VYTAUTO DIDŽIOJO UNIVERSITETAS

Gamtos mokslų fakultetas

Emilis DAMBAUSKAS

Fizikinė informacinės entropijos lygties interpretacija

(pavadinimas)

Fizikos bakalauro darbas

Vadovas: prof. Liudvikas Pranevičius

(pedagoginis ir mokslo vardas, vardas, pavardė)

Apginta: ....

(data, fakulteto dekano parašas)

KAUNAS 2003

Turinys

1 Įvadas

1.1 Aktualumas

Šiame darbe aptariamas entropijos sąvokos aptarimo būdas atsirado palyginus neseniai. Tam įtakos turėjo tai, kad termodinamikos bei statistinės fizikos mokslų pagrindai ir informacijos teorijos mokslo pagrindai kūrėsi skirtingu metu.

Entropijos sąvoką 1865 m. įvedė vokiečių fizikas R.J.E. Clausius (1822–1888), suformulavęs antrąjį termodinamikos dėsnį. 1876 m. J.W. Gibbs (1839–1903), o 1896 m. ir L. Boltzman (1844–1906) dirbdami nepriklausomai išvedė entropijos priklausomybės nuo termodinaminės sistemos mikrobūsenų lygtis. Taip entropija įsitvirtino fizikoje.

XX a. viduryje pradėjus kurti elektronines skaičiavimo mašinas, 1948 m. JAV matematikas C.E. Shanon (1916–2001) paskelbė savo straipsnį "A Mathematical Theory of Communication", kuris padėjo pamatus informacijos teorijai. Šiame darbe Shanon įvedė entropijos sąvoką, kuria apibrėžė informacijos kiekį, kuriamą duomenų srauto.

Svarbu pabrėžti, kad informacinės entropijos išraiška nebuvo išvesta, o tiesiog "pasiskolinta" iš statistinės fizikos. Visgi toks klasikinės fizikos dėsnio pritaikymas informacijos teorijoje sukėlė mokslininkų diskusijas, kurios aktualios ir šiandien. Tarp kitų darbų, norėčiau paminėti 1956 m. JAV mokslininko E.T. Jaynes (1922–1998) straipsnius "Information Theory and Statistical Mechanics" (1963 m.), kuriuose jis siūlė statistinės fizikos dėsnius išvesti naudojantis tik matematikos šaka statistika, bei iškėlė mintį, kad statistinė fizika yra išvada padaryta turint nepilnus duomenis.

"Matematinė komunikacijos teorija" taip pat paskatino matematikus tiksliau apibrėžti atsitiktinumo (*randomness*) sąvoką. Šioje srityje daug nuveikė rusų mokslininkas A.N. Kolmogorovas (1903–1987). Darbe remiuosi Per Martin-Löf bei P.Vitányi išvadomis.

Iš vėlesnių darbų, kuriais remiuosi svarbūs yra E.T. Jaynes straipsnis "The Gibbs Paradox", paskelbtas 1992 m., taip pat JAV matematiko J.D.H. Smith straipsnis "Some Observations on the Concepts of Information-Theoretic Entropy and Randomness", paskelbtas

2001 m..

1.2 Tikslas

Šio darbo tikslas — aprašyti kai kuriuos informacinės entropijos ir atsitiktinumo savokų aspektus, ypač jų taikymo fizikoje srityse. Taip pat parodyti ryšį tarp informacinės bei fizikinės entropijos.

1.3 Darbo struktūra

Darbe pateikiu fizikinės bei informacinės entropijų apibrėžimus, informacinės entropijos lygties išvedimą, pabrėžiant atsitiktinumo priklausomybę nuo turimos bandymų įrangos.

Kaip šios priklausomybės makroskopinis atitikmuo aptariamas Gibso paradoksas.

1.4 Naudojami terminai

Entropijų apibrėžimus ir pavadinimus pateikiu 2 skyriuje.

Toliau darbe žodį *atsitiktinumas* naudoju kaip pakaitalą anglų kalbos žodžiui *randomness*.

