskom intervalu t. Liouvilleov teorem tvrdi da mjera skupa M_t za bilo koji t koincidira s mjerom skupa M_0 [31, str. 15–16]. Ovaj teorem dokazuje da je mjera u faznom prostoru Γ,

$$\mu(M_t) = \int_{M_t} dq_1 \ldots dq_s dp_1 \ldots dp_s = \int_{M_t} d\Gamma, \tag{3.4}$$

invarijantna pri Hamiltonovom gibanju. U notaciji korištenoj u jednadžbi (3.4), element volumena $dq_1 \ldots dq_s dp_1 \ldots dp_s$ faznog prostora Γ označen je s $d\Gamma$. Neposredan korolar [31, str. 18–19] Liouvilleovog teorema je da, ako je M_0 Lebesgue izmjeriv skup točaka faznog prostora Γ, konačne mjere, i $f(q,p)$ funkcija Lebesgue integrabilna na Γ, tada je

$$\int_{M_t} f(q,p) d\Gamma = \int_{M_0} f(q(t; q_0, p_0), p(t; q_0, p_0)) d\Gamma_0. \tag{3.5}$$

Jednadžba (3.5) se dobije zamjenom varijabli u integralu i uvođenjem novih varijabli (q_0, p_0), povezanih s varijablama (q,p) transformacijom prostora Γ u sebe pri Hamiltonovom gibanju, koja je dana jednadžbama (3.2). Ako je, posebice, skup M invarijantan na Hamiltonovo gibanje, tada je korištenjem ovog korolara lako pokazati kako se integral funkcije $f(q,p)$ na invarijantnom skupu M može transformirati u integraciju na skupu $\Omega(M)$ svih putanja u M. Ova procedura će sada biti razvijena u ostatku ovog odjeljka. Ona se koristi u definiciji raspodjela vjerojatnosti u odjeljku 3.2.

U bilo kojem trenutku t kroz bilo koju točku $(q,p)_\omega \in \Gamma$ prolazi samo jedna putanja $\omega \in \Omega(\Gamma)$. Putanja ω prolazi također kroz točku $(q_0, p_0)_\omega \in \Gamma$ danu inverzijom jednadžbi (3.2). Infinitezimalni element volumena $d\Gamma_0$ oko točke $(q_0, p_0)_\omega$ može se zapisati kao $d\Gamma_0 = ds_{0\omega} dS_{0\omega}$. Ovdje je $ds_{0\omega}$ infinitezimalna udaljenost duž putanje ω. Infinitezimalni element $dS_{0\omega}$ plohe $S_0(M)$ presijeca putanju ω okomito u točki $(q_0, p_0)_\omega$. (Hiper)ploha $S_0(M)$ je okomita na sve putanje u skupu $\Omega(M)$ putanja u M.

Invarijantnost mjere $d\Gamma$ na Hamiltonovo gibanje i činjenica da je polje brzine $\mathbf{v}(q,p)$ u Γ stacionarno jer Hamiltonova funkcija $H(q,p)$ ne ovisi o vremenu, vode do slijedeće posljedice. Za bilo koju putanju u faznom prostoru $\omega \in \Omega(\Gamma)$, produkt brzine $v((q,p)_\omega)$ i infinitezimalnog elementa plohe dS_ω koji presijeca putanju ω okomito u točki $(q,p)_\omega$, konstantan je pri Hamiltonovom gibanju duž cijele duljine putanje ω, tj.

$$v((q,p)_\omega) dS_\omega = konst. \tag{3.6}$$

Za bilo koje dvije točke $(q_0, p_0)_\omega$ i $(q_a, p_a)_\omega$ na istoj putanji ω, slijedeća relacija se dobije iz (3.6):

$$v((q_0, p_0)_\omega) dS_{0\omega} = v((q_a, p_a)_\omega) dS_{a\omega}. \tag{3.7}$$

Infinitezimalni element $dS_{a\omega}$ plohe $S_a(M)$ presijeca putanju ω okomito u točki $(q_a, p_a)_\omega$. Poput plohe $S_0(M)$, ploha $S_a(M)$ je također okomita na sve putanje u $\Omega(M)$. Infinitezimalni elementi $dS_{0\omega}$ i $dS_{a\omega}$ dviju ploha $S_0(M)$ i $S_a(M)$ povezani su putanjom ω i susjednim putanjama određenim rješenjima Hamiltonovih jednadžbi. Integral na plohi $S_a(M)$ transformira se korištenjem relacije (3.7) u integraciju na plohi $S_0(M)$,

$$\int_{S_a(M)} dS_{a\omega} = \int_{S_0(M)} \frac{v((q_0, p_0)_\omega)}{v((q_a, p_a)_\omega)} dS_{0\omega}. \tag{3.8}$$

Funkcijska zavisnost između točaka $(q_0, p_0)_\omega$ i $(q_a, p_a)_\omega$ na putanji ω nije eksplicitno zapisana u integralu (3.8); implicitno se podrazumijeva da je ova funkcijska zavisnost određena iz rješenja Hamiltonovih jednadžbi.

Slijedeću notaciju uvodimo korištenjem jednadžbi (3.2), s fiksnim vremenima t and t_0, u funkciji $f(q, p)$:

$$f(q(t; q_0, p_0), p(t; q_0, p_0)) \equiv g(q_0, p_0, t_0). \tag{3.9}$$

Jednadžba (3.9) se potom supstituira (s fiksnim t i t_0 i indeksima u (q_0, p_0) zamijenjenim s indeksima (q_a, p_a)) u integral (3.5), uzet na skupu M koji je invarijantan na Hamiltonovo gibanje. To vodi do slijedeće jednakosti:

$$\int_M f(q, p) d\Gamma = \int_M g(q_a, p_a, t_0) d\Gamma_a. \tag{3.10}$$

Integral (3.10) se onda transformira korištenjem relacije (3.7) i $d\Gamma_a = ds_{a\omega} dS_{a\omega}$:

$$\int_M g(q_a, p_a, t_0) ds_{a\omega} dS_{a\omega} = \int_{S_0(M)} dS_{0\omega} v((q_0, p_0)_\omega) \int_\omega \frac{g(q_a, p_a, t_0)}{v(q_a, p_a)} ds_{a\omega}. \tag{3.11}$$

Funkcija

$$F((q_0, p_0)_\omega, t_0) = v((q_0, p_0)_\omega) \int_\omega \frac{g(q_a, p_a, t_0)}{v(q_a, p_a)} ds_{a\omega}, \tag{3.12}$$

je definirana na plohi $S_0(M)$ i naziva se *funkcija putanja* ili *raspodjela putanja*. Integral u relaciji (3.12) koja definira funkciju putanja $F((q_0, p_0)_\omega, t_0)$ je duž cijele duljine putanje ω koja je okomito presječena plohom $S_0(M)$ u točki $(q_0, p_0)_\omega$. Infinitezimalni element putanje ω koja prolazi kroz točku $(q_a, p_a)_\omega$ je $ds_{a\omega}$, a vrijeme t_0 u integralu (3.12) je fiksno.

Ako je funkcija $f(q, p)$ u integralu (3.10) funkcija gustoće vjerojatnosti u faznom prostoru, jednaka nuli svuda izvan invarijantnog skupa M, jednostavno se može dokazati da funkcija putanja $F((q_0, p_0)_\omega, t_0)$ definirana jednadžbom (3.12) zadovoljava uvjete nenegativnosti i normiranosti koji se zahtijevaju od raspodjela vjerojatnosti. Nenegativnost i normiranost funkcije $F((q_0,p_0)_\omega,t_0)$ koja tada predstavlja *raspodjelu vjerojatnosti putanja*, slijede iz svojstava nenegativnosti i normiranosti s njom povezane funkcije gustoće vjerojatnosti $f(q,p)$ u faznom prostoru. Zbog jednostavnosti će se za funkcije gustoće vjerojatnosti u ostatku teksta koristiti također i naziv raspodjele vjerojatnosti. Uz pomoć jednadžbi (3.10) i (3.11) i definicije $F((q_0,p_0)_\omega,t_0)$ u jednadžbi (3.12), dobije se onda svojstvo normiranosti

$$\int_M f(q,p)d\Gamma = \int_{S_0(M)} F((q_0,p_0)_\omega,t_0)dS_0 = 1. \tag{3.13}$$

Nenegativnost $F((q_0,p_0)_\omega,t_0)$ se ustanovi za sve $(q_0,p_0)_\omega \in S_0(M)$ na sličan način. Integral na bilo kojem invarijantnom i izmjerivom podskupu skupa M se transformira, na gore opisan način, u integral na odgovarajućem izmjerivom podskupu na plohi $S_0(M)$. Također je jasno da se mjera definirana na plohi $S_0(M)$ može iskoristiti kao mjera na skupu $\Omega(M)$ svih putanja u nekom invarijantnom skupu M. Korespondencija između točaka $(q_0,p_0)_\omega \in S_0(M)$ i putanja $\omega \in \Omega(M)$ je jedan na jedan.

3.2 Vjerojatnost mikrostanja i vjerojatnost putanja

Sada je moguće povezati vjerojatnost mikrostanja i vjerojatnost putanja u faznom prostoru sustava. Neka je $f(q,p,t)$ funkcija gustoće vjerojatnosti mikrostanja u Γ. Sve točke u faznom prostoru Γ gibaju se prema Hamiltonovim jednadžbama (3.1) i $f(q,p,t)$ zadovoljava Liouvilleovu jednadžbu

$$\frac{\partial f}{\partial t} + \sum_{i=1}^{s}\left(\frac{\partial f}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial H}{\partial q_i}\right) \equiv \frac{df}{dt} = 0. \tag{3.14}$$

Budući da je df/dt totalna odnosno "hidrodinamska" derivacija, jednadžba (3.14) izražava činjenicu da je brzina vremenske promjene $f(q,p,t)$ nula duž bilo koje putanje u faznom prostoru jedinstveno određene rješenjem Hamiltonovih jednadžbi. U notaciji koja se ovdje koristi, ova činjenica je zapisana na slijedeći način,

$$f((q,p)_\omega,t) = f((q_0,p_0)_\omega,t_0), \tag{3.15}$$

gdje su točke na putanji $\omega \in \Omega(\Gamma)$ povezane jednadžbama (3.2).

Uz definiciju raspodjele vjerojatnosti putanja $F((q_0,p_0)_\omega,t_0)$ preko funkcije gustoće vjerojatnosti mikrostanja $f(q,p,t)$, moguće je dati još jednu ekvivalentnu definiciju $F((q_0,p_0)_\omega,t_0)$. Da bi to postigli, uvodimo funkciju gustoće vjerojatnosti $\mathcal{F}(q,p,t;q_0,p_0,t_0)$ na $4s$-dimenzionalnom Euklidskom prostoru $\Gamma\times\Gamma$. Ova funkcija ima slijedeća posebna svojstva. Integracijom funkcije $\mathcal{F}(q,p,t;q_0,p_0,t_0)$ na faznom prostoru

Γ, s (q,p) kao varijablama integracije, dobije se funkcija gustoće vjerojatnosti mikrostanja $f(q_0,p_0,t_0)$ u trenutku t_0,

$$\int_\Gamma \mathcal{F}(q,p,t;q_0,p_0,t_0)d\Gamma = f(q_0,p_0,t_0). \tag{3.16}$$

Funkcija gustoće vjerojatnosti mikrostanja $f(q,p,t)$ u trenutku t dobije se na analogan način,

$$\int_\Gamma \mathcal{F}(q,p,t;q_0,p_0,t_0)d\Gamma_0 = f(q,p,t). \tag{3.17}$$

Jednostavno se može pokazati, korištenjem relacije (3.15), da su jednadžbe (3.16) i (3.17) zadovoljene ako funkcija $\mathcal{F}(q,p,t;q_0,p_0,t_0)$ ima slijedeći oblik:

$$\mathcal{F}(q,p,t;q_0,p_0,t_0) = f(q,p,t)\prod_{i=1}^{s}\delta(q_i - q_i(t;q_0,p_0))\delta(p_i - p_i(t;q_0,p_0)), \tag{3.18}$$

gdje su $q_i(t;q_0,p_0)$ i $p_i(t;q_0,p_0)$ dane jednadžbama (3.2), a δ-e su Diracove delta funkcije. U prostoru $\Gamma\times\Gamma$, funkcija $\mathcal{F}(q,p,t;q_0,p_0,t_0)$ dana jednadžbom (3.18) predstavlja gustoću vjerojatnosti da je točka koja korespondira stanju sustava u elementu $d\Gamma_0$ oko točke (q_0,p_0) u trenutku t_0 i u elementu $d\Gamma$ oko točke (q,p) u trenutku t.

Kao što je objašnjeno u jednadžbi (3.17), funkcija gustoće vjerojatnosti $f(q,p,t)$ je dana integralom funkcije $\mathcal{F}(q,p,t;q_0,p_0,t_0)$ po Γ, s (q_0,p_0) kao varijablama integracije. Pretpostavimo sada da je skup M svih točaka u Γ koje predstavljaju moguća mikrostanja sustava invarijantan na Hamiltonovo gibanje. Primjenom slične procedure i notacije koja je već bila uvedena u relacijama (3.10) i (3.11), integral (3.17) se sada može zapisati kao

$$f(q,p,t) = \int_{S_0(M)} dS_{0\omega}v((q_0,p_0)_\omega)\int_\omega \frac{\mathcal{F}(q,p,t;q_a,p_a,t_0)}{v(q_a,p_a)}ds_{a\omega}. \tag{3.19}$$

Zajedno s jednadžbom (3.19) uvodi se također i funkcija

$$G(q,p,t;(q_0,p_0)_\omega,t_0) = v((q_0,p_0)_\omega)\int_\omega \frac{\mathcal{F}(q,p,t;q_a,p_a,t_0)}{v(q_a,p_a)}ds_{a\omega}. \tag{3.20}$$

Integral u definiciji funkcije $G(q,p,t;(q_0,p_0)_\omega,t_0)$ u jednadžbi (3.20) je duž cijele duljine putanje ω koja je okomito presječena plohom $S_0(M)$ u točki $(q_0,p_0)_\omega$. Korištenjem jednadžbe (3.20), relacija (3.19) se potom može zapisati kao

$$f(q,p,t) = \int_{S_0(M)} G(q,p,t;(q_0,p_0)_\omega,t_0)dS_0. \tag{3.21}$$

Jasno je da izraz

$$G(q,p,t;(q_0,p_0)_\omega,t_0)dS_0d\Gamma \equiv \mathrm{P}(q,p,t\cap(q_0,p_0)_\omega,t_0), \tag{3.22}$$

predstavlja vjerojatnost da je točka koja korespondira stanju sustava u trenutku t_0 bilo gdje duž putanja koje prolaze kroz infinitezimalni element dS_0 oko (q_0,p_0) na plohi $S_0(M)$, i da je u nekom različitom trenutku t u elementu volumena $d\Gamma$ oko (q,p).

Još jedna definicija raspodjele vjerojatnosti putanja $F((q_0, p_0)_\omega, t_0)$, uz jednadžbu (3.12), sada je moguća na ovaj način. Ona je dana integralom

$$F((q_0, p_0)_\omega, t_0) = \int_\Gamma G(q, p, t; (q_0, p_0)_\omega, t_0) d\Gamma. \tag{3.23}$$

Tada, u skladu s teorijom vjerojatnosti, kvocijent

$$\frac{G(q, p, t; (q_0, p_0)_\omega, t_0) dS_0 d\Gamma}{F((q_0, p_0)_\omega, t_0) dS_0} \equiv \mathrm{P}(q, p, t | (q_0, p_0)_\omega, t_0), \tag{3.24}$$

predstavlja *uvjetnu vjerojatnost* da je u trenutku t točka koja korespondira stanju sustava u elementu $d\Gamma$ oko (q, p), ako je u trenutku t_0 bilo gdje duž putanja koje prolaze kroz infinitezimalni element dS_0 oko (q_0, p_0) na plohi $S_0(M)$. Relacija (3.23) tada dokazuje da integral izraza (3.24) po Γ zadovoljava uvjet normiranosti, tj.,

$$\int_\Gamma \frac{G(q, p, t; (q_0, p_0)_\omega, t_0) dS_0}{F((q_0, p_0)_\omega, t_0) dS_0} d\Gamma = 1. \tag{3.25}$$

Zbog pripreme potrebnih pojmova teorije vjerojatnosti, uvjetna raspodjela vjerojatnosti $D(q, p, t | (q_0, p_0)_\omega, t_0)$ koja odgovara uvjetnoj vjerojatnosti (3.24), definira se relacijom

$$D(q, p, t | (q_0, p_0)_\omega, t_0) = \frac{G(q, p, t; (q_0, p_0)_\omega, t_0)}{F((q_0, p_0)_\omega, t_0)}. \tag{3.26}$$

Relacija (3.24), poput relacije (3.22), predstavlja vjerojatnost, a ona je sačuvana u faznom prostoru Γ. Totalna vremenska derivacija (tj., brzina vremenske promjene uzduž linija Hamiltonovog toka) ove vjerojatnosti je jednaka nuli. U relaciji (3.24) za uvjetnu vjerojatnost, raspodjela vjerojatnosti putanja $F((q_0, p_0)_\omega, t_0)$ i element plohe dS_0 su neovisni o varijablama t i (q, p). Također, mjera $d\Gamma$ je invarijantna na Hamiltonovo gibanje. Dakle, slijedi da je totalna vremenska derivacija uvjetne vjerojatnosti (3.24) jednaka nuli ako i samo ako

$$\frac{dG}{dt} \equiv \frac{\partial G}{\partial t} + \sum_{i=1}^{s} \left(\frac{\partial G}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial G}{\partial p_i} \frac{\partial H}{\partial q_i} \right) = 0\ . \tag{3.27}$$

To je jednostavna demonstracija da raspodjela vjerojatnosti $G(q, p, t; (q_0, p_0)_\omega, t_0)$ zadovoljava jednadžbu analognu Liouvilleovoj jednadžbi (3.14) za raspodjelu vjerojatnosti mikrostanja $f(q, p, t)$.

Poglavlje 4

MODEL VREMENSKE EVOLUCIJE ZATVORENIH HAMILTONOVIH SUSTAVA

4.1 Informacijske entropije i MaxEnt algoritam

U Shannonovoj teoriji informacije [4] veličini oblika $H = -\sum_i p_i \log p_i$ dana je centralna uloga mjere informacije, izbora i nesigurnosti za različite raspodjele vjerojatnosti p_i. Na analogan način Shannon je definirao entropiju neprekidne raspodjele i entropiju N-dimenzionalne neprekidne raspodjele. S druge strane, Jaynes [5] je pokazao da veličini $-\sum_{i=1}^{n} p_i \log p_i$ za diskretnu raspodjelu vjerojatnosti p_i koja teži u limesu beskonačnog broja točaka u neprekidnu raspodjelu $w(x)$ (na način da se gustoća točaka, podijeljena s ukupnim brojem točaka, približava određenoj funkciji $m(x)$) odgovara veličina:

$$S_I = -\int w(x) \log \left[\frac{w(x)}{m(x)} \right] dx. \tag{4.1}$$

Shannon je pretpostavio da je ta veličina oblika $-\int w(x) \log[w(x)] dx$, ali je također istaknuo i važnu razliku između svojih definicija diskretnih i neprekidnih entropija. Ako promijenimo koordinate, entropija neprekidne raspodjele će se promijeniti na način koji je Shannon uzeo u obzir [4]. Invarijantnost entropije neprekidne raspodjele na zamjenu nezavisne varijable postiže se modifikacijom koja slijedi iz matematičke dedukcije koju je proveo Jaynes [5]. Opisani granični postupak uz definiciju funkcije mjere $m(x)$ rezultira invarijantnom mjerom informacije (4.1). Ako se pretpostavi da je funkcija mjere jednolika $m(x) = konst$, tada se invarijantna mjera informacije (4.1) razlikuje se od Shannonove definicije entropije neprekidne raspodjele [4] za irelevantnu aditivnu konstantu.

Shannon [4] je također definirao zajedničku entropiju i uvjetnu entropiju zajedničke raspodjele dvije neprekidne varijable (koje mogu same biti višedimenzionalne). U prethodnom odjeljku je bila uvedena *zajednička raspodjela* $G(q,p,t;(q_0,p_0)_\omega,t_0)$ dvije neprekidne višedimenzionalne varijable $(q,p)\in\Gamma$ i $(q_0,p_0)_\omega\in S_0(M)$. Iz detaljnog objašnjenja izraza (3.22) slijedi da $G(q,p,t;(q_0,p_0)_\omega,t_0)dS_0d\Gamma$ predstavlja vjerojatnost zajedničkog pojavljivanja dva događaja: prvi se događa u trenutku t_0 na skupu $\Omega(M)$ svih mogućih putanja u faznom prostoru, a drugi se događa u trenutku t na skupu M svih mogućih točaka u faznom prostoru, skupu koji je invarijantan na Hamiltonovo gibanje. U skladu s Shannonovom definicijom [4], *zajednička informacijska entropija* raspodjele $G(q,p,t;(q_0,p_0)_\omega,t_0)$ je dana s

$$S_I^G(t,t_0)=-\int_{S_0(M)}\int_\Gamma G\log G\ d\Gamma dS_0. \tag{4.2}$$

Notacija $S_I^G(t,t_0)$ označava da je riječ o funkciji varijabli t i t_0, kroz raspodjelu $G\equiv G(q,p,t;(q_0,p_0)_\omega,t_0)$. Slijedeći Shannonovu definiciju [4], *uvjetna informacijska entropija* zajedničke raspodjele $G(q,p,t;(q_0,p_0)_\omega,t_0)$ je dana s

$$S_I^{DF}(t,t_0)=-\int_{S_0(M)}\int_\Gamma G\log\left[\frac{G}{F}\right]\ d\Gamma dS_0, \tag{4.3}$$

gdje je $F\equiv F((q_0,p_0)_\omega,t_0)$ raspodjela vjerojatnosti putanja. Korištenjem definicije $D(q,p,t|(q_0,p_0)_\omega,t_0)$ dane jednadžbom (3.26), neposredno se dobije ekvivalentan oblik uvjetne informacijske entropije (4.3):

$$S_I^{DF}(t,t_0)=-\int_{S_0(M)}\int_\Gamma DF\log D\ d\Gamma dS_0. \tag{4.4}$$

Iz izraza (4.4) jasno je da je uvjetna informacijska entropija $S_I^{DF}(t,t_0)$ prosjek entropije uvjetne vjerojatnosti $D(q,p,t|(q_0,p_0)_\omega,t_0)$, uzet preko svih mogućih putanja $\omega\in\Omega(M)$ u skladu s težinom danom raspodjelom vjerojatnosti putanja $F((q_0,p_0)_\omega,t_0)$.

Relacija između informacijskih entropija $S_I^G(t,t_0)$ i $S_I^{DF}(t,t_0)$, uvedenih u jednadžbama (4.2) i (4.3), upotpunjena je uvođenjem informacijske entropije raspodjele $F((q_0,p_0)_\omega,t_0)$, ili alternativno, *informacijske entropije putanja*:

$$S_I^F(t_0)=-\int_{S_0(M)}F\log F\ dS_0. \tag{4.5}$$

Relacija između $S_I^G(t,t_0)$, $S_I^{DF}(t,t_0)$ i $S_I^F(t_0)$ dobije se jednostavno, korištenjem jednadžbe (3.26) u (4.2), i primjenom svojstava raspodjela vjerojatnosti. Na taj način se dobije

$$S_I^G(t,t_0)=S_I^{DF}(t,t_0)+S_I^F(t_0). \tag{4.6}$$

Relacija (4.6), u skladu s analognom relacijom u Shannonovoj teoriji [4], tvrdi da je nesigurnost (entropija) zajedničkog događaja jednaka zbroju nesigurnosti prvog i nesigurnosti drugog događaja kad je prvi poznat.

Važno je dati neke dodatne komentare uz jednadžbu (4.6). Općenito, nesigurnost zajedničkog događaja je manja od ili je jednaka zbroju nesigurnosti dva pojedinačna događaja, s jednakošću ako (i samo ako) su dva događaja nezavisna [4]. Raspodjela vjerojatnosti zajedničkog događaja je ovdje dana s $G(q, p, t; (q_0, p_0)_\omega, t_0)$. Informacijska entropija ili nesigurnost jednog od njih (u ovom slučaju nazvanim drugi događaj zbog njegove pojave u kasnijem trenutku) je jednaka

$$S_I^f(t) = -\int_\Gamma f \log f \; d\Gamma. \tag{4.7}$$

Veličina $S_I^f(t)$ je informacijska entropija raspodjele vjerojatnosti mikrostanja $f(q, p, t)$, ili kraće, *informacijska entropija*. Nesigurnost prvog događaja dana je informacijskom entropijom putanja $S_I^F(t_0)$ definiranom u jednadžbi (4.5). Prethodno spomenuto svojstvo informacijskih entropija ovdje je dano za $S_I^G(t, t_0)$, $S_I^f(t)$ i $S_I^F(t_0)$ slijedećom relacijom:

$$S_I^G(t, t_0) \leq S_I^f(t) + S_I^F(t_0), \tag{4.8}$$

s jednakošću ako (i samo ako) su dva događaja nezavisna. Osim toga, iz (4.6) i (4.8) se dobija važna relacija između $S_I^f(t)$ i $S_I^{DF}(t, t_0)$:

$$S_I^f(t) \geq S_I^{DF}(t, t_0), \tag{4.9}$$

s jednakošću ako (i samo ako) su dva događaja nezavisna.

U smislu vjerojatnosti, događaji u trenutku t_0 u skupu $\Omega(M)$ svih mogućih putanja u faznom prostoru i u bilo kojem trenutku t u skupu $M \subset \Gamma$ svih mogućih točaka u faznom prostoru, nisu nezavisni. Ako pretpostavimo da su vrijednosti zajedničke raspodjele vjerojatnosti $G(q, p, t; (q_0, p_0)_\omega, t_0)$ fizikalno dobro definirane (u smislu jednadžbe (3.22)) za sve točke $(q, p) \in \Gamma$ i $(q_0, p_0) \in S_0(M)$ u danom početnom trenutku $t = t_0$, onda su njene vrijednosti određene za sve t u cijelom faznom prostoru Γ preko Liouvilleove jednadžbe (3.27). Jednostavnom dedukcijom dolazi se do zaključka da se maksimizacijom uvjetne informacijske entropije $S_I^{DF}(t, t_0)$, uz ograničenja dana Liouvilleovom jednadžbom (3.27) i uvjetom normiranosti, ne može doseći gornja granica koja je dana (u bilo kojem trenutku t) vrijednošću informacijske entropije $S_I^f(t)$ u relaciji (4.9). Dosezanje ove gornje granice bi zahtijevalo statističku nezavisnost, što bi za logičku posljedicu imalo potpuni gubitak korelacije između skupa $\Omega(M)$ mogućih putanja u faznom prostoru u trenutku t_0 i skupa $M \subset \Gamma$ mogućih točaka u faznom prostoru u trenutku t. Međutim, statistička nezavisnost isključena je u bilo kojem trenutku t ograničenjem koje je implicirano Liouvilleovom jednadžbom (3.27) i zahtjevom da zajednička raspodjela vjerojatnosti $G(q, p, t; (q_0, p_0)_\omega, t_0)$ bude dobro definirana.

Ovdje može biti od pomoći razlučivanje između dva aspekta vremenske evolucije. Prvi je mikroskopski aspekt koji predstavlja problem dinamike impliciran u ovom radu Hamiltonovim jednadžbama. Rješenja su reprezentirana u Γ kao putanje u faznom prostoru.

S druge strane, predviđanje makroskopske vremenske evolucije predstavlja problem dostupne informacije i inferencije, odnosno izvođenja i donošenja zaključaka, iz parcijalne informacije. Zato su, i mikroskopska dinamika i odgovarajuće putanje u faznom prostoru također dio ovog problema nepotpune informacije. U slučaju makroskopskog sustava je informacija o mikroskopskoj dinamici vrlo vjerojatno nepotpuna zbog niza različitih mogućih razloga. Neki od njih će posebno biti analizirani u odjeljku 5.3. Detaljno znanje o mikroskopskim trajektorijama, odnosno o stanjima kroz koja trajektorije prolaze općenito nam nedostaje u slučaju makroskopskih sustava. Međutim, u nedostatku potpunijeg znanja, Hamiltonove jednadžbe (3.1) i skup mogućih putanja u faznom prostoru su reprezentacija naše prethodne informacije o mikroskopskoj dinamici. Prirodno je pretpostaviti da je makroskopska vremenska evolucija koju predviđamo i tada konzistentna s našim znanjem o mikroskopskoj dinamici, čak i kad to znanje nije potpuno.

Svi argumenti koji su ranije spomenuti vode do zaključka da je razmatranje Liouvilleove jednadžbe (3.27) kao strogog *mikroskopskog ograničenja* na vremensku evoluciju, u smislu predikcije ekvivalentno posjedovanju potpune informacije o mikroskopskoj dinamici. Slijedeći prethodno navedene pretpostavke, Liouvilleova jednadžba (3.27) se također može razmatrati i kao *makroskopsko ograničenje* na vremensku evoluciju. Ako naša informacija o mikroskopskoj dinamici nije dovoljno iscrpna da se detaljno odredi vremenska evolucija, uobičajeno je uzeti prosjek po svim slučajevima koji su mogući na osnovi parcijalne informacije. U prediktivnoj statističkoj mehanici koju je formulirao Jaynes, zaključci se izvode iz raspodjela vjerojatnosti čiji prostori elementarnih događaja reprezentiraju poznato o strukturi mikrostanja, i koje maksimiziraju informacijsku entropiju uz dostupne makroskopske podatke kao ograničenja [8]. Na ovaj način "objektivnost" pridjeljivanja vjerojatnosti i predikcija je sigurna od uvođenja dodatnih pretpostavki koje nisu nužno sadržane u dostupnim podacima. U ovom modelu se ista osnovna ideja uvodi u fazu (2) problema predikcije za zatvorene Hamiltonove sustave. Uvjetna informacijska entropija $S_I^{DF}(t, t_0)$ se tada maksimizira uz ograničenje dano Liouvilleovom jednadžbom (3.27), uvedeno kao prosjek po faznom prostoru, ili preciznije, integral po faznom prostoru slično drugim makroskopskim ograničenjima. Ovaj pristup nam omogućuje da razmotrimo nepotpunu prirodu naše informacije o mikroskopskoj dinamici na racionalan način, i vodi do gubitka korelacije između početnih putanja u faznom prostoru i konačnih mikrostanja i do odgovarajuće nesigurnosti u predikciji. Uvjetna informacijska entropija $S_I^{DF}(t, t_0)$ je mjera te nesigurnosti, povezane s gubitkom informacije o stanju sustava.

4.2 MaxEnt inferencije i vremenska evolucija

U prvom pristupu koji će biti sada razmotren, vremenska evolucija uvjetne raspodjele vjerojatnosti $D(q, p, t|(q_0, p_0)_\omega, t_0)$ u intervalu $t_0 \leq t \leq t_a$ treba zadovoljavati slijedeća ograničenja: uvjet normiranosti

$$\int_M D(q, p, t|(q_0, p_0)_\omega, t_0) d\Gamma = 1, \tag{4.10}$$

i Liouvilleovu jednadžbu za $D(q, p, t|(q_0, p_0)_\omega, t_0)$,

$$\frac{\partial D}{\partial t} + \sum_{i=1}^{s} \left(\frac{\partial D}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial D}{\partial p_i} \frac{\partial H}{\partial q_i} \right) = 0. \tag{4.11}$$

Iz relacije (3.26) slijedi da su ograničenja dana jednadžbama (3.27) i (4.11) ekvivalentna. Prema definiciji uvedenoj u odjeljku 3.2, skup $M \subset \Gamma$ svih mogućih mikrostanja je invarijantan na Hamiltonovo gibanje. Ograničenje normiranosti (4.10) sadrži informaciju o strukturi mogućih mikrostanja u Γ, u vremenskom intervalu $t_0 \leq t \leq t_a$ koji se razmatra. Informacija o mikroskopskoj dinamici je reprezentirana skupom $\Omega(M)$ svih mogućih putanja u Γ. Dodatno je ta informacija je sadržana u Liouvilleovoj jednadžbi (4.11). Pridružena raspodjela vjerojatnosti putanja $F((q_0, p_0)_\omega, t_0)$ je kompatibilna s dostupnom informacijom.

Vremenska derivacija uvjetne informacijske entropije $S_I^{DF}(t, t_0)$ dane jednadžbom (4.4) je jednaka

$$\frac{dS_I^{DF}(t, t_0)}{dt} = -\int_{S_0(M)} \int_M \frac{\partial D}{\partial t} F \log D \; d\Gamma dS_0 - \int_{S_0(M)} \int_M \frac{\partial D}{\partial t} F \; d\Gamma dS_0. \tag{4.12}$$

Zbog uvjeta normiranosti (4.10), posljednji član u (4.12) je jednak nuli. U trenutku t_a uvjetna informacijska entropija $S_I^{DF}(t_a, t_0)$ je dana izrazom,

$$S_I^{DF}(t_a, t_0) = -\int_{t_0}^{t_a} \int_{S_0(M)} \int_M \frac{\partial D}{\partial t} F \log D \; d\Gamma dS_0 dt + S_I^{DF}(t_0, t_0). \tag{4.13}$$

Prikladno je formirati slijedeći funkcional

$$J[D] = S_I^{DF}(t_a, t_0) - S_I^{DF}(t_0, t_0) = \int_{t_0}^{t_a} \int_{S_0(M)} \int_M K(D, \partial_t D) d\Gamma dS_0 dt, \tag{4.14}$$

s funkcijom $K(D, \partial_t D)$ danom s

$$K(D, \partial_t D) = -\frac{\partial D}{\partial t} F \log D. \tag{4.15}$$

U varijacijskom problemu koji se ovdje razmatra, traži se stacionarnost funkcionala $J[D]$ danog jednadžbom (4.14) s obzirom na varijacije uz ograničenja (4.10) i (4.11).

Ne zahtjeva se da na rubu područja integracije $M \times (t_0, t_a)$ u integralu (4.14) funkcija $D(q,p,t|(q_0,p_0)_\omega,t_0)$ poprima propisane vrijednosti. Ograničenja dana jednadžbama (4.10) i (4.11) su zapisana ovdje u ekvivalentnom, ali prikladnijem obliku:

$$\varphi_1((q_0,p_0)_\omega,t_0;t,D) = F\int_M D\; d\Gamma - F = 0, \tag{4.16}$$

i

$$\varphi_2((q_0,p_0)_\omega,t_0;q,p,t,\partial_q D,\partial_p D,\partial_t D) =$$

$$= \left[\frac{\partial D}{\partial t} + \sum_{i=1}^{s}\left(\frac{\partial D}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial D}{\partial p_i}\frac{\partial H}{\partial q_i}\right)\right] F = 0. \tag{4.17}$$

Metode za varijacijske probleme s ovakvim vrstama ograničenja postoje i mogu se razviti i primjeniti u praktičnim problemima, što je detaljno objašnjeno u referenci [32]. Ovdje se, u notaciji koja je prilagođena ovom specifičnom problemu, uvode slijedeći funkcionali:

$$C_1[D,\lambda_1] = \int_{S_0(M)}\int_{t_0}^{t_a} \lambda_1\varphi_1\; dt dS_0, \tag{4.18}$$

i

$$C_2[D,\lambda_2] = \int_{S_0(M)}\int_{t_0}^{t_a}\int_M \lambda_2\varphi_2\; d\Gamma dt dS_0. \tag{4.19}$$

Lagrangeovi multiplikatori $\lambda_1 \equiv \lambda_1((q_0,p_0)_\omega,t_0;t)$ i $\lambda_2 \equiv \lambda_2((q_0,p_0)_\omega,t_0;q,p,t)$ su funkcije koje su definirane u područjima integracije u (4.18) i (4.19). Za bilo koju funkciju s neprekidnim prvim parcijalnim derivacijama, Eulerova jednadžba za ograničenje $\varphi_2 \equiv \varphi_2((q_0,p_0)_\omega,t_0;q,p,t,\partial_q D,\partial_p D,\partial_t D)$ je jednaka nuli. Slijedeći najopćenitije pravilo multiplikatora za ovu vrstu problema koje je detaljno objašnjeno u referenci [32], uvodimo dodatni konstantni Lagrangeov multiplikator λ_0 za funkciju K,

$$J[D,\lambda_0] = \int_{t_0}^{t_a}\int_{S_0(M)}\int_M \lambda_0 K(D,\partial_t D)\; d\Gamma dS_0 dt. \tag{4.20}$$

Funkcional $I[D,\lambda_0,\lambda_1,\lambda_2]$ se formira od $J[D,\lambda_0]$, $C_1[D,\lambda_1]$ i $C_2[D,\lambda_2]$:

$$I[D,\lambda_0,\lambda_1,\lambda_2] = J[D,\lambda_0] - C_1[D,\lambda_1] - C_2[D,\lambda_2]. \tag{4.21}$$

Egzistencija Lagrangeovih multiplikatora $\lambda_0 \neq 0$, λ_1 i λ_2, takvih da je varijacija $I[D,\lambda_0,\lambda_1,\lambda_2]$ stacionarna $\delta I = 0$, predstavlja dokaz da je moguće učiniti $J[D]$ u jednadžbi (4.14) stacionarnim uz ograničenja (4.16) i (4.17). Funkcija $D(q,p,t|(q_0,p_0)_\omega,t_0)$ koja čini $J[D]$ stacionarnim uz ograničenja (4.16) i (4.17) mora zadovoljavati Eulerovu jednadžbu:

$$\lambda_0 \left\{ \frac{\partial K}{\partial D} - \frac{d}{dt}\left(\frac{\partial K}{\partial(\partial_t D)}\right) - \sum_{i=1}^{s}\left[\frac{d}{dq_i}\left(\frac{\partial K}{\partial(\partial_{q_i} D)}\right) + \frac{d}{dp_i}\left(\frac{\partial K}{\partial(\partial_{p_i} D)}\right)\right]\right\}$$

$$-\lambda_1 F + \left[\frac{\partial \lambda_2}{\partial t} + \sum_{i=1}^{s}\left(\frac{\partial \lambda_2}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial \lambda_2}{\partial p_i}\frac{\partial H}{\partial q_i}\right)\right] F = 0. \tag{4.22}$$

Lako se može provjeriti da je član pomnožen s λ_0 u Eulerovoj jednadžbi (4.22) jednak nuli. Stacionarnost funkcionala $I[D, \lambda_0, \lambda_1, \lambda_2]$ u (4.21) je zato moguća i s $\lambda_0 \neq 0$. Iz jednadžbe (4.22) slijedi da Lagrangeovi multiplikatori λ_1 i λ_2 zadovoljavaju jednadžbu

$$\frac{\partial \lambda_2}{\partial t} + \sum_{i=1}^{s}\left(\frac{\partial \lambda_2}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial \lambda_2}{\partial p_i}\frac{\partial H}{\partial q_i}\right) = \lambda_1. \tag{4.23}$$

U ovom varijacijskom problemu, od funkcije $D(q, p, t|(q_0, p_0)_\omega, t_0)$ koja čini $J[D]$ u (4.14) stacionarnim uz ograničenja (4.16) i (4.17), ne zahtjeva se da poprima propisane vrijednosti na rubu područja integracije $M \times (t_0, t_a)$. Zato je nužno, da uz zadovoljavanje Eulerove jednadžbe (4.22), ona također zadovoljava i Eulerov rubni uvjet na granici $M \times (t_0, t_a)$, kao što je to objašnjeno u referenci [32]. Za sve točke na dijelu granice područja integracije $M \times (t_0, t_a)$ gdje je $t = t_0$ ili $t = t_a$, Eulerov rubni uvjet nezavisno daje:

$$\left[\frac{\partial K}{\partial(\partial_t D)} - \lambda_2 F\right]_{t=t_0,t_a} = -\left[\log D + \lambda_2\right]_{t=t_0,t_a} F = 0. \tag{4.24}$$

Za sve točke na dijelu granice područja integracije $M \times (t_0, t_a)$ gdje je vrijeme t u intervalu $t_0 < t < t_a$, isti Eulerov rubni uvjet daje:

$$F\left[\lambda_2 \mathbf{v} \cdot \mathbf{n}\right]_{\text{na rubu } M} = 0. \tag{4.25}$$

U jednadžbi (4.25), $\mathbf{v} \cdot \mathbf{n}$ je skalarni produkt polja brzine $\mathbf{v}(q, p)$ u faznom prostoru Γ (definiranog u odjeljku 3.1) i jedinične normale $\mathbf{n}$ rubne plohe invarijantnog skupa M, uzet na toj plohi. Jednadžba (4.25) je zadovoljena prirodno zbog Hamiltonovog gibanja, budući da je skup M invarijantan po definiciji, i zato je $\mathbf{v} \cdot \mathbf{n} = 0$ za sve točke na rubnoj plohi od M. To je posljedica činjenice da putanje u faznom prostoru ne prelaze preko rubne plohe invarijantnog skupa M.