2 Fizikinės bei informacinė entropijos

2.1 Fizikinės entropijos

2.1.1 Termodinaminė entropija

Termodinaminė entropija — tai termodinaminės sistemos būsenos funkcija S, apibūdinanti izoliuotoje sistemoje vykstančių procesų negrįžtamumą. Sistemos entropijos nykstamas pokytis dS, suteikus sistemai šilumos kiekį δQ, reiškiamas sąryšiu:

$$dS \geq \frac{\delta Q}{T}. \tag{1}$$

T — sistemos absoliuti temperatūra. Lygybės ženklas galioja grįžtamiesiems procesams, nelygybės — negrįžtamiesiems. [1, p. 359]

2.1.2 Statistinė entropija

Entropijos prasmę atskleidžia statistinės fizikos formuluotė: sistemos entropija yra tiesiog proporcinga termodinaminės tikimybės, sistemai turėti vidinę energiją U, natūriniam logaritmui:

$$S = k \ln \Omega(U). \tag{2}$$

Čia $\Omega(U)$ — energijos vertės U išsigimimo eilė arba laipsnis. Kadangi šis dydis nusako energiją U atitinkantį mikrobūsenų skaičių (užimamą fazinės erdvės tūrį), tai jį dar galima vadinti energijos U mikrobūsenos termodinamine tikimybe. [2, p. 112,109]

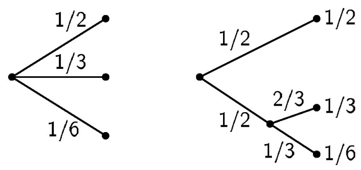
2.2 Informacinė entropija

Informacijos teorijoje entropija — tai dydis, nusakantis duomenyse esantį informacijos kiekį.

Tarkime, kad turime galimų įvykių aibę, kurių pasitaikymo tikimybės yra $p_1, p_2, \ldots, p_n$. Šios tikimybės yra žinomos, bet tai ir yra viskas, ką mes žinome apie tai, kuris įvykis galėtų atsitikti. Ar galėtume surasti tokį dydį, kuris aprašytų kiek "pasirinkimo" yra įvykių atsitikime arba kokia yra mūsų nežinomybė apie rezultatus?

Būtų pagrįsta iš tokio dydžio $H(p_1, p_2, \ldots, p_n)$ tikėtis tokių savybių:

1. H turėtų būti tolydus visiems p_i.
2. Jei visos tikimybės p_i yra lygios ($p_i = \frac{1}{n}$), tai H turėtų būti funkcija, kurios rezultatas tiesiogiai priklauso nuo n. Esant vienodai tikėtiniems įvykiams yra daugiau pasirinkimo ir nežinomybės, kai įvykių skaičius didėja.
3. Jei pasirinkimas būtų suskaidytas į du pasirinkimus, einančius vienas po kito, tai bendras H turėtų būti atskirų pasirinkimų H reikšmių suma.

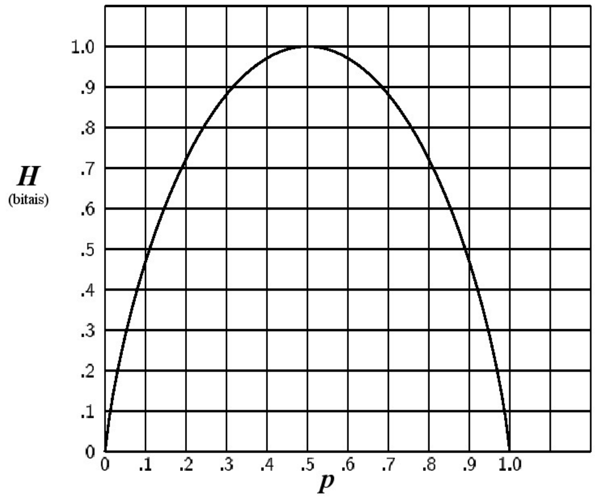


1: Pasirinkimo suskaidymas

Vienintelis dydis H atitinkantis šiuos kriterijus gali būti išreiškiamas lygtimi:

$$H = -K \sum_{i=1}^{n} p_i \ln p_i. \tag{3}$$

Čia H — informacinė entropija, K — konstanta (priklausanti tik nuo matavimo vienetų pasirinkimo), p_i — konkretaus įvykio tikimybė.



2: Entropija esant dviems galimybėms su tikimybėmis p ir $q = 1 - p$. Parodyta $H = -(p \ln p + q \ln q)$ priklausomybė nuo p.