Funkcije $D(q, p, t|(q_0, p_0)_\omega, t_0)$ koje čine $J[D]$ u (4.14) stacionarnim uz ograničenja (4.16) i (4.17), određuju se iz tih ograničenja i rubnog uvjeta danog jednadžbom (4.24). Iz jednadžbe (4.24) dobije se oblik $D(q, p, t|(q_0, p_0)_\omega, t_0)$ u vremenskim trenucima t_0 i t_a,

$$D(q, p, t|(q_0, p_0)_\omega, t_0)\big|_{t=t_0,t_a} = \exp\left[-\lambda_2((q_0, p_0)_\omega, t_0; q, p, t)\right]\big|_{t=t_0,t_a}. \tag{4.26}$$

Budući da se zahtjeva jedino da je $t_a \geq t_0$, odabir vremena t_a proizvoljan je u svim drugim aspektima. Rubni uvjet (4.24) tada vrijedi za bilo koje vrijeme $t \geq t_0$:

$$D(q, p, t|(q_0, p_0)_\omega, t_0) = \exp\left[-\lambda_2((q_0, p_0)_\omega, t_0; q, p, t)\right]. \tag{4.27}$$

Iz ograničenja (4.17), korištenjem jednadžbe (4.27), izravno se dobije jednadžba za Lagrangeov multiplikator $\lambda_2((q_0, p_0)_\omega, t_0; q, p, t)$:

$$\frac{\partial \lambda_2}{\partial t} + \sum_{i=1}^{s} \left(\frac{\partial \lambda_2}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial \lambda_2}{\partial p_i} \frac{\partial H}{\partial q_i} \right) = 0. \tag{4.28}$$

Usporedbom jednadžbe (4.23) s jednadžbom (4.28), slijedi da je za sve $t \geq t_0$,

$$\lambda_1((q_0, p_0)_\omega, t_0; t) = 0. \tag{4.29}$$

Pravi smisao Lagrangeovog multiplikatora $\lambda_1((q_0, p_0)_\omega, t_0; t)$ postaje puno jasniji kad se promotri u sklopu drugačijeg pristupa. Kao što je objašnjeno u odjeljku 4.1, uvjetna raspodjela vjerojatnosti $D(q, p, t|(q_0, p_0)_\omega, t_0)$ mora biti fizikalno dobro definirana u smislu jednadžbi (3.24) i (3.26). Za bilo koju takvu funkciju, gornja granica za $S_I^{DF}(t, t_0)$ dana relacijom (4.9), ne doseže se maksimizacijom uz ograničenja (4.16) i (4.17).

Zaključci koji slijede iz interpretacije relacije (4.9) i svojstva $S_I^{DF}(t, t_0)$ kao mjere nesigurnosti povezane s gubitkom informacije, argumentirani su u odjeljku 4.1. Ovi zaključci se sada uzimaju u obzir kao osnova za drugačiji pristup u skladu s temeljnim načelima prediktivne statističke mehanike. Kao što je detaljno bilo obrazloženo u odjeljku 4.1, ograničenje koje ima strogi oblik jednakosti (4.17), sada će biti zamijenjeno s ograničenjem u obliku integrala po faznom prostoru,

$$\varphi_2((q_0, p_0)_\omega, t_0; t, D) = \int_M \left[\frac{\partial D}{\partial t} + \sum_{i=1}^{s} \left(\frac{\partial D}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial D}{\partial p_i} \frac{\partial H}{\partial q_i} \right) \right] F \, d\Gamma = 0. \tag{4.30}$$

Ograničenje (4.30) je jednim dijelom u obliku izoperimetrijskog ograničenja, premda ne u potpunosti. Funkcional (4.19) treba zatim zamijeniti odgovarajućim funkcionalom, koji je dan integralima preostalim nakon integrala (4.30) po faznom prostoru,

$$C_2[D, \lambda_2] = \int_{S_0(M)} \int_{t_0}^{t_a} \lambda_2 \varphi_2 \, dt dS_0. \tag{4.31}$$

Lagrangeov multiplikator $\lambda_2 \equiv \lambda_2((q_0, p_0)_\omega, t_0; t)$ je sada funkcija definirana u području integracije u integralu (4.31).

Informacija da je skup M mogućih mikrostanja u faznom prostoru Γ invarijantan na Hamiltonovo gibanje sadržana je u ograničenju (4.30). Analogija s prethodnim pristupom nije potpuna, jer mnogo šira klasa funkcija zadovoljava ograničenje (4.30), uključujući u to i sve funkcije koje zadovoljavaju ograničenje (4.17). Ova činjenica omogućuje maksimizaciju uvjetne informacijske entropije $S_I^{DF}(t_a, t_0)$, uz ograničenja (4.16) i (4.30), čak ako je $D(q, p, t|(q_0, p_0)_\omega, t_0)$ propisana u početnom trenutku t_0. U ovom varijacijskom problemu, od funkcije $D(q, p, t|(q_0, p_0)_\omega, t_0)$ se ne zahtjeva da poprima propisane vrijednosti na preostalom dijelu ruba područja integracije $M \times (t_0, t_a)$ u integralu (4.13). Funkcija

$D(q,p,t|(q_0,p_0)_\omega,t_0)$ je propisana jedino u početnom trenutku t_0, i pri tome mora biti fizikalno dobro definirana u smislu jednadžbi (3.24) i (3.26).

Da bi funkcija $D(q,p,t|(q_0,p_0)_\omega,t_0)$ maksimizirala $S_I^{DF}(t_a,t_0)$ uz ograničenja (4.16) i (4.30), nužno je da zadovoljava Eulerovu jednadžbu:

$$\lambda_0 \left\{ \frac{\partial K}{\partial D} - \frac{d}{dt}\left(\frac{\partial K}{\partial(\partial_t D)}\right) - \sum_{i=1}^{s}\left[\frac{d}{dq_i}\left(\frac{\partial K}{\partial(\partial_{q_i} D)}\right) + \frac{d}{dp_i}\left(\frac{\partial K}{\partial(\partial_{p_i} D)}\right)\right]\right\}$$
$$-\lambda_1 F + \frac{\partial \lambda_2}{\partial t} F = 0. \tag{4.32}$$

Još jedan nužni uvjet za maksimum, uz jednadžbu (4.32), postoji ako se od funkcije $D(q,p,t|(q_0,p_0)_\omega,t_0)$ ne zahtjeva da poprima propisane vrijednosti na dijelu ruba $M \times (t_0,t_a)$: tada je nužno i da $D(q,p,t|(q_0,p_0)_\omega,t_0)$ zadovoljava Eulerov rubni uvjet na dijelu ruba $M \times (t_0,t_a)$ gdje njene vrijednosti nisu propisane, ref. [32]. U skladu s tim, za sve točke na dijelu ruba $M \times (t_0,t_a)$ gdje je $t = t_a$, Eulerov rubni uvjet daje:

$$\left[\frac{\partial K}{\partial(\partial_t D)} - \lambda_2 F\right]_{t=t_a} = -\left[\log D + \lambda_2\right]_{t=t_a} F = 0. \tag{4.33}$$

Eulerov rubni uvjet je prirodno zadovoljen za sve točke na dijelu ruba $M \times (t_0,t_a)$ gdje je vrijeme t u intervalu $t_0 < t < t_a$. Skup M je invarijantan na Hamiltonovo gibanje, i jednadžba analogna jednadžbi (4.25) ovdje je također prirodno zadovoljena zbog Hamiltonovog gibanja.

Na način analogan onome koji je u prvom pristupu vodio do jednadžbe (4.23), sada iz Eulerove jednadžbe (4.32) slijedi jednadžba za Lagrangeove multiplikatore $\lambda_1((q_0,p_0)_\omega,t_0;t)$ i $\lambda_2((q_0,p_0)_\omega,t_0;t)$:

$$\frac{\partial \lambda_2}{\partial t} = \lambda_1. \tag{4.34}$$

Oblik MaxEnt uvjetne raspodjele vjerojatnosti u trenutku t_a slijedi iz jednadžbe (4.33):

$$D(q,p,t_a|(q_0,p_0)_\omega,t_0) = \exp\left[-\lambda_2((q_0,p_0)_\omega,t_0;t_a)\right]. \tag{4.35}$$

Za svaku, u početnom trenutku t_0 dobro definiranu uvjetnu raspodjelu vjerojatnosti, postoji cijela klasa jednako vjerojatnih rješenja $\{D(q,p,t|(q_0,p_0)_\omega,t_0)\}$ dobivenih MaxEnt algoritmom, od kojih svako zadovoljava makroskopsko ograničenje (4.30). U trenutku t_a, sve funkcije u toj klasi MaxEnt rješenja međusobno su jednake i dane jednadžbom (4.35). Dakle, osim u rubnim vremenskim trenucima t_0 i t_a, uvjetna raspodjela vjerojatnosti dobivena MaxEnt algoritmom nije jedinstveno određena u intervalu $t_0 < t < t_a$. To je posljedica činjenice da makroskopsko ograničenje (4.30) ne određuje vremensku evoluciju $D(q,p,t|(q_0,p_0)_\omega,t_0)$ jedinstveno, kao što to čini strogo mikroskopsko ograničenje (4.17). Međutim, MaxEnt rješenja još uvijek predviđaju jedino vremenske evolucije u potpunosti unutar invarijantnog skupa M, zbog jednadžbe (4.25). To svojstvo slijedi iz ograničenja (4.30), i u obzir uzima informaciju o konstantama gibanja koje određuju invarijantan skup M, i na taj način, o odgovarajućim zakonima sačuvanja.

Iz normiranosti (4.10) uvjetne raspodjele vjerojatnosti, dane u trenutku t_a jednadžbom (4.35), dobije se relacija:

$$W(M) \exp\left[-\lambda_2((q_0,p_0)_\omega, t_0; t_a)\right] = 1, \tag{4.36}$$

gdje je $W(M)$ mjera, tj., volumen invarijantnog skupa M u faznom prostoru. Jednadžba (4.36) implicira da Lagrangeov multiplikator $\lambda_2((q_0,p_0)_\omega, t_0; t)$ u trenutku $t = t_a$ ne ovisi o varijablama $(q_0,p_0)_\omega$:

$$\lambda_2((q_0,p_0)_\omega, t_0; t_a) = \lambda_2(t_a). \tag{4.37}$$

Raspodjela vjerojatnosti mikrostanja $f(q,p,t)$ u trenutku $t = t_a$ se onda izračuna korištenjem: jednadžbi (3.21) i (3.26), MaxEnt uvjetne raspodjele vjerojatnosti $D(q,p,t|(q_0,p_0)_\omega, t_0)$ u trenutku $t = t_a$ dane jednadžbama (4.35) i (4.37), i raspodjele vjerojatnosti putanja $F((q_0,p_0)_\omega, t_0)$ u početnom trenutku t_0:

$$f(q,p,t_a) = \exp\left[-\lambda_2(t_a)\right]. \tag{4.38}$$

Iz jednadžbi (4.35–4.38) slijedi da su u trenutku t_a, MaxEnt uvjetna raspodjela vjerojatnosti i odgovarajuća raspodjela vjerojatnosti mikrostanja jednake,

$$D(q,p,t_a|(q_0,p_0)_\omega, t_0) = f(q,p,t_a) = \exp\left[-\lambda_2(t_a)\right] = \frac{1}{W(M)}. \tag{4.39}$$

Iz jednadžbi (4.4), (4.7) i (4.39), dobiju se vrijednosti informacijskih entropija $S_I^{DF}(t,t_0)$ i $S_I^f(t)$ u trenutku t_a,

$$S_I^f(t_a) = S_I^{DF}(t_a,t_0) = \log W(M). \tag{4.40}$$

Jednakosti (4.39) i (4.40) su moguće jedino u slučaju *statističke nezavisnosti.* Logička posljedica statističke nezavisnosti je *potpuni gubitak korelacije* između putanja u faznom prostoru u trenutku t_0, i mikrostanja u trenutku t_a. Općenito gledajući, svojstvo makrosopskih sustava je da njihovo ponašanje kao da postaje nasumično između opažanja, uz uvjet da opažanja slijede jedno za drugim u vremenskim intervalima duljim od određenog karakterističnog vremena τ nazvanog vremenom relaksacije [33]. U intepretaciji koja se daje ovdje, vrijeme relaksacije τ za zatvoreni Hamiltonov sustav predstavlja karakteristično vrijeme koje je potrebno za opisani gubitak korelacije između početnih putanja u faznom prostoru i konačnih mikrostanja. Osim toga, τ također predstavlja vremenski interval tokom kojeg predviđanja, zasnovana na nepotpunoj informaciji o mikroskopskoj dinamici, postaju u najvećoj mogućoj mjeri nesigurna na način koji je kompatibilan s podacima koji su dostupni. Ova nesigurnost je povezana s gubitkom informacije o stanju sustava.

Ta interpretacija se odražava u ulozi Lagrangeovih multiplikatora $\lambda_1((q_0, p_0)_\omega, t_0; t)$ i $\lambda_2((q_0, p_0)_\omega, t_0; t)$. Oni moraju zadovoljavati jednadžbu (4.34), pa se integriranjem te jednadžbe dobije slijedeća relacija,

$$\lambda_2((q_0, p_0)_\omega, t_0; t) = \int_{t_0}^{t} \lambda_1((q_0, p_0)_\omega, t_0; t')dt' + \lambda_2((q_0, p_0)_\omega, t_0; t_0), \tag{4.41}$$

za sve t u intervalu $t_0 \leq t \leq t_a$. Korištenjem relacije (4.41), s jednadžbama (4.36), (4.37) i (4.40), dobije se

$$S_I^f(t_a) = S_I^{DF}(t_a, t_0) = \log W(M) =$$

$$= \int_{t_0}^{t_a} \lambda_1((q_0, p_0)_\omega, t_0; t)dt + \lambda_2((q_0, p_0)_\omega, t_0; t_0). \tag{4.42}$$

Jasno je, iz jednadžbi (4.37), (4.41) i (4.42), da je u trenutku t_a Lagrangeov multiplikator $\lambda_2((q_0, p_0)_\omega, t_0; t_a) = \lambda_2(t_a)$ određen mjerom $W(M)$ invarijantnog skupa M svih mogućih mikrostanja, tj., volumenom dostupnog dijela faznog prostora. Uzastopna primjena MaxEnt algoritma opisanog tipa za zatvoreni sustav s Hamiltonovom dinamikom, bez uvođenja dodatnih ograničenja, rezultira povećanjem $W(M)$. Iz jednadžbi (4.37), (4.41) i (4.42) tada slijedi da je $\lambda_2(t_a) \geq \lambda_2(t_0)$.

Informacija o strukturi mogućih mikrostanja ograničava odgovarajući skup, i postavlja gornju granicu na volumen dostupnog faznog prostora. Vrijednosti $S_I^{DF}(t_a, t_0)$ i $S_I^f(t_a)$ u trenutku t_a, dane jednadžbom (4.42), jednake su maksimalnoj vrijednosti *Boltzmann-Gibbsove entropije*, kompatibilnoj s ovom informacijom. Lagrangeov multiplikator $\lambda_1((q_0, p_0)_\omega, t_0; t)$, integriran u jednadžbi (4.42) po vremenu $t_0 \leq t \leq t_a$, određen je tada brzinom kojom se doseže maksimum Boltzmann-Gibbsove entropije u reproducibilnoj vremenskoj evoluciji. Integral u izrazu (4.42), i veličina $\lambda_1((q_0, p_0)_\omega, t_0; t)$, mogu se identificirati kao promjena entropije, i *brzina promjene entropije* za zatvoren Hamiltonov sustav. Ako se informacija o mikroskopskoj dinamici zatvorenog Hamiltonovog sustava smatra potpunom, ostaje otvoreno pitanje može li i tada produkcija entropije biti definirana bez uvođenja coarse-graining procedura, ili makroskopskih, fenomenoloških pristupa. Općenito, dio informacije odbacuje se u svim takvim modelima, u nekoj fazi, da bi se postiglo slaganje s onim što se opaža u prirodi pri različitim pojavama drugog zakona termodinamike.

Poglavlje 5

GENERALIZACIJA I POVEZANOST S NERAVNOTEŽNOM TEORIJOM

5.1 Prepostavke i uvjeti daljne generalizacije

U pristupu koji je uveden u prethodnom poglavlju posebno se po svojoj važnosti ističu veličine τ i $\lambda_1((q_0, p_0)_\omega, t_0; t)$, koje su interpretacijom logičkih posljedica tog pristupa bile dovedene u neposrednu vezu. Iz tog razloga će uz detaljniju analizu biti uzete kao osnova za daljnu generalizaciju pristupa. Bitan element koji se uzima u obzir ovim pristupom je nepotpunost informacije o mikroskopskoj dinamici, odnosno nepotpunost informacije o svim mikroskopskim interakcijama koje karakteriziraju makroskopski sustav. Ovaj element je uveden maksimizacijom uvjetne informacijske entropije $S_I^{DF}(t, t_0)$, uz ograničenja dana uvjetom normiranosti i Liouvilleovom jednadžbom usrednjenom po dostupnom faznom prostoru. Razlozi za to su bili argumentirani u odjeljku 4.1. Tim korakom u predviđanju vremenske evolucije efektivno se uzima u obzir jedino prethodna informacija o mikroskopskoj dinamici, zajedno sa svim drugim podacima koji su dostupni na makroskopskoj razini. Zbog logički ispravne interpretacije bilo je potrebno uvesti karakteristično vrijeme τ, koje je određeno svim fizikalnim procesima o kojima nemamo informaciju, a koji dovode do gubitka korelacije između putanja u faznom prostoru i mikrostanja. Izvedeni zaključci su iz tog razloga bili relevantni za vremenske intervale koji su dulji od karakterističnog vremena τ.

Posebno se to odražava u činjenici da su uvjetna raspodjela vjerojatnosti koja je dobivena primjenom MaxEnt algoritma, i odgovarajuća raspodjela vjerojatnosti mikrostanja, jedinstveno određene jedino u vremenskim trenucima t_0 i t_a, koji zadovoljavaju uvjet $t_a - t_0 \gg \tau$. Svakoj uvjetnoj raspodjeli vjerojatnosti koja je u početnom trenutku t_0

dobro definirana odgovara jedna klasa rješenja $\{D(q,p,t|(q_0,p_0)_\omega,t_0)\}$, što je bilo detaljno objašnjeno u odjeljku 4.2. U trenutku t_a sva rješenja su jedinstveno određena iz Eulerovog rubnog uvjeta (4.33), koji je nužan uvjet za stacionarnost varijacija. Za kraće vremenske intervale, moguće vremenske evolucije su dane klasom jednako vjerojatnih MaxEnt rješenja $\{D(q,p,t|(q_0,p_0)_\omega,t_0)\}$. Bitno je napomenuti da za bilo koji izbor početnih vrijednosti, u odgovarajuću klasu MaxEnt rješenja ne pripadaju funkcije koje u svim vremenskim trenucima intervala (t_0,t_a) zadovoljavaju Liouvilleovu jednadžbu (4.11). Sve funkcije iz bilo koje klase MaxEnt rješenja $\{D(q,p,t|(q_0,p_0)_\omega,t_0)\}$ ne zadovoljavaju Liouvilleovu jednadžbu (4.11), s obzirom da ona isključuje mogućnost statističke nezavisnosti početnih putanja i konačnih mikrostanja, a time i gubitak korelacije kao logičku posljedicu statističke nezavisnosti.

MaxEnt rješenja $\{D(q,p,t|(q_0,p_0)_\omega,t_0)\}$ su karakterizirana upravo ovakvim statističkim svojstvima: to su funkcije efektivno reducirane informacije o mikroskopskoj dinamici, koja po temeljnom načelu makroskopske reproducibilnosti nije uvijek nužno potrebna u svojoj potpunosti za predikciju reproducibilne makroskopske vremenske evolucije. Dakle, dobivena MaxEnt rješenja $\{D(q,p,t|(q_0,p_0)_\omega,t_0)\}$ nisu funkcije koje egzaktno zadovoljavaju Liouvilleovu jednadžbu (4.11), jer bi takve funkcije implicitno sadržavale i vrlo detaljnu informaciju o mikroskopskoj dinamici. To je najasnije u slučaju funkcije gustoće vjerojatnosti dane jednadžbom (3.18), iz koje se mogu izvesti i sve druge ovdje uvedene raspodjele, a koja u danom obliku (3.18) egzaktno zadovoljava odgovarajuću Liouvilleovu jednadžbu.

MaxEnt algoritam nam omogućuje da problem rješavanja iznimno složene Liouvilleove jednadžbe zamijenimo znatno jednostavnijim varijacijskim problemom nalaženja maksimuma uvjetne informacijske entropije uz reduciran skup makroskopskih ograničenja. Naravno da je na taj način znatno umanjena mogućnost predikcije. U konkretnom slučaju prethodnog pristupa to znači da ne možemo dati pouzdana predviđanja u vremenskim intervalima za koje uvjet $t_a - t_0 \gg \tau$ nije zadovoljen. Daljnjom analizom i generalizacijom pristupa će biti pokazano da je to svojstvo njegova općenita karakteristika. Izvan tog ograničenja, očekujemo da se naše predikcije za makroskopske veličine dobro slažu s egzaktnim rješenjima koja bi dobili rješavanjem Liouvilleove jednadžbe. U prediktivnoj statističkoj mehanici ovaj reducirani pristup temelji se prvenstveno na načelu makroskopske reproducibilnosti, prema kojem je za predviđanje reproducibilnih pojava dovoljno poznavati vrijednosti relevantnih makroskopskih podataka [9, 10]. To je moguće jer je neka makroskopska pojava reproducibilna upravo zato što je ogromna većina mikroskopskih realizacija te pojave karakterizirana istim makroskopskim ponašanjem. Navedena svojstva i očekivanja potvrđuju i zaključci do kojih je u okviru MaxEnt formalizma došao Grandy [16, 17, 18]. Uz poopćenje Liouville–von Neumannove jednadžbe za matricu gustoće, Grandy je primjenom MaxEnt modela vremenski ovisnih vjerojatnosti dao opis tipičnih procesa u neravnotežnoj termodinamici i hidrodinamici koji je konzistentan sa standardnom neravnotežnom teorijom.

Makroskopski sustav karakterizira uz vrijeme τ i integral po vremenu Lagrangeovog multiplikatora $\lambda_1((q_0,p_0)_\omega,t_0;t)$, koji je pomoću jednadžbe (4.42) identificiran kao promjena entropije. Lagrangeov multiplikator $\lambda_1((q_0,p_0)_\omega,t_0;t)$ interpretiran je kao brzina promjene entropije, odnosno kao produkcija entropije zatvorenog Hamiltonovog sustava. Ova povezanost je jasnija ako se usporede rezultati dobiveni maksimizacijom uvjetne informacijske entropije $S_I^{DF}(t,t_0)$, uz dva različita tipa ograničenja koja su bila uvedena u prethodnom poglavlju. U prvom slučaju, gdje je Liouvilleova jednadžba bila uvedena kao strogo mikroskopsko ograničenje, dobiven je rezultat dan jednadžbom (4.29), koja precizno određuje da je tada Lagrangeov multiplikator $\lambda_1((q_0,p_0)_\omega,t_0;t)$ identično jednak nuli. U vremenskoj evoluciji koja je za Hamiltonove sustave egzaktno opisana Liouvilleovom jednadžbom, Gibbs-Shannonov i Boltzmannov izraz za entropiju ostaju konstantni u vremenu. U tom smislu interpretacija Lagrangeovog multiplikatora $\lambda_1((q_0,p_0)_\omega,t_0;t)$ koja je proizišla u drugom pristupu na osnovi razmatranja porasta entropije, postaje u granici koja se razmatra prvim pristupom konzistentna s Liovilleovom vremenskom evolucijom u kojoj nema promjene entropije. U toj granici nije moguća statistička nezavisnost početnih putanja i konačnih mikrostanja i nema s tim povezanog gubitka korelacije, pa se odgovarajuće karakteristično vrijeme τ ne može definirati.

Očito je da iz rezultata prethodno opisanog pristupa nije bilo moguće odrediti precizno Lagrangeov multiplikator $\lambda_1((q_0,p_0)_\omega,t_0;t)$ koji u jednadžbi (4.42) određuje promjenu entropije. Također, kao što je prethodno detaljno objašnjeno, nije bilo moguće jedinstveno odrediti relevantne statističke raspodjele za vremenske intervale koji su kraći od karakterističnog vremena τ. Sa stajališta prediktivne statističke mehanike to je jednostavno objasniti činjenicom da početni osnovni model nije sadržavao dovoljno podataka koji su nužni za precizna predviđanja. Generalizacija modela zato nužno mora uključivati one makroskopske veličine koje su relevantne za predikciju vrijednosti drugih veličina koje karakteriziraju makroskopsku vremensku evoluciju na određenoj vremenskoj skali. Iz razloga što želimo pokazati da je u tom slučaju prethodni pristup konzistentan s poznatim rezultatima neravnotežne statističke mehanike i termodinamike ireverzibilnih procesa, potrebno je najprije razmotriti neke od uvjeta koje treba ispuniti da bi se takva generalizacija i veza mogle uspostaviti. To se u prvom redu odnosi na određivanje relevantnih makroskopskih veličina.

5.2 Reducirani opis neravnotežnih sustava

U ovom odjeljku uvode se i dijelom slijede osnove reduciranog opisa neravnotežnih sustava iz reference [24]. Prema tom pristupu [24, str. 89], ako nas zanima ponašanje sustava za vremenske intervale koji nisu pretkratki u specificiranom smislu, detalji početnog stanja postaju nebitni i broj parametara koji su nužni za opis stanja sustava je reduciran. U navedenom pristupu [24, str. 90] se korištenjem načela najveće informacijske entropije konstruiraju generalizirani Gibbsovi ansambli koji su blisko povezani s termodinamičkim opisom neravnotežnih sustava kad opservabilne makroskopske veličine ovise o vremenu.

Razlika u odnosu na MaxEnt pristup je u stajalištu [24, str. 90] da tako dobivene relevantne statističke raspodjele još uvijek nisu tražene statističke raspodjele jer, općenito, ne zadovoljavaju Liouvilleovu jednadžbu, nego služe kao pomoćne raspodjele za odabir posebnih rješenja Liouvilleove jednadžbe koji opisuju ireverzibilne makroskopske procese. Neravnotežni ansambli [24, str. 136] se konstruiraju na osnovi retardiranih rješenja Liouvilleove jednadžbe metodom neravnotežnog statističkog operatora. Metoda se zasniva na principu jednakih vjerojatnosti početnih stanja opisanih relevantnim raspodjelama [24, str. 119]: neravnotežna raspodjela je u bilo kojem trenutku rezultat vremenskog usrednjavanja Liouville evoluiranih početnih relevantnih raspodjela po svim početnim trenucima vremenskog intervala dovoljno dugog za iščezavanje mikroskopskih "nefizikalnih stanja" i pojavu nužnih korelacija u sustavu. Prema referenci [24, str. 119], fizikalno se ovo usrednjavanje rješenja Liouvilleove jednadžbe po početnim vremenskim trenucima može smatrati generalizacijom usrednjavanja po vremenu opažanja.

Na ovaj način se u neravnotežni ansambl uvodi se dobro poznato svojstvo [24, str. 118] da makroskopski sustav "zaboravlja" irelevantne detalje svog početnog stanja nakon nekog mikroskopskog vremena τ koje je karakteristično za dani sustav. Ovisno o odabiru skupa relevantnih varijabli metoda neravnotežnog statističkog operatora [24, str. 136] dopušta izvod kinetičkih, hidrodinamičkih i relaksacijskih jednadžbi, koje opisuju makroskopsku evoluciju sustava na različitim vremenskim skalama. Uz ovu metodu u literaturi su poznate [34, 35, 36, 37, 38] i druge metode koje koriste reducirane opise neravnotežnih stanja zajedno s kvantnim ili klasičnim Liouvilleovim jednadžbama.

Sa stajališta prediktivne statističke mehanike, osim kvantnomehaničke vjerojatnosti, nemamo razloga smatrati bilo koju raspodjelu vjerojatnosti jedino ispravnom raspodjelom. Takvo stajalište je u izrazitoj suprotnosti s interpretacijom koja vjerojatnost definira isključivo u frekvencijskom smislu kao objektivno svojstvo promatranog sustava. U frekvencijskoj interpretaciji su vjerojatnosti empirijski provjerljive, i sukladno tome [5] temeljni problem statističke mehanike bi bio da ih izvede odnosno opravda u frekvencijskom smislu. Jaynes je zastupao suprotno stajalište [6], da ako opisujemo stupanj našeg znanja o pojedinačnom sustavu onda ne može biti ništa fizikalno stvarno u frekvencijama ansambla velikog broja sustava, niti ima smisla pitati se koji je ansambl sustava jedino ispravan. Ono što nazivamo različitim ansamblima u stvari korespondira različitim stupnjevima poznavanja pojedinačnog sustava odnosno neke fizikalne situacije. Jaynes se u argumentaciji tog stajališta poziva na Gibbsovu [3] tvrdnju prema kojoj su ansambli odabrani da ilustriraju vjerojatnosti događaja u stvarnom svijetu.

Interpretacija Gibbsovog formalizma slijedi iz ranije spomenute činjenice da se maksimizacijom informacijske entropije uz dana ograničenja predviđa upravo makroskopsko ponašanje koje se može dogoditi na najveći mogući broj načina kompatibilno s ograničenjima. Ne ulazeći dublje u problem interpretacije vjerojatnosti, koji je još izraženiji u slučaju neravnotežnih stanja, važno je spomenuti da raspodjele koje slijede iz primjene načela najveće informacijske entropije ovise jedino o dostupnoj informaciji. S obzirom da ne ovise o bilo kojoj drugoj informaciji koja nije dostupna, nije potpuno jasno u kojem smislu bi te raspodjele vjerojatnosti mogle predstavljati objektivno svojstvo danog sustava. Ako se pri tome misli na predikcije, s istog stajališta se o objektivnosti može govoriti u mjeri u kojoj je uzeta u obzir nepotpunost informacije. Dosljedno tom načinu razmišljanja, pristup iz reference [24] ćemo primjenjivati u dijelu u kojem se primjenom načela najveće informacijske entropije dolazi do relevantnih statističkih raspodjela.

U pristupu danom u referenci [24] važnu ulogu u opisu ireverzibilnosti ima uvođenje karakterističnog mikroskopskog vremena nakon kojeg makroskopski sustav "zaboravlja" irelevantne detalje svog početnog stanja. U našem pristupu toj veličini bi po analogiji odgovaralo vrijeme τ, koje je karakteristično za gubitak korelacije između putanja u faznom prostoru i mikrostanja. Iz analize našeg početnog modela bilo je jasno da za vremenske intervale kraće od τ, MaxEnt rješenja nisu bila jedinstveno određena, nego su dana klasama jednako vjerojatnih rješenja $\{D(q,p,t|(q_0,p_0)_\omega,t_0)\}$.

Općenito, prvi korak u izvodu jednadžbi koje opisuju ireverzibilnu vremensku evoluciju iz reverzibilne Liouvilleove jednadžbe, sastoji se od odabira skupa dinamičkih varijabli i pridruženih opservabilnih veličina koje su relevantne za opis na odgovarajućoj vremenskoj skali. Dakle, reducirani opis neravnotežnog sustava ovisi o odgovarajućoj vremenskoj skali. Zato će u generalizaciji našeg početnog modela izbor relevantnih veličina, odnosno uvođenje odgovarajućih dodatnih ograničenja pri maksimizaciji uvjetne informacijske entropije $S_I^{DF}(t,t_0)$, omogućiti u fizikalnom smislu precizniju definiciju vremena τ.

Zbog orijentacije ovdje navodimo primjer preuzet iz reference [24], gdje se za razrijeđen klasični plin N identičnih čestica zatvorenih u konstantan volumen V, uvodi hijerarhija vremenskih skala koja u ovom slučaju razlikuje tri karakteristična vremenska stadija (skale):

- *dinamički stadij* (skala) za vremenske intervale Δt kraće od vremena sudara $\Delta t \ll \tau_0$. Za opis vremenske evolucije na tako kratkim intervalima potrebna je potpuna

N-čestična raspodjela vjerojatnosti mikrostanja $f(x, p)$, koja daje najdetaljniji statistički opis sustava. Na ovoj skali se reducirani opis ne može koristiti.

- *kinetički stadij* (skala) za vremenske intervale koji zadovoljavaju uvjet $\tau_0 \ll \Delta t \ll \tau_r$, gdje je τ_r vrijeme potrebno za uspostavljanje lokalne ravnoteže u makroskopski malom volumenu koji sadrži velik broj čestica. Prepostavka je da se nakon vremena koje je veliko u usporedbi s trajanjem sudara, sustav može prikladno opisati jednočestičnom raspodjelom vjerojatnosti koordinata i impulsa $f_1(\mathbf{r}, \mathbf{p}, t)$. Evolucija plina u kinetičkom stadiju opisana je kinetičkom jednadžbom za jednočestičnu raspodjelu.

- *hidrodinamički stadij* (skala) za vremenske intervale koji zadovoljavaju uvjet $\tau_r \ll \Delta t \ll \tau_{eq}$, gdje je τ_{eq} vrijeme potrebno za relaksaciju sustava ka globalnoj ravnoteži. Na ovoj skali sustav je došao u stanje lokalne ravnoteže. Lokalne makroskopske veličine kao što su lokalna gustoća broja čestica $\langle n(\mathbf{r}) \rangle_t$, lokalna gustoća impulsa $\langle \mathbf{P}(\mathbf{r}) \rangle_t$ i lokalna gustoća energije $\langle h(\mathbf{r}) \rangle_t$ dovoljne su za opis sustava.

Uz ova karakteristična vremena važno je i vrijeme između dva uzastopna sudara τ_f. Na osnovi elementarnih kinetičkih razmatranja može se pokazati da je za razrijeđene plinove zadovoljen uvjet $\tau_0 \ll \tau_f$. Do lokalne ravnoteže sustav dolazi kroz puno sudara pa vrijedi $\tau_f \ll \tau_r$, dok je uvjet $\tau_r \ll \tau_{eq}$ očito zadovoljen. Iz ovih procjena [24, str. 91] slijedi da za razrijeđene plinove postoji hijerarhija osnovnih relaksacijskih vremena $\tau_0 \ll \tau_f \ll \tau_r \ll \tau_{eq}$, koja nam dopušta da kažemo da se približavanje razrijeđenog plina ravnoteži odvija u tri stadija.

Treba naglasiti da ove vremenske skale nisu uvijek definirane. Neki od primjera su navedeni [24, str. 93]:

- iznimno razrijeđeni plinovi, za koje vrijeme τ_f između sudara može postati reda veličine τ_{eq}, pa hidrodinamički stadij gubi svoje značenje.

- kinetička teorija kvantnih plinova, u kojoj jednočestična matrica gustoće ima ulogu jednočestične raspodjele. Mnogi kvantni sustavi se mogu opisati kao slabo međudjelujući plinovi kvazičestica.

- tekućine, gdje je vrijeme sudara τ_0 istog reda veličine kao τ_f. Kinetički stadij se ne može razlučiti u ovom slučaju, dok je hidrodinamički opis prikladan do skale usporedive s vremenom molekularnih interakcija.

Kao bitan element našeg pristupa, maksimizacijom uvjetne informacijske entropije $S_I^{DF}(t, t_0)$ uz ograničenja dana uvjetom normiranosti i Liouvilleovom jednadžbom usrednjenom po dostupnom faznom prostoru, uvedena je nepotpunost informacije o mikroskopskoj dinamici sustava, odnosno o interakcijama koje je karakteriziraju. Zato je prirodno da za početni korak u generalizaciji pristupa odaberemo hidrodinamički stadij, vremensku skalu na kojoj je za reducirani opis neravnotežnog makroskopskog sustava potrebna manje detaljna informacija o mikroskopskoj dinamici nego što je to slučaj s ostale dvije skale.