Informacinė entropija pasižymi šiomis savybėmis:

1. $H = 0$ tada ir tik tada, kai visos tikimybės p_i išskyrus vieną yra lygios nuliui. Tik kai esame užtikrinti bandymo rezultatais (negauname jokios naujos informacijos) informacinė entropija išnyksta.
2. Duotam skaičiui rezultatų n, H vertė yra maksimali ir lygi $\ln n$, kai visos tikimybės p_i yra vienodos (lygios $\frac{1}{n}$). Tai — pati neapibrėžčiausia situacija.
3. Dviems įvykiams x ir y bendra entropija $H(x, y) \leq H(x) + H(y)$. Lygybė galioja tik tada, kai įvykiai yra nepriklausomi.
4. Bet koks pokytis vedantis prie tikimybių p_i suvienodinimo didina entropiją.

[3, p. 10–15]

3 Atsitiktinumas, tikimybė ir entropija

Martin-Löf (1966) pasiūlytas atsitiktinumo sąvokos apibrėžimas [4, p. 24–25] padeda nuosekliau prieiti prie informacinės entropijos, ką toliau ir nagrinėju.

Įsivaizduokime mokslinį eksperimentą su N galimais rezultato variantais. Modernaus mokslo susitarimu, “mokslinis” eksperimentas turi būti įmanomas pakartoti ir atgaminti. Pagal Martin-Löf Eksperimento rezultatai laikomi atsitiktiniais jei nėra jokio statistinio patikrinimo, prieinamo eksperimentuotojui, kuriuo jis galėtų atpažinti bet kokį dėsningumą, po pakartotinių eksperimento kartojimų. Labai svarbu pabrėžti, kad šis apibrėžimas priklauso nuo turimos įrangos galingumo. Žaidimų kauliuko ridenimas kazino salygomis turi būti atsitiktinis, iš kitos pusės, jei turėtume tokią įrangą kaip didelio greičio viedeokamera, būtų įmanoma iš anksto nuspėti kiekvieno ridenimo rezultatus, nuo pradinio kauliuko judėjimo momento ir eksperimento rezultatai nebegalėtų būti laikomi atsitiktiniais. Subtilesnis skirtumas atsirastų, jei ridentume truputį išcentruotą kauliuką kazino sąlygomis. Jei ridenimų skaičius reikalingas net ir sudėtingiausiam statistiniam patikrinimui parodyti rezultatus yra didesnis už tą nuo kurio kauliukas susidėvėtų, eksperimento rezultatai būtų vistiek laikomi atsitiktiniais.

Pagrindinės entropijos bei tikimybės sąvokos atsiranda iš atsitiktinumo sąvokos [5, p. 3-4]. Eksperimento, kurio rezultatai yra atsitiktiniai ir kuris turi N galimų rezultatų variantų (natūrali) entropija yra:

$$H = \ln(N), \tag{4}$$

o dvejetainė entropija (bitais):

$$H = \log_2(N). \tag{5}$$

Naudojant logaritmą, kurio pagrindas 2, entropija matuojama bitais. Bet kurio eksperimento rezultato x tikimybė $\pi(x)$ yra:

$$\pi(x) = N^{-1}. \tag{6}$$

Tikimybė $\pi(x)$ ir natūrali entropija H yra susiję abipusiai priešingomis lygybėmis:

$$H = -\ln \pi(x), \tag{7a}$$

$$\pi(x) = e^{-H}. \tag{7b}$$

Dvejetainei entropijai H:

$$H = -\log_2 \pi(x), \tag{8a}$$

$$\pi(x) = 2^{-H}. \tag{8b}$$

Tikimybė $\pi(x)$ parodo, kokią sumą apsimoka statyti tokiame žaidime, kuriame laimėtume, pastatę už eksperimento rezultatą x. Eksperimento atsitiktinumas, reiškia, kad nėra jokios strategijos kuri padėtų mums laimėti šį žaidimą. Entropija (4) ir (5) lygybėse apibrėžia eksperimento rezultatų nežinojimą. Jei eksperimentą atliktų kasnors kitas, o vėliau paklaustume kokie buvo eksperimento rezultatai, tai reikalingas atsakymų į kuriuos būtų galima atsakyti tik “taip” arba “ne” skaičius atitiktų (5) lygybės rezultatą.