5.3 Hidrodinamičke jednadžbe kontinuiteta

Kao što je spomenuto u prethodnom odjeljku, za vremenske intervale dulje od vremena τ_r potrebnog za uspostavljanje lokalne ravnoteže, lokalne makroskopske veličine, kao što su *lokalna gustoća broja čestica* $\langle n(\mathbf{r})\rangle_t$, *lokalna gustoća impulsa* $\langle \mathbf{P}(\mathbf{r})\rangle_t$ i *lokalna gustoća energije* $\langle h(\mathbf{r})\rangle_t$, dovoljne su za opis neravnotežnog sustava. Vrijednosti ovih veličina dobiju se usrednjavanjem odgovarajućih dinamičkih varijabli po raspodjeli vjerojatnosti mikrostanja $f(x, p, t)$ u trenutku t. Za klasični fluid N identičnih čestica, koji se uzima ovdje kao osnova za analizu, dinamičke varijable koje odgovaraju tim veličinama su gustoća broja čestica

$$n(\mathbf{r}) \equiv n(\mathbf{r}; \mathbf{r}_1, \ldots, \mathbf{r}_N) = \sum_{i=1}^{N} \delta(\mathbf{r} - \mathbf{r}_i) \ , \tag{5.1}$$

gustoća impulsa

$$\mathbf{P}(\mathbf{r}) \equiv \mathbf{P}(\mathbf{r}; \mathbf{r}_1, \mathbf{p}_1, \ldots, \mathbf{r}_N, \mathbf{p}_N) = \sum_{i=1}^{N} \mathbf{p}_i \delta(\mathbf{r} - \mathbf{r}_i) \ , \tag{5.2}$$

i gustoća energije

$$\begin{aligned} h(\mathbf{r}) &\equiv h(\mathbf{r}; \mathbf{r}_1, \mathbf{p}_1, \ldots, \mathbf{r}_N, \mathbf{p}_N) \\ &= \sum_{i=1}^{N} \left[\frac{\mathbf{p}_i^2}{2m} + \frac{1}{2} \sum_{j=1,\ j\neq i}^{N} \Phi(|\mathbf{r}_i - \mathbf{r}_j|) \right] \delta(\mathbf{r} - \mathbf{r}_i) \ . \end{aligned} \tag{5.3}$$

Klasični fluid N indentičnih čestica opisan je translacijski i rotacijski invarijantnom Hamiltonovom funkcijom:

$$H(x, p) = \sum_{i=1}^{N} \left[\frac{\mathbf{p}_i^2}{2m} + \frac{1}{2} \sum_{j=1,\ j\neq i}^{N} \Phi(|\mathbf{r}_i - \mathbf{r}_j|) \right] \ , \tag{5.4}$$

gdje je $\Phi(|\mathbf{r}_i - \mathbf{r}_j|)$ potencijalna energija interakcije para čestica s indeksima i, j. Notacija (x, p) u jednadžbi (5.4) označava skup $6N$ dinamičkih varijabli danih koordinatama $(x_1, \ldots, x_{3N}) = (x_1, y_1, z_1, \ldots, x_N, y_N, z_N)$ i konjugiranim impulsima $(p_1, \ldots, p_{3N}) =$

$(p_{1x}, p_{1y}, p_{1z}, \ldots, p_{Nx}, p_{Ny}, p_{Nz})$. Skup varijabli (x, p) čine Kartezijeve komponente N vektora položaja čestica $(\mathbf{r}_1, \ldots, \mathbf{r}_N)$ i Kartezijeve komponente N vektora impulsa čestica $(\mathbf{p}_1, \ldots, \mathbf{p}_N)$. Vremenska ovisnost varijabli (q, p) je određena Hamiltonovim jednadžbama

$$\dot{x}_i = \frac{\partial H}{\partial p_i}, \qquad \dot{p}_i = -\frac{\partial H}{\partial x_i}, \qquad 1 \leq i \leq 3N \ , \tag{5.5}$$

što se može ekvivalentno zapisati i u obliku

$$\dot{\mathbf{r}}_k = \frac{\partial H}{\partial \mathbf{p}_k}, \qquad \dot{\mathbf{p}}_k = -\frac{\partial H}{\partial \mathbf{r}_k}, \qquad 1 \leq k \leq N \ . \tag{5.6}$$

Ako proizvoljno dodijelimo polovinu potencijalne energije interakcije parova svakoj od pripadnih čestica tih parova, u izrazima (5.3) i (5.4) možemo prepoznati funkcije koje se tada mogu interpretirati kao energije pojedinih čestica označenih s indeksom i,

$$H_i(\mathbf{p}_i, \mathbf{r}_1, \ldots, \mathbf{r}_N) = \frac{\mathbf{p}_i^2}{2m} + \frac{1}{2} \sum_{j=1,\ j \neq i}^{N} \Phi(|\mathbf{r}_i - \mathbf{r}_j|) \ . \tag{5.7}$$

Korištenjem izraza (5.7), gustoću energije (5.3) i Hamiltonovu funkciju (5.4) možemo zapisati i u drukčijim oblicima koji su pogodni u računima:

$$h(\mathbf{r}; \mathbf{r}_1, \mathbf{p}_1, \ldots, \mathbf{r}_N, \mathbf{p}_N) = \sum_{i=1}^{N} H_i(\mathbf{p}_i, \mathbf{r}_1, \ldots, \mathbf{r}_N) \delta(\mathbf{r} - \mathbf{r}_i) \ , \tag{5.8}$$

i

$$H(x, p) = \sum_{i=1}^{N} H_i(\mathbf{p}_i, \mathbf{r}_1, \ldots, \mathbf{r}_N) \ . \tag{5.9}$$

Hamiltonova funkcija (5.4) je jednaka integralu gustoće energije (5.3) po cijelom volumenu sustava,

$$H(x, p) = \int h(\mathbf{r}; \mathbf{r}_1, \mathbf{p}_1, \ldots, \mathbf{r}_N, \mathbf{p}_N) \, d^3\mathbf{r} \ . \tag{5.10}$$

Također, lako je pokazati da integrali dinamičkih varijabli (5.1) i (5.2) po cijelom volumenu sustava daju ukupan broj čestica,

$$N = \int n(\mathbf{r}; \mathbf{r}_1, \ldots, \mathbf{r}_N) \, d^3\mathbf{r} \ , \tag{5.11}$$

i ukupan impuls sustava čestica

$$\mathbf{P}_{ukup} = \sum_{i=1}^{N} \mathbf{p}_i = \int \mathbf{P}(\mathbf{r}; \mathbf{r}_1, \mathbf{p}_1, \ldots, \mathbf{r}_N, \mathbf{p}_N) \, d^3\mathbf{r} \ . \tag{5.12}$$

Lokalne vrijednosti makroskopskih veličina koje opisuju klasični fluid identičnih čestica, dobiju se usrednjavanem dinamičkih varijabli $n(\mathbf{r})$, $\mathbf{P}(\mathbf{r})$ i $h(\mathbf{r})$ po raspodjeli vjerojatnosti mikrostanja $f(x,p,t)$ u trenutku t:

$$\begin{aligned}
\langle n(\mathbf{r})\rangle_t &= \int_M f(x,p,t)n(\mathbf{r};\mathbf{r}_1,\ldots,\mathbf{r}_N)\,d\Gamma \;, &(5.13)\\
\langle \mathbf{P}(\mathbf{r})\rangle_t &= \int_M f(x,p,t)\,\mathbf{P}(\mathbf{r};\mathbf{r}_1,\mathbf{p}_1,\ldots,\mathbf{r}_N,\mathbf{p}_N)\,d\Gamma \;, &(5.14)\\
\langle h(\mathbf{r})\rangle_t &= \int_M f(x,p,t)\,h(\mathbf{r};\mathbf{r}_1,\mathbf{p}_1,\ldots,\mathbf{r}_N,\mathbf{p}_N)\,d\Gamma \;, &(5.15)
\end{aligned}$$

gdje je $d\Gamma = dx_1\ldots dx_{3N}dp_1\ldots dp_{3N}$ element volumena $6N$-dimenzionalnog faznog prostora Γ. Usrednjavanja po $f(x,p,t)$ dana su integralima po skupu $M\subset\Gamma$ koji odgovara svim mogućim mikrostanjima, a za kojeg se ovdje po definiciji uzima da je invarijantan na Hamiltonovo gibanje. Raspodjela vjerojatnosti mikrostanja $f(x,p,t)$ sustava N identičnih čestica normalizirana je u skladu s definicijom mikrostanja u faznom prostoru koja slijedi u klasičnoj granici kvantne statističke mehanike [33].

Integracijom lokalne gustoće broja čestica $\langle n(\mathbf{r})\rangle_t$, dane jednadžbom (5.13), po cijelom volumenu sustava, uz korištenje jednadžbe (5.1) i uvjeta normiranosti $f(x,p,t)$, dobije se

$$N = \int \langle n(\mathbf{r})\rangle_t\, d^3\mathbf{r} \;. \tag{5.16}$$

Vrijednosti dobivene iz izraza (5.11) i (5.16) su jednake; ukupan broj čestica sustava je fiksan i jednak N. Hamiltonova funkcija $H(x,p)$ konstantna je u vremenu; u slučaju translacijski invarijantne Hamiltonove funkcije (5.4) isto vrijedi i za ukupni impuls $\mathbf{P}_{ukup}$, koji je dan jednadžbom (5.12). To se može zapisati pomoću Poissonovih zagrada:

$$\begin{aligned}
\frac{d\mathbf{P}_{ukup}}{dt} &= \{\mathbf{P}_{ukup},H\} = 0 \;, &(5.17)\\
\frac{dH}{dt} &= \{H,H\} = 0 \;. &(5.18)
\end{aligned}$$

Za bilo koje dvije funkcije $\varphi_1(x,p)$ i $\varphi_2(x,p)$, Poissonova zagrada definirana je slijedećim izrazom

$$\begin{aligned}
\{\varphi_1,\varphi_2\} &= \sum_{i=1}^{3N}\left(\frac{\partial\varphi_1}{\partial x_i}\frac{\partial\varphi_2}{\partial p_i} - \frac{\partial\varphi_1}{\partial p_i}\frac{\partial\varphi_2}{\partial x_i}\right)\\
&= \sum_{k=1}^{N}\left(\frac{\partial\varphi_1}{\partial\mathbf{r}_k}\cdot\frac{\partial\varphi_2}{\partial\mathbf{p}_k} - \frac{\partial\varphi_1}{\partial\mathbf{p}_k}\cdot\frac{\partial\varphi_2}{\partial\mathbf{r}_k}\right) \;. &(5.19)
\end{aligned}$$

U posljednoj liniji jednadžbe (5.19) Poissonova zagrada je napisana u pogodnom alternativnom obliku, korištenjem skalarnog produkta gradijenata, čije komponente su parcijalne derivacije funkcija $\varphi_1(x,p)$ i $\varphi_2(x,p)$ po Kartezijevim komponentama vektora položaja i impulsa pojedinih čestica. Iz jednadžbi (5.17) i (5.18) je očito da su ukupni impuls i energija konstante gibanja. Hamiltonovom funkcijom (5.4) volumen sustava N identičnih čestica nije preciziran, što će ponovo biti analizirano u diskusiji rubnih uvjeta koji vrijede na granici sustava.

Kao što je prethodno bilo pokazano, lokalne dinamičke varijable $n(\mathbf{r})$, $\mathbf{P}(\mathbf{r})$ i $h(\mathbf{r})$ gustoće su odgovarajućih sačuvanih veličina N, $\mathbf{P}_{ukup}$ i $H(x,p)$. Jednadžbe gibanja tih dinamičkih varijabli mogu se zato zapisati u obliku lokalnih mikroskopskih zakona sačuvanja [24, str. 93],

$$\begin{aligned}
\frac{dn(\mathbf{r})}{dt} &= \{n(\mathbf{r}),H\} = -\nabla\cdot\mathbf{J}(\mathbf{r}) \ , \\
\frac{dP_\alpha(\mathbf{r})}{dt} &= \{P_\alpha(\mathbf{r}),H\} = -\nabla\cdot\mathbf{J}_{P_\alpha}(\mathbf{r}) \ , \\
\frac{dh(\mathbf{r})}{dt} &= \{h(\mathbf{r}),H\} = -\nabla\cdot\mathbf{J}_h(\mathbf{r}) \ . \qquad (5.20)
\end{aligned}$$

Kartezijeve komponente vektora gustoće impulsa $\mathbf{P}(\mathbf{r})$ označene su s $P_\alpha(\mathbf{r})$, $\alpha = 1,2,3$. Dinamičke varijable $\mathbf{J}(\mathbf{r})$, $\mathbf{J}_{P_\alpha}(\mathbf{r})$ i $\mathbf{J}_h(\mathbf{r})$ predstavljaju gustoće struja odgovarajućih sačuvanih veličina čije su gustoće $n(\mathbf{r})$, $P_\alpha(\mathbf{r})$ i $h(\mathbf{r})$. Divergencije ovih gustoća struja uzete su s obzirom na vektor $\mathbf{r}$ u fiksnoj točki prostora. Jednadžbe (5.20) su standardni izrazi; izvodi eksplicitnih izraza za gustoće struja mogu se pronaći u literaturi [13], [24], [28], [39].

Jednostavno je pokazati da je srednja vrijednost vremenske derivacije bilo koje dinamičke varijable A jednaka vremenskoj derivaciji srednje vrijednosti $\langle A\rangle_t = \int_M Af\,d\Gamma$ iste varijable:

$$\begin{aligned}
\left\langle\frac{dA}{dt}\right\rangle_t &= \int_M\left(\frac{\partial A}{\partial t}f + \{A,H\}f\right)d\Gamma \\
&= \int_M\left(\frac{\partial A}{\partial t}f - A\{f,H\}\right)d\Gamma \\
&= \int_M\left(\frac{\partial A}{\partial t}f + A\frac{\partial f}{\partial t}\right)d\Gamma \\
&= \frac{d\langle A\rangle_t}{dt} \ . \qquad (5.21)
\end{aligned}$$

U izvodu jednakosti (5.21) korištene su: jednadžba gibanja za dinamičku varijablu A (varijabla A može i eksplicitno ovisiti o vremenu),

$$\frac{dA}{dt} = \frac{\partial A}{\partial t} + \{A,H\} \ , \qquad (5.22)$$

Liouvilleova jednadžba za raspodjelu vjerojatnosti mikrostanja $f(x,p,t)$,

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \{f,H\} = 0 \ , \qquad (5.23)$$

te n-dimenzionalna generalizacija teorema o divergenciji (formula Ostrogradskog) [40] čijom primjenom slijedi drugi red jednadžbe (5.21) uz iščezavanje doprinosa ruba invarijantnog skupa M (zbog jednakosti $\mathbf{v}\cdot\mathbf{n} = 0$; objašnjenje je dato uz jednadžbu (4.25)).

Usrednjavanjem jednadžbi (5.20) po raspodjeli vjerojatnosti mikrostanja $f(x, p, t)$, zbog jednakosti (5.21) dobiju se slijedeći izrazi:

$$\begin{aligned} \frac{\partial \langle n(\mathbf{r}) \rangle_t}{\partial t} &= \langle \{n(\mathbf{r}), H\} \rangle_t = -\nabla \cdot \langle \mathbf{J}(\mathbf{r}) \rangle_t \ , \\ \frac{\partial \langle P_\alpha(\mathbf{r}) \rangle_t}{\partial t} &= \langle \{P_\alpha(\mathbf{r}), H\} \rangle_t = -\nabla \cdot \langle \mathbf{J}_{P_\alpha}(\mathbf{r}) \rangle_t \ , \\ \frac{\partial \langle h(\mathbf{r}) \rangle_t}{\partial t} &= \langle \{h(\mathbf{r}), H\} \rangle_t = -\nabla \cdot \langle \mathbf{J}_h(\mathbf{r}) \rangle_t \ . \end{aligned} \tag{5.24}$$

Derivacija po vremenu u jednadžbama (5.24) je označena kao parcijalna derivacija iz razloga što gustoće i njihove srednje vrijednosti ovise i o prostornoj koordinati $\mathbf{r}$. Ove jednadžbe predstavljaju *lokalne makroskopske zakone sačuvanja*, koji služe kao osnova za izvod jednadžbi hidrodinamike [13], [24], [28], [39].

Važan korak u izvodu jednakosti (5.21) za proizvoljnu varijablu A, a onda i u izvodu jednadžbi (5.24) bilo je korištenje Liouvilleove jednadžbe (5.23). Za sustav za koji vrijede lokalni mikroskopski zakoni sačuvanja u obliku (5.20), uvjet da raspodjela vjerojatnosti mikrostanja $f(x, p, t)$ zadovoljava Liouvilleovu jednadžbu (5.23) dovoljan je ali *nije nužan* za jednakosti (5.24).

To se može pokazati na slijedeći način. Ako proizvoljnu dinamičku varijablu A u jednakosti (5.21) zamijenimo gustoćama $n(\mathbf{r})$, $\mathbf{P}(\mathbf{r})$ i $h(\mathbf{r})$, pripadajuće jednakosti (5.21) će i dalje vrijediti ako je ispunjeno

$$\begin{aligned} \int_M n(\mathbf{r}) \left(\frac{\partial f}{\partial t} + \{f, H\} \right) d\Gamma &= 0 \ , \\ \int_M \mathbf{P}(\mathbf{r}) \left(\frac{\partial f}{\partial t} + \{f, H\} \right) d\Gamma &= 0 \ , \\ \int_M h(\mathbf{r}) \left(\frac{\partial f}{\partial t} + \{f, H\} \right) d\Gamma &= 0 \ . \end{aligned} \tag{5.25}$$

Jednadžbe (5.25) ne predstavljaju tako strog uvjet na raspodjelu vjerojatnosti mikrostanja $f(x, p, t)$ kao Liouvilleova jednadžba (5.23).

Osim toga, izravno se može pokazati da su jednadžbe (5.25) *ekvivalentne* lokalnim makroskopskim zakonim sačuvanja (5.24). Korištenjem teorema o divergenciji u izrazima (5.25) uz iščezavanje doprinosa ruba skupa M, na način opisan u izvodu (5.21), te zatim uz korištenje desne strane jednakosti (5.20), dobiju se ekvivalentni izrazi

$$\begin{aligned} &\int_M \left(\frac{\partial f}{\partial t} n(\mathbf{r}) + f \nabla \cdot \mathbf{J}(\mathbf{r}) \right) d\Gamma = 0 \ , \\ &\int_M \left(\frac{\partial f}{\partial t} \mathbf{P}(\mathbf{r}) + f \nabla \cdot \mathbf{J}_{P_\alpha}(\mathbf{r}) \right) d\Gamma = 0 \ , \\ &\int_M \left(\frac{\partial f}{\partial t} h(\mathbf{r}) + f \nabla \cdot \mathbf{J}_h(\mathbf{r}) \right) d\Gamma = 0 \ . \end{aligned} \tag{5.26}$$

Ovi izrazi su upravo lokalni makroskopski zakoni sačuvanja (5.24).

Dakle, može se rezimirati slijedeći zaključak: za sustav za koji vrijede lokalni mikroskopski zakoni sačuvanja u obliku (5.20), jednadžbe (5.25) su ekvivalentne lokalnim makroskopskim zakonima sačuvanja (5.24). Naravno, ovaj zaključak uključuje implicitno mogućnost da raspodjela vjerojatnosti mikrostanja $f(x, p, t)$ zadovoljava općenitiji, odnosno manje strog "zakon gibanja" nego su to Liouvilleova jednadžba (5.23) i zakoni gibanja za dinamičke varijable. Pretpostavku sličnu toj je u kvantnomehaničkom pristupu neravnotežnim sustavima i ireverzibilnim procesima uveo Grandy [16, 17, 18] kroz poopćenje Liouville–von Neumannove jednadžbe gibanja za matricu gustoće. U prisustvu vanjskih utjecaja na sustav, takvim se poopćenjem jednadžbe gibanja za matricu gustoće, uz unitarnu kvantnomehaničku evoluciju, može opisati i promjena klasičnih vjerojatnosti u matrici gustoće do koje dolazi zbog promjene makroskopskih ograničenja. Međutim, iako su ideje u osnovi vrlo slične, ističe se jedna temeljna razlika. Ovdje su osnovne jednadžbe gibanja sustava Hamiltonove jednadžbe klasične mehanike, i te jednadžbe određuju vremensku evoluciju klasičnih raspodjela vjerojatnosti. Poopćenje Liouvilleove jednadžbe za klasične raspodjele vjerojatnosti zato bi imalo smisla jedino na način koji u obzir uzima nepotpunost informacije o mikroskopskoj dinamici. Tu se prirodno postavljaju pitanja da li je poopćenje Liouvilleove jednadžbe za klasične sustave doista potrebno, i je li ostvarivo bez uvođenja dodatnih prepostavki u koje ne možemo biti sigurni.

Ono u što možemo biti sigurni je slijedeće. Pretpostavimo da, uz Hamiltonovu funkciju $H(x, p)$ koja je dana jednadžbom (5.4), potpuna Hamiltonova funkcija $H_{tot}(x, p, t)$ uključuje i dodatni član $H_{ni}(x, p, t)$, o kojem nemamo prethodnu informaciju,

$$H_{tot}(x, p, t) = H(x, p) + H_{ni}(x, p, t) \ . \tag{5.27}$$

Pretpostavimo sada da neka raspodjela vjerojatnosti mikrostanja $\tilde{f}(x, p, t)$ doista zadovoljava "potpunu" Liouvilleovu jednadžbu

$$\frac{\partial \tilde{f}}{\partial t} + \{\tilde{f}, H_{tot}\} = \frac{\partial \tilde{f}}{\partial t} + \{\tilde{f}, H\} + \{\tilde{f}, H_{ni}\} = 0 \ . \tag{5.28}$$

Kao dodatak ovome, pretpostavimo da je invarijantan skup M svih mogućih mikrostanja u faznom prostoru, invarijantan također i na vremensku evoluciju opisanu potpunom Hamiltonovom funkcijom $H_{tot}(x, p, t)$. Takvu je situaciju moguće zamisliti u slučaju da postoji skup dinamičkih varijabli koje su konstante gibanja za obje Hamiltonove funkcije $H(x, p)$ i $H_{tot}(x, p, t)$. Ako se takva pretpostavka čini nerealističnom, možemo umjesto toga uzeti znatno jednostavniju mogućnost da je invarijantan skup mikrostanja cijeli fazni prostor $M = \Gamma$.

Ako su zadovoljene jednadžbe

$$\begin{aligned}
\int_M \tilde{f}\{n(\mathbf{r}), H_{ni}\}\, d\Gamma &= 0 \;, \\
\int_M \tilde{f}\{P_\alpha(\mathbf{r}), H_{ni}\}\, d\Gamma &= 0 \;, \\
\int_M \tilde{f}\{h(\mathbf{r}), H_{ni}\}\, d\Gamma &= 0 \;,
\end{aligned} \tag{5.29}$$

tada lokalni makroskopski zakoni sačuvanja vrijede u obliku koji je identičan jednadžbama (5.24).

Ako su zadovoljene jednadžbe

$$\begin{aligned}
\{n(\mathbf{r}), H_{ni}\} &= 0 \;, \\
\{P_\alpha(\mathbf{r}), H_{ni}\} &= 0 \;, \\
\{h(\mathbf{r}), H_{ni}\} &= 0 \;,
\end{aligned} \tag{5.30}$$

tada lokalni mikroskopski zakoni sačuvanja vrijede u obliku koji je identičan jednadžbama (5.20).

Uvjet (5.30) je restriktivniji za $H_{ni}(x,p,t)$ nego što je to uvjet (5.29); ako je ispunjen uvjet (5.30) tada je također zadovoljen i uvjet (5.29). Prethodne tvrdnje možemo bitno sažeti na slijedeći način. Ako vrijedi potpuna Liouvilleova jednadžba (5.28), i ako su zadovoljene jednadžbe (5.30) ili jednadžbe (5.29), tada lokalni makroskopski zakoni sačuvanja i dalje vrijede u istom obliku (5.24), koji je ekvivalentan jednadžbama (5.25).

Bitno je spomenuti da jednadžbe (5.29) i (5.30) nije moguće koristiti u predikcijama pomoću načela najveće informacijske entropije, jer ne raspolažemo informacijom o članu $H_{ni}(x,p,t)$ potpune Hamiltonove funkcije $H_{tot}(x,p,t)$. Tu je važno napomenuti i slijedeće: ako neka funkcija $\tilde{f}(x,p,t)$ zadovoljava Liouvilleovu jednadžbu (5.28) i jednadžbe (5.29), tada ta funkcija zadovoljava i jednadžbe (5.25). Obrat ne vrijedi.

Jednadžbe (5.29) i (5.30) se mogu interpretirati na slijedeći način. Jednadžbama (5.30) tvrdi se da informacija o mikroskopskoj dinamici *koja nam nedostaje* nije relevantna za opis vremenske evolucije lokalnih dinamičkih varijabli $n(\mathbf{r})$, $P_\alpha(\mathbf{r})$ i $h(\mathbf{r})$. Jednadžbama (5.29) tvrdi se da nedostajuća informacija o mikroskopskoj dinamici nije relevantna za opis vremenske evolucije lokalnih makroskopskih veličina $\langle n(\mathbf{r})\rangle_t$, $\langle P_\alpha(\mathbf{r})\rangle_t$ i $\langle h(\mathbf{r})\rangle_t$. Obje tvrdnje su u skladu s pretpostavkom da je reduciran opis neravnotežnih makroskopskih sustava moguć na specificiranim vremenskim skalama, što je detaljno raspravljeno u prethodnom odjeljku. Ova pretpostavka se može prihvatiti kao posljedica temeljnog načela makroskopske reproducibilnosti.

Iz jednadžbe (3.21) i definicije (3.26) uvjetne raspodjele vjerojatnosti $D(q,p,t|(q_0,p_0)_\omega,t_0)$, slijedi da se jednadžbe (5.25) mogu zapisati i u slijedećem obliku:

$$\begin{aligned}
\int_M \int_{S_0(M)} n(\mathbf{r}) \left(\frac{\partial D}{\partial t} + \{D, H\} \right) F \, dS_0 \, d\Gamma &= 0 \, , \\
\int_M \int_{S_0(M)} \mathbf{P}(\mathbf{r}) \left(\frac{\partial D}{\partial t} + \{D, H\} \right) F \, dS_0 \, d\Gamma &= 0 \, , \\
\int_M \int_{S_0(M)} h(\mathbf{r}) \left(\frac{\partial D}{\partial t} + \{D, H\} \right) F \, dS_0 \, d\Gamma &= 0 \, .
\end{aligned} \tag{5.31}$$

Najvažniji cilj diskusije u ovom odjeljku bilo je pokazati *ekvivalenciju* makroskopskih jednadžbi kontinuiteta (5.24) s jednadžbama (5.31). Također, jednako važno je bilo provjeriti uvjete pod kojim ta ekvivalencija vrijedi. Pokazano je da su to upravo uvjeti koji su najvažniji za reduciran opis sustava. S obzirom da makroskopske jednadžbe kontinuiteta omogućuju daljnje izvođenje jednadžbi hidrodinamike, na taj način te jednažbe predstavljaju osnovne elemente reduciranog opisa vremenske evolucije sustava na hidrodinamičkoj skali. U odnosu na početni MaxEnt pristup koji je dan u odjeljku 4.2, makroskopske jednadžbe kontinuiteta (5.24) predstavljaju relevantnu dodatnu informaciju koja je temeljna za opis hidrodinamičke vremenske evolucije sustava. U slijedećem odjeljku će biti pokazano da je, uz prethodno uvedena ograničenja (4.16) i (4.30), uvođenje dodatnih ograničenja (5.31) na maksimizaciju uvjetne informacijske entropije $S_I^{DF}(t, t_0)$, dovoljno za određenije predikcije i za elementaran opis ireverzibilnosti na hidrodinamičkoj vremenskoj skali. U posebnom dijelu rasprave će preciznije biti definirani rubni uvjeti na granicama volumena sustava.

5.4 MaxEnt i hidrodinamička ireverzibilna vremenska evolucija

U osnovnom MaxEnt pristupu jedina ograničenja na maksimizaciju uvjetne informacijske entropije $S_I^{DF}(t, t_0)$ bila su uvjet normiranosti (4.16) i Liouvilleova jednadžba usrednjena po dostupnom faznom prostoru (4.30). Pristup koji će sada biti izložen uključuje oba navedena ograničenja. Jedina razlika u odnosu na osnovni pristup je uvođenje dodatnih ograničenja (5.31). Rubni uvjeti na granici područja integracije $M \times (t_0, t_a)$ integrala (4.13) jednaki su rubnim uvjetima koji su uvedeni u osnovnom pristupu. Ovi rubni uvjeti su detaljno obrazloženi u odjeljku 4.2 i zbog toga ovdje neće biti posebno komentirani. Zbog navedenih razloga, rezultati koji će ovdje biti izloženi razlikovat će se odnosu na rezultate osnovnog pristupa isključivo kao posljedica uvođenja dodatnih ograničenja (5.31). Jednadžbe koje u ovom odjeljku slijede iz varijacijskog računa mogu se i izravno dobiti iz jednadžbi koje su izvedene u osnovnom pristupu, uz odgovarajuću modifikaciju kojom se uvažavaju i dodatna ograničenja (5.31).

Poštujući notaciju koja je uvedena u odjeljku 4.2, ograničenja (5.31) su napisana ovdje u obliku koji je jednostavan za korištenje u varijacijskom računu,

$$
\begin{aligned}
&\varphi_n(\mathbf{r}, t, D) = \\
&= \int_M \int_{S_0(M)} n(\mathbf{r}) \left[\frac{\partial D}{\partial t} + \sum_{i=1}^{3N} \left(\frac{\partial D}{\partial x_i} \frac{\partial H}{\partial p_i} - \frac{\partial D}{\partial p_i} \frac{\partial H}{\partial x_i} \right) \right] F \, dS_0 \, d\Gamma = 0 \ , \\
&\varphi_{P_\alpha}(\mathbf{r}, t, D) = \\
&= \int_M \int_{S_0(M)} P_\alpha(\mathbf{r}) \left[\frac{\partial D}{\partial t} + \sum_{i=1}^{3N} \left(\frac{\partial D}{\partial x_i} \frac{\partial H}{\partial p_i} - \frac{\partial D}{\partial p_i} \frac{\partial H}{\partial x_i} \right) \right] F \, dS_0 \, d\Gamma = 0 \ , \\
&\varphi_h(\mathbf{r}, t, D) = \\
&= \int_M \int_{S_0(M)} h(\mathbf{r}) \left[\frac{\partial D}{\partial t} + \sum_{i=1}^{3N} \left(\frac{\partial D}{\partial x_i} \frac{\partial H}{\partial p_i} - \frac{\partial D}{\partial p_i} \frac{\partial H}{\partial x_i} \right) \right] F \, dS_0 \, d\Gamma = 0 \ ,
\end{aligned}
\tag{5.32}
$$

gdje indeks $\alpha = 1, 2, 3$ označava Kartezijeve komponente vektora $\mathbf{P}(\mathbf{r})$. U varijacijski problem se dodatna ograničenja (5.32) uvode pomoću odgovarajućih dodatnih Lagrangeovih multiplikatora $\lambda_n \equiv \lambda_n(\mathbf{r}, t)$, $\lambda_{P_\alpha} \equiv \lambda_{P_\alpha}(\mathbf{r}, t)$, $\lambda_h \equiv \lambda_h(\mathbf{r}, t)$ i pripadajućih funkcionala

$$
\begin{aligned}
C_n[D, \lambda_n] &= \int_{t_0}^{t_a} \int \lambda_n(\mathbf{r}, t) \, \varphi_n(\mathbf{r}, t, D) \, d^3\mathbf{r} \, dt \ , \\
C_{P_\alpha}[D, \lambda_{P_\alpha}] &= \int_{t_0}^{t_a} \int \lambda_{P_\alpha}(\mathbf{r}, t) \, \varphi_{P_\alpha}(\mathbf{r}, t, D) \, d^3\mathbf{r} \, dt \ , \\
C_h[D, \lambda_h] &= \int_{t_0}^{t_a} \int \lambda_h(\mathbf{r}, t) \, \varphi_h(\mathbf{r}, t, D) \, d^3\mathbf{r} \, dt \ ,
\end{aligned}
\tag{5.33}
$$

gdje su $\varphi_n(\mathbf{r}, t, D)$, $\varphi_{P_\alpha}(\mathbf{r}, t, D)$ $(\alpha = 1, 2, 3)$ i $\varphi_h(\mathbf{r}, t, D)$ ograničenja dana jednadžbama (5.32).

Rezultati dobiveni varijacijskim računom će ovdje biti samo navedeni. Objašnjenja potrebnih pravila varijacijskog računa su dana u odjeljku 4.2. Eulerova jednadžba, analogna jednadžbi (4.32), sada je

$$
\begin{aligned}
&\lambda_0 \left\{ \frac{\partial K}{\partial D} - \frac{d}{dt}\left(\frac{\partial K}{\partial(\partial_t D)}\right) - \sum_{i=1}^{3N}\left[\frac{d}{dx_i}\left(\frac{\partial K}{\partial(\partial_{x_i} D)}\right) + \frac{d}{dp_i}\left(\frac{\partial K}{\partial(\partial_{p_i} D)}\right)\right]\right\} \\
&-F\lambda_1 + F\frac{\partial \lambda_2}{\partial t} + F\int \left(\frac{\partial \lambda_n}{\partial t}n + \frac{\partial \lambda_h}{\partial t}h + \sum_{\alpha=1}^{3}\frac{\partial \lambda_{P_\alpha}}{\partial t}P_\alpha\right) d^3\mathbf{r} \\
&+F\int \left(\lambda_n\{n,H\} + \lambda_h\{h,H\} + \sum_{\alpha=1}^{3}\lambda_{P_\alpha}\{P_\alpha,H\}\right) d^3\mathbf{r} \\
&= 0 \ . \qquad (5.34)
\end{aligned}
$$

Član pomnožen s λ_0 u Eulerovoj jednadžbi (5.34) jednak je nuli, pa iz nje slijedi

$$
\begin{aligned}
&\frac{\partial \lambda_2}{\partial t} + \int \left(\frac{\partial \lambda_n}{\partial t}n + \frac{\partial \lambda_h}{\partial t}h + \sum_{\alpha=1}^{3}\frac{\partial \lambda_{P_\alpha}}{\partial t}P_\alpha\right) d^3\mathbf{r} \\
&+\int \left(\lambda_n\{n,H\} + \lambda_h\{h,H\} + \sum_{\alpha=1}^{3}\lambda_{P_\alpha}\{P_\alpha,H\}\right) d^3\mathbf{r} \\
&= \lambda_1 \ . \qquad (5.35)
\end{aligned}
$$

Eulerov rubni uvjet, analogan jednadžbi (4.33), daje

$$
\begin{aligned}
&\left[\frac{\partial K}{\partial(\partial_t D)} - F\lambda_2 - F\int \left(\lambda_n n + \lambda_h h + \sum_{\alpha=1}^{3}\lambda_{P_\alpha}P_\alpha\right) d^3\mathbf{r}\right]_{t=t_a} \\
&= -F\left[\log D + \lambda_2 + \int \left(\lambda_n n + \lambda_h h + \sum_{\alpha=1}^{3}\lambda_{P_\alpha}P_\alpha\right) d^3\mathbf{r}\right]_{t=t_a} \\
&= 0 \ . \qquad (5.36)
\end{aligned}
$$

Iz razloga koji su bili objašnjeni u odjeljcima 4.2, 5.1 i 5.2 u daljnjem tekstu se podrazumijeva da je $t - t_0 \gg \tau$, osim ako se eksplicitno ne tvrdi suprotno. Iz jednadžbe (5.36) slijedi

$$
\begin{aligned}
&D(x,p,t|(x_0,p_0)_\omega,t_0) = \exp\left[-\lambda_2((x_0,p_0)_\omega,t_0;t)\right] \times \\
&\times \exp\left\{-\int \lambda_n(\mathbf{r},t)\, n(\mathbf{r};\mathbf{r}_1,\mathbf{p}_1,\ldots,\mathbf{r}_N,\mathbf{p}_N)\, d^3\mathbf{r}\right\} \times \\
&\times \exp\left\{-\int \lambda_h(\mathbf{r},t)\, h(\mathbf{r};\mathbf{r}_1,\mathbf{p}_1,\ldots,\mathbf{r}_N,\mathbf{p}_N)\, d^3\mathbf{r}\right\} \times \\
&\times \exp\left\{-\int \left[\sum_{\alpha=1}^{3}\lambda_{P_\alpha}(\mathbf{r},t)\, P_\alpha(\mathbf{r};\mathbf{r}_1,\mathbf{p}_1,\ldots,\mathbf{r}_N,\mathbf{p}_N)\right] d^3\mathbf{r}\right\} \ . \qquad (5.37)
\end{aligned}
$$

Varijable u uvjetnoj raspodjeli vjerojatnosti (5.37) su eksplicitno naznačene (na način objašnjen jednadžbama (5.1), (5.2) i (5.3)) zbog boljeg razumijevanja izvoda koji slijede.

Iz uvjeta normiranosti (4.16) i jednadžbe (5.37) slijedi da je

$$\lambda_2((x_0, p_0)_\omega, t_0; t) = \lambda_2(t) \ . \tag{5.38}$$

Iz jednadžbi (5.37) i (5.38) slijedi da je Lagrangeov muliplikator $\lambda_2(t)$ povezan s normalizacijskim faktorom $Z(t)$ uvjetne raspodjele vjerojatnosti na slijedeći način:

$$\lambda_2(t) = \log Z(t) = \log \left\{ \int_M d\Gamma \exp \left[- \int d^3\mathbf{r} \left(\lambda_n n + \lambda_h h + \sum_{\alpha=1}^{3} \lambda_{P_\alpha} P_\alpha \right) \right] \right\} \ . \tag{5.39}$$

U standardnom MaxEnt formalizmu normalizacijski faktor $Z(t)$ naziva se particijska funkcija, odnosno u ovom slučaju, *particijski funkcional*

$$Z(t) \equiv Z_t \left[\lambda_n(\mathbf{r}, t), \lambda_h(\mathbf{r}, t), \lambda_{P_\alpha}(\mathbf{r}, t) \right] \equiv Z_t \ . \tag{5.40}$$

Korištenjem (5.37), (5.38), (5.39) i (5.40), MaxEnt uvjetna raspodjela vjerojatnosti može se zapisati u standardnom obliku

$$\begin{aligned} D(x, p, t | (x_0, p_0)_\omega, t_0) &= \frac{1}{Z_t} \times \\ &\times \exp \left\{ - \int d^3\mathbf{r} \left[\lambda_n(\mathbf{r}, t)\, n(\mathbf{r}) + \lambda_h(\mathbf{r}, t)\, h(\mathbf{r}) \right] \right\} \times \\ &\times \exp \left\{ - \int d^3\mathbf{r} \left[\sum_{\alpha=1}^{3} \lambda_{P_\alpha}(\mathbf{r}, t)\, P_\alpha(\mathbf{r}) \right] \right\} \ . \end{aligned} \tag{5.41}$$

Raspodjela vjerojatnosti mikrostanja $f(x, p, t)$, koja slijedi iz MaxEnt uvjetne raspodjele vjerojatnosti (5.41), dobije se pomoću jednadžbe (3.21), korištenjem jednadžbe (3.26) i uvjeta normiranosti raspodjele vjerojatnosti putanja $F((x_0, p_0)_\omega, t_0)$. Odgovarajuća raspodjela vjerojatnosti mikrostanja je jednaka

$$\begin{aligned} f(x, p, t) &= \frac{1}{Z_t} \times \\ &\times \exp \left\{ - \int d^3\mathbf{r} \left[\lambda_n(\mathbf{r}, t)\, n(\mathbf{r}) + \lambda_h(\mathbf{r}, t)\, h(\mathbf{r}) \right] \right\} \times \\ &\times \exp \left\{ - \int d^3\mathbf{r} \left[\sum_{\alpha=1}^{3} \lambda_{P_\alpha}(\mathbf{r}, t)\, P_\alpha(\mathbf{r}) \right] \right\} \ . \end{aligned} \tag{5.42}$$

Iz jednadžbi (5.41) i (5.42) odmah se uoči slijedeća jednakost

$$D(x, p, t | (x_0, p_0)_\omega, t_0) = f(x, p, t) \ . \tag{5.43}$$

Također, iz jednadžbi (4.4), (4.7), (5.41) i (5.42) slijedi da su u vremenskom trenutku t informacijska entropija $S_I^f(t)$ i uvjetna informacijska entropija $S_I^{DF}(t, t_0)$ jednake:

$$
\begin{aligned}
S_I^f(t) &= S_I^{DF}(t, t_0) = \log Z_t \\
&+ \int d^3\mathbf{r} \left(\lambda_n(\mathbf{r}, t)\langle n(\mathbf{r})\rangle_t + \lambda_h(\mathbf{r}, t)\langle h(\mathbf{r})\rangle_t + \sum_{\alpha=1}^{3} \lambda_{P_\alpha}(\mathbf{r}, t)\langle P_\alpha(\mathbf{r})\rangle_t \right) . \qquad (5.44)
\end{aligned}
$$

Iz jednadžbi (5.43) i (5.44) slijedi da su početne putanje u faznom prostoru u trenutku t_0 i konačna mikrostanja u trenutku t *statistički nezavisni*, što za logičku posljedicu ima *potpuni gubitak korelacije.* Ovaj rezultat dodatno potvrđuje ispravnost uvođenja uvjeta $t - t_0 \gg \tau$, gdje τ predstavlja vrijeme potrebno za gubitak korelacije između početnih putanja u faznom prostoru i konačnih mikrostanja. Razlozi za uvođenje ovog uvjeta bili su detaljno argumentirani u odjeljcima 4.2, 5.1 i 5.2.

Raspodjela vjerojatnosti mikrostanja $f(x, p, t)$ dana jednadžbom (5.42) identična je po obliku *relevantnoj raspodjeli za klasični fluid u lokalnoj ravnoteži* poznatoj iz literature [24, str. 101, jed. 2.2.5]. Uz pretpostavku lokalne ravnoteže, jednostavnom usporedbom dviju raspodjela dobiju se slijedeće identifikacije Lagrangeovih multiplikatora:

$$
\begin{aligned}
\lambda_n(\mathbf{r}, t) &= -\beta(\mathbf{r}, t) \left(\mu(\mathbf{r}, t) - \frac{1}{2} m\mathbf{u}^2(\mathbf{r}, t) \right) \\
\lambda_{P_\alpha}(\mathbf{r}, t) &= -\beta(\mathbf{r}, t) u_\alpha(\mathbf{r}, t) \\
\lambda_h(\mathbf{r}, t) &= \beta(\mathbf{r}, t) . \qquad (5.45)
\end{aligned}
$$

U referenci [24, str. 101–104] je pokazano da funkcija $k^{-1}\beta(\mathbf{r}, t)^{-1} = T(\mathbf{r}, t)$ ima ulogu lokalne temperature, $\mu(\mathbf{r}, t)$ lokalnog kemijskog potencijala po čestici, a da je $\mathbf{u}(\mathbf{r}, t)$ brzina hidrodinamičkog gibanja. Zbog fokusa na rezultate koji se izlažu u ovom odjeljku taj dokaz nema potrebe reproducirati.

U skladu s diskusijom iz odjeljka 5.2, pretpostavka lokalne ravnoteže daje precizniju fizikalnu definiciju uvjeta $t - t_0 \gg \tau$; vrijeme τ potrebno za potpuni gubitak korelacije između početnih putanja u faznom prostoru i konačnih mikrostanja dovedeno je u vezu s vremenom τ_r potrebnim za uspostavljanje stanja lokalne ravnoteže fluida s odgovarajućom raspodjelom koja je dana jednadžbama (5.42) i (5.45).

Deriviranjem po vremenu izraza (5.39) za $\log Z_t$, gdje je Z_t particijski funkcional (5.40), dobije se izraz

$$
\frac{d \log Z_t}{dt} = \frac{1}{Z_t}\frac{dZ_t}{dt} = -\int d^3\mathbf{r} \left(\frac{\partial \lambda_n}{\partial t}\langle n\rangle_t + \frac{\partial \lambda_h}{\partial t}\langle h\rangle_t + \sum_{\alpha=1}^{3} \frac{\partial \lambda_{P_\alpha}}{\partial t}\langle P_\alpha\rangle_t \right) . \qquad (5.46)
$$

Derivacija po vremenu informacijske entropije $S_I^f(t)$ dane jednadžbom (5.44) izračuna se pomoću jednadžbe (5.46), ograničenja (5.31) i jednadžbi (5.24) koje su ekvivalentne tim ograničenjima,

$$\begin{aligned}\frac{dS_I^f(t)}{dt} &= \int d^3\mathbf{r}\left(\lambda_n\frac{\partial\langle n\rangle_t}{\partial t}+\lambda_h\frac{\partial\langle h\rangle_t}{\partial t}+\sum_{\alpha=1}^{3}\lambda_{P_\alpha}\frac{\partial\langle P_\alpha\rangle_t}{\partial t}\right)\\ &= \int d^3\mathbf{r}\left(\lambda_n\langle\{n,H\}\rangle_t+\lambda_h\langle\{h,H\}\rangle_t+\sum_{\alpha=1}^{3}\lambda_{P_\alpha}\langle\{P_\alpha,H\}\rangle_t\right)\\ &= -\int d^3\mathbf{r}\left(\lambda_n\nabla\cdot\langle\,\mathbf{J}\,\rangle_t+\lambda_h\nabla\cdot\langle\mathbf{J}_h\rangle_t+\sum_{\alpha=1}^{3}\lambda_{P_\alpha}\nabla\cdot\langle\mathbf{J}_{P_\alpha}\rangle_t\right)\ .\end{aligned} \tag{5.47}$$

Iz posljednjeg reda jednadžbe (5.47) se dobije da je derivacija po vremenu informacijske entropije $S_I^f(t)$ jednaka

$$\begin{aligned}\frac{dS_I^f(t)}{dt} = &-\int d^3\mathbf{r}\left[\nabla\cdot(\lambda_n\langle\,\mathbf{J}\,\rangle_t)+\nabla\cdot(\lambda_h\langle\mathbf{J}_h\rangle_t)+\sum_{\alpha=1}^{3}\nabla\cdot(\lambda_{P_\alpha}\langle\mathbf{J}_{P_\alpha}\rangle_t)\right]\\ &+\int d^3\mathbf{r}\left[\nabla(\lambda_n)\cdot\langle\,\mathbf{J}\,\rangle_t+\nabla(\lambda_h)\cdot\langle\mathbf{J}_h\rangle_t+\sum_{\alpha=1}^{3}\nabla(\lambda_{P_\alpha})\cdot\langle\mathbf{J}_{P_\alpha}\rangle_t\right]\ .\end{aligned} \tag{5.48}$$

Usrednjavanjem jednadžbe (5.35) po raspodjeli vjerojatnosti mikrostanja i korištenjem jednadžbi (5.38), (5.39), (5.46) i (5.47) jednostavno se pokaže da Lagrangeov multiplikator λ_1 ovisi jedino o vremenu

$$\lambda_1((x_0,p_0)_\omega,t_0;t)=\lambda_1(t)\ , \tag{5.49}$$

i da je jednak derivaciji po vremenu informacijske entropije $S_I^f(t)$,

$$\lambda_1(t)=\frac{dS_I^f(t)}{dt}\ . \tag{5.50}$$

Jednadžba (5.50) potvrđuje interpretaciju koja je Lagrangeovom multiplikatoru λ_1 dana u osnovnom pristupu iz odjeljka 4.2. U generalizaciji osnovnog MaxEnt pristupa koja je uvedena u ovom odjeljku, uvođenje dodatnih ograničenja (5.31) ekvivalentnih hidrodinamičkim jednadžbama kontinuiteta (5.24) preciznije je odredilo brzinu promjene entropije danu Lagrangeovim multiplikatorom λ_1; ona je sada određena jednadžbom (5.48). Korištenjem jednadžbe (5.48) sada će biti pokazano da je ova brzina promjene entropije jednaka odgovarajućem standardnom izrazu iz termodinamike ireverzibilnih procesa.

Iz literature [13], [39] je poznato da se gustoće struja u makroskopskim zakonima sačuvanja (5.24) mogu napisati u slijedećem obliku:

$$\begin{aligned}m\langle\mathbf{J}(\mathbf{r})\rangle_t &= \rho(\mathbf{r},t)\mathbf{u}(\mathbf{r},t)\ ,\\ \langle J_{P_\alpha,\ \beta}(\mathbf{r})\rangle_t &= \rho(\mathbf{r},t)u_\alpha(\mathbf{r},t)u_\beta(\mathbf{r},t)+T_{\beta\alpha}(\mathbf{r},t)\ ,\\ \langle J_{h,\ \alpha}(\mathbf{r})\rangle_t &= \rho(\mathbf{r},t)e(\mathbf{r},t)u_\alpha(\mathbf{r},t)+T_{\alpha\beta}(\mathbf{r},t)u_\beta(\mathbf{r},t)+J_{Q,\ \alpha}(\mathbf{r},t)\ .\end{aligned} \tag{5.51}$$

U prvoj jednadžbi $\rho(\mathbf{r},t)$ je gustoća mase. Brzina fluida $\mathbf{u}(\mathbf{r},t)$ je ranije uvedena brzina hidrodinamičkog gibanja. U drugoj i trećoj jednadžbi veličina $\mathbf{T}(\mathbf{r},t)$ odgovara tenzoru tlaka čije su komponente $T_{\beta\alpha}(\mathbf{r},t)$. Uz korištenje Einsteinove konvencije o sumiranju za indekse α,β uvedene u gornjim relacijama, tenzor tlaka je definiran relacijom $dF_\alpha \equiv -dS_\beta T_{\beta\alpha}$, gdje je dF_α Kartezijeva komponenta sile $d\mathbf{F}$ preko infinitezimalnog elementa površine $d\mathbf{S}$. Tenzor tlaka se razlikuje od tenzora napetosti za negativan predznak. U trećoj jednadžbi $\rho(\mathbf{r},t)e(\mathbf{r},t)$ je gustoća energije, gdje je $e(\mathbf{r},t)$ energija po jedinici mase. Gustoća struje topline je označena s $\mathbf{J}_Q(\mathbf{r},t)$.

Gustoće broja čestica, impulsa i energije koje su ranije uvedene u makroskopskim zakonima sačuvanja (5.24), sada su analogno jednadžbama (5.51) dane obliku

$$\begin{aligned} m\langle n(\mathbf{r})\rangle_t &= \rho(\mathbf{r},t) \ , \\ \langle \mathbf{P}(\mathbf{r})\rangle_t &= \rho(\mathbf{r},t)\mathbf{u}(\mathbf{r},t) \ , \\ \langle h(\mathbf{r})\rangle_t &= \rho(\mathbf{r},t)e(\mathbf{r},t) \ . \end{aligned} \tag{5.52}$$

Brzina fluida $\mathbf{u}(\mathbf{r},t)$ može se konzistentno definirati i relacijom

$$\langle \mathbf{J}(\mathbf{r})\rangle_t = \langle n(\mathbf{r})\rangle_t \mathbf{u}(\mathbf{r},t) \ . \tag{5.53}$$

Zbog jednostavnijeg računa uvodimo nove varijable (x',p') koje su sa starim varijablama (x,p) povezane kanonskom transformacijom koja je slijedećeg oblika

$$\begin{aligned} \mathbf{r}_k &= \mathbf{r}'_k \ , \\ \mathbf{p}_k &= \mathbf{p}'_k + m\mathbf{u}(\mathbf{r}'_k,t) \ , \end{aligned} \tag{5.54}$$

gdje $\mathbf{u}(\mathbf{r}'_k,t)$ predstavlja brzinu fluida na položaju $\mathbf{r}'_k$. Jednostavno je provjeriti da je Jacobijan transformacije jednak jedinici. Uz zamjenu varijabli (5.54), izrazi (5.1), (5.2) i (5.3) transformiraju se na slijedeći način

$$\begin{aligned} n(\mathbf{r}) &= n'(\mathbf{r}) \\ \mathbf{P}(\mathbf{r}) &= \mathbf{P}'(\mathbf{r}) + m\mathbf{u}(\mathbf{r},t)n'(\mathbf{r}) \\ h(\mathbf{r}) &= h'(\mathbf{r}) + \mathbf{u}(\mathbf{r},t)\cdot\mathbf{P}'(\mathbf{r}) + \frac{1}{2}m\mathbf{u}^2(\mathbf{r},t)n'(\mathbf{r}) \ . \end{aligned} \tag{5.55}$$

S obzirom da je Jacobijan transformacije (5.54) jednak jedinici, raspodjela vjerojatnosti koja u novim varijablama (x',p') odgovara raspodjeli vjerojatnosti (5.42), dobije se jednostavnim uvrštavanjem izraza (5.45) i (5.55):

$$f'(x',p',t) = \frac{1}{Z_t}\exp\left\{-\int d^3\mathbf{r}\beta(\mathbf{r},t)\left[h'(\mathbf{r}) - \mu(\mathbf{r},t)n'(\mathbf{r})\right]\right\} \ . \tag{5.56}$$

S obzirom da su dane integralima po faznom prostoru, informacijska entropija $S_I^f(t)$ u jednadžbi (5.44) i odgovarajuća particijska funkcija (5.39) invarijantne su na transformaciju (5.54) za koju je Jacobijan transformacije jednak jedinici.[1] S eksplicitno naznačenim novim koordinatama (x', p') informacijska entropija $S_I^f(t)$ je dana izrazom

$$S_I^f(t) = \log Z_t + \int d^3\mathbf{r}\beta(\mathbf{r},t)\left[\langle h'(\mathbf{r})\rangle_t - \mu(\mathbf{r},t)\langle n'(\mathbf{r})\rangle_t\right] \ , \tag{5.57}$$

dok je particijska funkcija dana izrazom

$$\log Z_t = \log\left\{\int_M d\Gamma' \exp\left[-\int d^3\mathbf{r}\beta(\mathbf{r},t)\left[h'(\mathbf{r}) - \mu(\mathbf{r},t)n'(\mathbf{r})\right]\right]\right\} \ . \tag{5.58}$$

Usrednjavanjem lijeve strane jednadžbi (5.55) po raspodjeli vjerojatnosti mikrostanja koja je dana jednadžbom (5.42) i usrednjavanjem desne strane po odgovarajućoj raspodjeli (5.56), te zatim korištenjem relacija (5.52), dobiju se slijedeće jednakosti

$$\begin{aligned}
\langle n'(\mathbf{r})\rangle_t &= \langle n(\mathbf{r})\rangle_t \\
\langle \mathbf{P}'(\mathbf{r})\rangle_t &= 0 \\
\langle h'(\mathbf{r})\rangle_t &= \langle h(\mathbf{r})\rangle_t - \frac{1}{2}m\mathbf{u}^2(\mathbf{r},t)\langle n'(\mathbf{r})\rangle_t \ .
\end{aligned} \tag{5.59}$$

Posljednja od jednadžbi (5.59) nije ništa drugo nego izraz za gustoću unutarnje energije $U(\mathbf{r},t)$. Transformacijom (5.54) je pokazano da particijska funkcija (5.58), raspodjela vjerojatnosti (5.56) i informacijska entropija (5.57) ne ovise o brzini fluida $\mathbf{u}(\mathbf{r},t)$.[2] Ovisnost o vremenu se u izrazu za informacijsku entropiju (5.57) pojavljuje jedino kroz veličine $\beta(\mathbf{r},t)$ i $\mu(\mathbf{r},t)$, na eksplicitan i implicitan način. Iz tog razloga niti vremenske derivacije informacijske entropije (5.57) ne mogu ovisiti o $\mathbf{u}(\mathbf{r},t)$. Zbog invarijantnosti informacijske entropije $S_I^f(t)$ na transformaciju (5.54) isto mora vrijedi vrijediti i za odgovarajuće izraze (5.44), (5.47) i (5.48). To znači da zbog jednostavnijeg računa, svugdje u izrazima (5.44), (5.45), (5.47) i (5.48) slobodno možemo uzeti da je $\mathbf{u}(\mathbf{r},t) = 0$.

Kad se to napravi u jednadžbi (5.48), uz korištenje izraza za Lagrangeove multiplikatore (5.45) i za gustoće struja (5.51), dobije se slijedeći izraz za derivaciju po vremenu informacijske entropije $S_I^f(t)$:

$$\frac{dS_I^f(t)}{dt} = -\int d^3\mathbf{r}\,\nabla\cdot(\beta\mathbf{J}_Q) + \int d^3\mathbf{r}\,\nabla(\beta)\cdot\mathbf{J}_Q \ . \tag{5.60}$$

Množenjem izraza (5.60) Boltzmannovom konstantom k i uvrštavanjem $\beta(\mathbf{r},t) = k^{-1}T(\mathbf{r},t)^{-1}$ dobije se

$$\frac{dS(t)}{dt} = -\int d^3\mathbf{r}\,\nabla\cdot\left(\frac{\mathbf{J}_Q(\mathbf{r},t)}{T(\mathbf{r},t)}\right) + \int d^3\mathbf{r}\,\nabla\left(\frac{1}{T(\mathbf{r},t)}\right)\cdot\mathbf{J}_Q(\mathbf{r},t) \ . \tag{5.61}$$

---

[1] Zanimljivi detalji povezani s tim dani su u referencama [4], [5, 25]

[2] Pri tome se pretpostavlja da je dostupni dio faznog prostora, invarijantni skup M svih mogućih mikrostanja, invarijantan također i na transformaciju (5.54). Naprimjer, ta pretpostavka će biti zadovoljena ako je skup M dan produktom $M = V\times\cdots\times V\times(-\infty,\infty)\times\cdots\times(-\infty,\infty)$, gdje skup V predstavlja volumen sustava kojim su ograničene moguće vrijednosti koordinata svake od N čestica, a u intervalima $(-\infty,\infty)$ dane su moguće vrijednosti komponenti vektora impulsa svake od N čestica.

Jednadžbu (5.61) možemo prepoznati kao jednadžbu iz termodinamike ireverzibilnih procesa koja daje *brzinu promjene entropije sustava.* Prvi član jednadžbe (5.61) je promjena zbog entropije koja se izmjenjuje kroz granicu volumena sustava, pa je sukladno tome predznak tog člana negativan. S obzirom na početnu pretpostavku da je broj čestica u sustavu fiksan, nema izmjene čestica s okolinom; sustav je zatvoren. Čestice ne mogu napuštati volumen sustava pa brzina fluida na granici volumena sustava iščezava; iz tog razloga entropija ne može prolaziti kroz granicu volumena sustava strujanjem fluida. Nadalje, može se u razmatranje uzeti granični slučaj u kojem je volumen sustava beskonačan. Tada nema niti izmjene topline s okolinom, pa prvi član jednadžbe (5.61) u potpunosti iščezava; to je granica u kojoj je sustav izoliran.

Drugi član jednadžbe (5.61) je integral veličine koja je poznata u termodinamci ireverzibilnih procesa kao *gustoća produkcije entropije*, ili alternativno, *izvor entropije.* Jedan od temeljnih postulata termodinamike ireverzibilnih procesa [39] je da taj član uvijek ima kanonski oblik

$$\sigma = \sum_i \mathbf{X}_i \cdot \mathbf{J}_i \ . \tag{5.62}$$

Ovaj kanonski oblik definira termodinamičke struje $\mathbf{J}_i$ i konjugirane termodinamičke sile $\mathbf{X}_i$ označene indeksom i.

Usporedba jednadžbe (5.62) s drugim članom u jednadžbi (5.61) ukazuje da je moguće, slijedeći ovaj temeljni postulat, definirati gustoću produkcije entropije za klasični fluid identičnih čestica koji smo razmatrali, u obliku

$$\sigma(\mathbf{r},t) = \nabla\left(\frac{1}{T(\mathbf{r},t)}\right) \cdot \mathbf{J}_Q(\mathbf{r},t) \ . \tag{5.63}$$

Pri tome su uočljive odgovarajuće identifikacije termodinamičkih sila i struja.

Izraz za produkciju entropije koji je dan jednadžbom (5.63) u skladu je s pristupom iz reference [41], gdje se gustoća struje entropije definira relacijom

$$\mathbf{J}_S = \sum_i F_i \mathbf{J}_i \ . \tag{5.64}$$

U jednadžbi (5.64), $\mathbf{J}_i$ je gustoća struje ekstenzivne veličine A_i. Za lokalne intenzivne veličine F_i se uzima da su to iste funkcije lokalnih ekstenzivnih veličina kao u ravnoteži.[3] To slijedi iz relacije

$$dS = \sum_i F_i dA_i \ , \tag{5.65}$$

gdje je S lokalna entropija, za koju se po definiciji uzima da ima identičnu funkcionalnu ovisnost o lokalnim ekstenzivnim veličinama A_i kao u ravnoteži. Za lokalnu gustoću

[3] Ovdje koristimo oznake koje se koriste u referenci [41]. Zbog toga oznaka F_i sada označava lokalne intenzivne veličine dane jednadžbom (5.65), pa je treba razlikovati od iste oznake koja je u odjeljcima 2.3 i 2.4 korištena za druge veličine.

entropije s uzima se relacija

$$ds = \sum_i F_i da_i \ , \tag{5.66}$$

gdje su a_i gustoće lokalnih ekstenzivnih veličina A_i. Važno je napomenuti da relacija (5.66) ne uključuje član koji sadrži volumen, za razliku od odgovarajuće relacije (5.65).

Rezultati koji su dobiveni ovdje generaliziranim MaxEnt pristupom u skladu su s ovim definicijama. U pristupu danom u referenci [41], ako su ekstenzivne veličine sačuvane, slijedi da je tada gustoća produkcije entropije jednaka

$$\sigma = \sum_i \nabla F_i \cdot \mathbf{J}_i \ . \tag{5.67}$$

Za entropiju vrijedi relacija

$$\sigma = \frac{\partial s}{\partial t} + \nabla \cdot \mathbf{J}_S \ , \tag{5.68}$$

gdje $\partial s/\partial t$ je brzina promjene entropije po jedinici volumena. Izrazi (5.64), (5.67) i (5.68) po svome obliku odgovaraju izrazima (5.61) i (5.63) koji su dobiveni ovdje generaliziranim MaxEnt pristupom. Rasprava iz reference [41] samo dodatno oslikava temeljni problem definicije termodinamičke entropije neravnotežnog sustava.

Poglavlje 6

DISKUSIJA OTVORENIH PITANJA

Ako razmotrimo mogućnost da je naša informacija o mikroskopskoj dinamici nepotpuna, reproducibilnost i gubitak informacije postaju dio opisa makroskopske vremenske evolucije. Ovaj pristup tada vodi do jednostavne definicije produkcije entropije. Ideju da je ireverzibilnost povezana s postupnim gubitkom informacije razvio je Jaynes u okviru formalizma matrice gustoće [2]. Nedavno, Duncan i Semura [42, 43] su predložili ideju da se informacija doista gubi na fundamentalnom nivou. Uzajamno djelovanje kvantne dekoherencije i dinamike se razmatra kao jedan od mogućih razloga iza drugog zakona termodinamike, gdje je produkcija entropije uzrokovana curenjem informacije u okolinu [44, 45]. U našem klasičnom pristupu, gubitak informacije je povezan s nepotpunom informacijom o mikroskopskoj dinamici. Ako se ova ideja pažljivo razmotri, čak i u ovom jednostavnom modelu, nepotpunost informacije mora se uzeti u obzir na neki nepristran način.

Pitanja koja su povezana s nepotpunom informacijom mogu se razmotriti na objektivan način. Analitički principi za te svrhe, tj. za razdvajanje subjektivnih i objektivnih aspekata teorijskog formalizma, mogu se pronaći u teoriji vjerojatnosti. Filozofija ovog pristupa se zasniva na interpretaciji teorije vjerojatnosti kao prirodnog proširenja deduktivne logike. Takvu generalizaciju je na aksiomatski način razvio Jaynes u svojoj studiji o teoriji vjerojatnosti [25]. Standardna aksiomatska teorija vjerojatnosti izvedena je iz ove generalizirane teorije, što samo po sebi upućuje da je generalizirana teorija ispravan alat za uključivanje nove informacije u naše raspodjele vjerojatnosti. Raspodjele vjerojatnosti su u tom smislu interpretirane kao nosači nepotpune informacije. Ovaj pristup se možda najbolje razumije iz opisa koji je dao Jaynes [25]: "...teorija vjerojatnosti kao generalizirana logika vjerodostojnog zaključivanja koja bi trebala biti primjenjiva, u principu, u bilo kojoj situaciji gdje nemamo dovoljnu informaciju koja bi dopustila deduktivno zaključivanje.” Zbog potpunosti ćemo citirati i slijedeće linije iz reference [25], koje su važne za daljnu diskusiju: “Ali to je jednako tako istina i za apstraktne matematičke sustave; kada je tvrdnja neodlučiva u takvom sustavu, to jedino znači da njegovi aksiomi ne pružaju informaciju koja je dovoljna za odlučivost. Ali novi aksiomi, eksterni izvornom skupu, mogu dati informaciju koja nedostaje i ipak učiniti tvrdnju odlučivom.” Možemo zaključiti na osnovi rezultata koji su izloženi u ovom radu, da kada se vjerojatnosti interpretiraju na povezan način kao svojstvo stupnja našeg znanja, uz primjenu MaxEnt algoritma i prediktivne statističke mehanike, da je tada matematički opis ireverzibilnosti kompatibilan s pojmovima Shannonove teorije informacije.

Još jedan objektivan aspekt spomenutog problema povezan je s pitanjima koja je na zanimljiv i spekulativan način postavio Zwick. U njegovom radu [46] o problemu mjerenja u kvantnoj mehanici, teškoće u opisivanju mjerenja na nivou kvantne dinamike uspoređene su, i pronađene su sličnosti s nepotpunosti određenih aksiomatskih sustava u matematici, koje su otkrili i detaljno elaborirali Gödel i drugi [47, 48]. Prema Zwicku, opsežni paralelizam između fizikalnih i matematičkih slučajeva ukazuje na mogućnost da je proces mjerenja samoreferentan kao što je bila Gödelova specijalna formula, i da mjerenje može biti neodlučivo unutar dinamike (formalizma vremenski ovisne Schrödingerove jednadžbe), događajući se jedino na meta-nivou formalizma. Po toj liniji razmišljanja, fizikalna teorija imala bi tada barem dva nivoa; proces mjerenja bio bi opisan na meta-nivou, ali neodlučiv na osnovnom nivou koji je opisan zakonom dinamike. U isto vrijeme je osnovni nivo inherentno nepotpun i ne generiraju se kontradikcije. Zakon dinamike i svi procesi opisani njim su reverziblni, ali proces mjerenja je ireverzibilan; na ovaj način ireverzibilnost bi bila prisutna u opisu procesa mjerenja unutar takve teorije s dva nivoa. Zwick navodi slične sugestije koje je dao Pattee [49] o neophodnosti dva nivoa strukuture i opisa za bilo koji proces predikcije i kontrole (tj. mjerenja). Pitanja koja se postavljaju o ireverzibilnom ponašanju sustava koji su deterministički određeni vremenski simetričnim jednadžbama gibanja, tada bi se činila paradoksalna jedino u kontekstu teorije s jednim nivoom [49].

Bez dubljeg ulaženja u ova pitanja, može se primijetiti u izloženoj primjeni prediktivne statističke mehanike na problem predikcije makroskopske vremenske evolucije sustava s Hamiltonovom dinamikom, da su određene značajke teorije s dva nivoa jasno uočljive. U ovom jednostavnom modelu one se pojavljuju jedino kao rezultat prepoznavanja nepotpunosti naše vlastite informacije o mikroskopskoj dinamici. S pragmatične točke gledišta to nam dopušta daljnju diskusiju o povezanim pitanjima o uzajamnom djelovanju našeg znanja i ograničenja mjerenja na sustav i njegovu “stvarnu” dinamiku.

Poglavlje 7

ZAKLJUČAK

Ovim radom se pokazuje da Jaynesova interpretacija ireverzibilnosti kao posljedice postupnog gubitka informacije o stanju sustava zbog nemogućnosti praćenja njegove točne vremenske evolucije tijekom procesa, ima jasnu matematičku formulaciju u konceptima koji su uvedeni u radu. Najvažniji teorijski koncept u radu bila je maksimizacija uvjetne informacijske entropije uz zadana ograničenja, i njena relacija s informacijskom entropijom, preuzeta iz Shannonove teorije informacije. Pri tome je ključan element teorijskog pristupa bilo uvođenje Liouvilleove jednadžbe za uvjetnu raspodjelu vjerojatnosti kao makroskopskog ograničenja, tj. kao ograničenja koje je dano usrednjavanjem te jednadžbe integralom po dostupnom faznom prostoru. U problem predikcije makroskopske vremenske evolucije zatvorenog Hamiltonovog sustava je time uključena nepotpunost naše informacije o detaljnoj mikroskopskoj dinamici, na način koji je konzistentan s temeljnim načelima prediktivne statističke mehanike. Pokazano je da takav matematički opis rezultira potpunim gubitkom korelacije između početnih putanja u faznom prostoru i konačnih mikrostanja. Ovaj gubitak korelacije je povezan s gubitkom informacije o mogućim mikrostanjima sustava, koji je ovim pristupom doveden u vezu s promjenom entropije sustava. Ta je povezanost omogućila definiciju promjene entropije i brzine promjene entropije za zatvoren Hamiltonov sustav u okviru razmotrenog pristupa bez dodatnih pretpostavki.

Konstrukcijom raspodjele vjerojatnosti uz korištenje načela najveće informacijske entropije, tj. maksimizacijom informacijske entropije uz zadana ograničenja, uključuje se u raspodjelu vjerojatnosti jedino informacija koja je reprezentirana zadanim ograničenjima. Predikcije izvedene iz takve raspodjele vjerojatnosti najbolje su predikcije koje su moguće na osnovi dostupne informacije, bez uvođenja dodatnih pretpostavki u koje nismo sigurni. Ako se kontrolom određenih makroskopskih veličina u eksperimentu reproducira neka makroskopska pojava, u skladu s temeljnim načelom makroskopske reproducibilnosti, tada je informacija o vrijednostima tih veličina upravo ona informacija koja je relevantna za predikciju makroskopske pojave.

Dakle, može se reći da je razmatranje relevantnosti dostupne informacije o sustavu za predikcije i za reproducibilnost makroskopske vremenske evolucije, ključno za bolje razumijevanje pojave ireverzibilnosti. Na primjeru zatvorenog Hamiltonovog sustava je pokazano da se za taj sustav može dati elementaran opis ireverzibilne makroskopske vremenske evolucije, ako se uvođenjem odgovarajućih dodatnih ograničenja na maksimizaciju uvjetne informacijske entropije, u raspodjelu vjerojatnosti uključi ona informacija koja je relevantna za opis neravnotežnog sustava. U okviru pristupa koji je razvijen u ovom radu je tako uvođenjem hidrodinamičkih jednadžbi kontinuiteta kao relevantne informacije na hidrodinamičkoj vremenskoj skali, dobivena brzina promjene entropije i gustoća produkcije entropije za klasični fluid identičnih čestica. Dobiveni izrazi u skladu su s definicijama koje te veličine imaju u termodinamici ireverzibilnih procesa. Ako se u obzir uzme da je prediktivna statistička mehanika općeniti oblik zaključivanja iz dostupne informacije, bez dodatnih pretpostavki, dobiveni rezultati upućuju na važnost njenih temeljnih načela za teoriju ireverzibilnosti.

Bibliografija

[1] Jaynes, E.T.: Information theory and statistical mechanics. Phys. Rev. 106, 620–630 (1957)

[2] Jaynes, E.T.: Information theory and statistical mechanics. II. Phys. Rev. 108, 171–190 (1957)

[3] Gibbs, J.W.: Elementary Principles in Statistical Mechanics. Yale University Press, New Haven (1902)

[4] Shannon, C.E.: A mathematical theory of communication. Bell Syst. Tech. J. 27, 379–423, 623–656 (1948). Reprinted In: Shannon C.E., Weaver, W.: The Mathematical Theory of Communication. University of Illinois Press, Urbana (1949)

[5] Jaynes, E.T.: Information theory and statistical mechanics. In: Ford, K.W. (ed.) 1962 Brandeis Lectures in Theoretical Physics, vol. 3, pp. 181–218. W. A. Benjamin, Inc., New York (1963)

[6] Jaynes, E.T.: Where do we stand on maximum entropy? In: Levine, R.D., Tribus, M. (eds.) The Maximum Entropy Formalism, pp. 15–118. MIT Press, Cambridge (1979)

[7] Jaynes, E.T.: Gibbs vs Boltzmann entropies. Am. J. Phys. 33, 391–398 (1965)

[8] Jaynes, E.T.: The minimum entropy production principle. Ann. Rev. Phys. Chem. 31, 579–601 (1980)

[9] Jaynes, E.T.: Macroscopic prediction. In: Haken, H. (ed.) Complex Systems – Operational Approaches in Neurobiology, Physics, and Computers, pp. 254–269. Springer, Berlin (1985)

[10] Grandy, W.T.: Principle of maximum entropy and irreversible processes. Phys. Rep. 62, 175–266 (1980)

[11] Garrod, C.: Statistical Mechanics and Thermodynamics. Oxford University Press, New York (1995)

[12] Jaynes, E.T.: The second law as physical fact and as human inference. Unpublished manuscript (1990). http://bayes.wustl.edu/etj/node2.html

[13] Grandy, W.T.: Entropy and the Time Evolution of Macroscopic Systems. Oxford University Press, Oxford (2008)

[14] Zurek, W.H.: Algorithmic randomness and physical entropy. Phys. Rev. A 40, 4731–4751 (1989)

[15] Kuić, D., Županović, P., Juretić, D.: Macroscopic time evolution and MaxEnt inference for closed systems with Hamiltonian dynamics. Found. Phys. 42, 319–339 (2012)

[16] Grandy, W.T.: Time evolution in macroscopic systems. I. Equations of motion. Found. Phys. 34, 1–20 (2004)

[17] Grandy, W.T.: Time evolution in macroscopic systems. II. The entropy. Found. Phys. 34, 21–57 (2004)

[18] Grandy, W.T.: Time evolution in macroscopic systems. III. Selected applications. Found. Phys. 34, 771–813 (2004)

[19] Tishby, N.Z., Levine, R.D.: Time evolution via a self-consistent maximal-entropy propagation: the reversible case. Phys. Rev. A 30, 1477–1490 (1984)

[20] Plastino, A.R., Plastino, A.: Statistical treatment of autonomous systems with divergenceless flows. Physica A 232, 458–476 (1996)

[21] Plastino, A., Plastino, A.R., Miller, H.G.: Continuity equations, H-theorems, and maximum entropy. Phys. Lett. A 232, 349–355 (1997)

[22] Plastino, A.R., Plastino, A.: Universality of Jaynes' approach to the evolution of time-dependent probability distributions. Physica A 258, 429–445 (1998)

[23] Schönfeldt, J-H., Jimenez, N., Plastino, A.R., Plastino, A., Casas, M.: Maximum entropy principle and classical evolution equations with source terms. Physica A 374, 573–584 (2007)

[24] Zubarev, D., Morozov, V., Röpke, G.: Statistical Mechanics of Nonequilibrium Processes, Vol. 1: Basic Concepts, Kinetic Theory. Akademie Verlag (1996)

[25] Jaynes, E.T.: Probability Theory: The Logic of Science. Cambridge University Press, Cambridge (2003). Bretthorst, G.L. (ed.)

[26] Jensen inequality. E.K. Godunova (originator), *Encyclopedia of Mathematics*. URL: http://www.encyclopediaofmath.org/

[27] Brillouin, L: Science and Information Theory, 2nd edn. Academic Press, New York (1962)

[28] Balian, R.: From Microphysics to Macrophysics, Vol. 2: Methods and Applications of Statistical Physics. Springer, Berlin (2007)

[29] Outer measure. *Encyclopedia of Mathematics*.
URL: http://www.encyclopediaofmath.org/

[30] Zemansky, M.W., Dittman, R.H.: Heat and Thermodynamics, 7th ed. McGraw-Hill, New York (1997)

[31] Khinchin, A.I.: Mathematical Foundations of Statistical Mechanics. Dover, New York (1949)

[32] Wan, F.Y.M.: Introduction to the Calculus of Variations and Its Applications. Chapman & Hall, New York (1995)

[33] Kittel, C.: Elementary Statistical Physics. Wiley, New York (1958)

[34] Zwanzig, R.: Ensemble method in the theory of irreversibility. J. Chem. Phys. 33, 1338–1341 (1960)

[35] Zwanzig, R.: On the identity of three generalized master equations. Physica 30, 1109–1123 (1964)

[36] Robertson, B.: Equations of motion in nonequilibrium statistical mechanics. Phys. Rev. 144, 151–161 (1966)

[37] Robertson, B.: Application of maximum entropy to nonequilibrium statistical mechanics. In: Levine, R.D., Tribus, M. (eds.) The Maximum Entropy Formalism. MIT Press, Cambridge (1979)

[38] Zubarev, D.N., Kalashnikov, V.P: Equivalence of various methods in the statistical mechanics of irreversible processes. Teoret. Mat. Fiz. 7, 372–394 (1971)

[39] Evans, D.J., Morriss, G.: Statistical Mechanics of Nonequilibrium Liquids, 2nd edn. Cambridge University Press, Cambridge (2008)

[40] Ostrogradski formula. L.D. Kudryavtsev (originator), *Encyclopedia of Mathematics*. URL: http://www.encyclopediaofmath.org/

[41] Callen, H.B.: Thermodynamics and an Introduction to Thermostatistics, 2nd edn. Wiley, New York (1985)

[42] Duncan, T.L., Semura, J.S.: The deep physics behind the second law: information and energy as independent forms of bookkeeping. Entropy 6, 21–29 (2004)

[43] Duncan, T.L., Semura, J.S.: Information loss as a foundational principle for the second law of thermodynamics. Found. Phys. 37, 1767–1773 (2007)

[44] Zurek, W.H., Paz, J.-P.: Decoherence, chaos, and the second law. Phys. Rev. Lett. 72, 2508–2512 (1994)

[45] Zurek, W.H.: Decoherence, einselection, and the quantum origins of the classical. Rev. Mod. Phys. 75, 715–775 (2003)

[46] Zwick, M.: Quantum measurement and Gödel's proof. Speculat. Sci. Technol. 1, 135–145 (1978)

[47] Gödel, K.: On Formally Undecidable Propositions of Principia Mathematica and Related Systems. Basic Books, Inc., New York (1962)

[48] Nagel, E., Newman, J.R.: Gödel's Proof. New York University Press, New York (1960)

[49] Pattee, H.H.: Postscript. In: Pattee, H.H. (ed.) Hierarchy Theory, pp. 131–156. George Braziller, New York (1973)

ŽIVOTOPIS AUTORA

Domagoj Kuić je rođen u Sinju 16. listopada 1976. Osnovnu školu i opću gimnaziju je završio u Splitu. Diplomirao je 2005. godine na smjeru diplomirani inženjer fizike na Prirodoslovno-matematičkom fakultetu Sveučilišta u Zagrebu. Diplomski rad pod naslovom “Produkcija fotona u neravnotežnoj kvark-gluon plazmi” izradio je pod vodstvom mentora dr. sc. Ivana Dadića sa Instituta Ruđer Bošković u Zagrebu.

U razdoblju 2005.-2006. radio je na Institutu za oceanografiju i ribarstvo u Splitu kao stručni suradnik, te u Državnom zavodu za zaštitu od zračenja u Zagrebu, također kao stručni suradnik.

Godine 2005. izabran je kao vanjski suradnik u naslovno zvanje asistent na Prirodoslovno-matematičkom fakultetu Sveučilišta u Splitu, gdje sudjeluje u nastavi na Odjelu za fiziku. Od 2006. zaposlen je na istom fakultetu kao znanstveni novak - asistent. Prijavljen je na projektu Ministarstva znanosti “Razvoj i primjene principa maksimalne proizvodnje entropije”, čiji je voditelj prof. dr. sc. Davor Juretić. U okviru nastave na fakultetu je kao asistent držao vježbe i seminare iz nekoliko različitih kolegija iz fizike.

Poslijediplomski doktorski studij fizike, smjer atomska i molekularna fizika, na Prirodoslovno-matematičkom fakultetu u Zagrebu upisao je 2006. godine.

Područje znanstvenog interesa mu je atomska i molekularna fizika, statistička fizika, termodinamika ireverzibilnih procesa i biofizika. Posebnu pažnju u istraživanju je usmjerio na primjenu načela najveće informacijske entropije i prediktivne statističke mehanike u razvoju teorije ireverzibilnosti. Interesira ga i područje teorije polja, uključujući metode neravnotežne teorije polja i njihove primjene u fizici mnoštva čestica.

POPIS OBJAVLJENIH RADOVA

[1] Kuić, D.: Quantum mechanical virial theorem in systems with translational and rotational symmetry. Int. J. Theor. Phys. 52, 1221-1239 (2013)
[2] Kuić, D., Županović, P., Juretić, D.: Macroscopic time evolution and MaxEnt inference for closed systems with Hamiltonian dynamics. Found. Phys. 42, 319-339 (2012)
[3] Dobovišek, A., Županović, P., Brumen, M., Bonačić-Lošić, Ž., Kuić, D., Juretić, D.: Enzyme kinetics and the maximum entropy production principle. Biophys. Chem. 154, 49-55 (2011)
[4] Županović, P., Kuić, D., Juretić, D., Dobovišek, A.: On the problem of formulating principles in nonequilibrium thermodynamics. Entropy 12, 926-931 (2010)
[5] Županović, P., Botrić, S., Juretić, D., Kuić, D.: Relaxation processes and the maximum entropy production principle. Entropy 12, 473-479 (2010)
[6] Županović, P., Kuić, D., Bonačić-Lošić, Ž., Petrov, D., Juretić, D., Brumen, M.: The maximum entropy production principle and linear irreversible processes. Entropy 12, 996-1005 (2010)