Aukščiau aprašytas modelis yra per daug siauras bendram naudojimui, kai norima jį taikyti eksperimentams su neatsitiktiniais rezultatais (šiuo turime omenyje bet kokius eksperimentus, kurių rezultatus statistiškai būtų galima laikyti atsitiktiniais, bet turinčius netolygų baigtinį ir dėsningą tikimybių pasiskirstymą). Šiems eksperimentams modeliuoti galima naudoti vidinį eksperimentą su atsitiktiniais rezultatais (kaip praeitoje pastraipoje), kurio rezultatų aibė, vadinama *fazine erdve*, turi N elementų. Fazinę erdvę galima visiškai suskaidyti į aibę $\xi = \{C_1, \ldots, C_r\}$ susidedančią iš skirtingų poaibių, vadinamų *būsenomis*. Skaidymas ξ atspindi eksperimentą su neatsitiktiniais rezultatais (taip pat žymimą ξ), kuriuo atrenkamas rezultatas x iš fazinės erdvės ir randama būsena C_i, kuriai jis priklauso. Jei būsenoje C_i yra n_i eksperimento, kurio kiekvieno rezultato tikimybė yra N^{-1} (pagal (6), rezultatų, tuomet būsenos C_i tikimybė $p(C_i)$ yra:

$$p(C_i) = n_i N^{-1}. \tag{9}$$

Būsena C_i gali būti laikoma atskiru eksperimentu su atsitikitniais rezultatais. Tuomet šio eksperimento entropija (iš lygybės (7a)):

$$H(C_i) = n_i N^{-1}. \tag{10}$$

Jei vykdome eksperimentą ξ ir rezultatas yra busena C_i, tai mūsų nežinomybė sumažėjo $\ln N - \ln n_i = -\ln p(C_i)$ kartų. Tai atsitinka su tikimybe $p(C_i)$. Taigi vidutinis nežinojimo praradimas (arba žinių gavimas) įgytas vykdant eksperimentą ξ yra jo entropija:

$$H(\xi) = -\sum_{i=1}^{r} p(C_i) \ln p(C_i) \tag{11}$$

(Imant logaritmus su pagrindu 2 rezultatas būtų matuojamas bitais). Matematikos šaka matavimų teorija išplecia šiuos tikimybės bei entropijos apibrėžimus atitinkamai iki begalinių fazinių erdvių, kuriose "rezultatų variantų skaičiavimas" gali būti keičiamas "tūrio matavimu". Entropija $H(\xi)$ tenkina nelygybę:

$$0 \leq H(\xi) \leq \ln r. \tag{12}$$

Kairioji nelygybės dalis tampa lygybe tada ir tik tada, kai $p(C_i)$ bent vienam i: jei žinome iš anksto, jog eksperimento ξ rezultatų būsena bus C_i, tai jokių naujų žinių vykdydami eksperimentą neįgysime. Dešinioji nelygybės dalis tampa lygybe tada ir tik tada, kai $p(C_i) = r^{-1}$ kiekvienam i: daugiausiai žinių suteikiantys eksperimentai yra tokie, kurių visi rezultatai yra vienodai tikėtini. Tokiu būdu eksperimento rezultatų atsitiktinį pasiskirstymą apibūdina entropija. Dar daugiau, trys savokos (entropija, tikimybė ir atsitiktinumas) pasirodo esančios ekvivalenčios griežtai matematine prasme, nes vienos iš jų įvedimas suponuoja kitų dviejų įvedimą. Pavyzdžiui, tokios formulės kaip (7) parodo ryšį tarp entropijos ir tikimybės.

4 Kanoninis ansamblis: sąryšis su statistine fizika

Eksperimentui $\xi = \{C_1, \ldots, C_r\}$, aptartam praėjusioje dalyje, absoliutus atsitiktinumas — visiškas nežinojimas apie eksperimento rezultatus — buvo apibūdintas, kaip entropijos $H(\xi)$ didėjimas, įgaunantis reikšmę $\ln r$ (12). Praktikoje, kiekvienai C_i reikšmei galima priskirti reikšmę E_i. Tai gali būti taškų skaičius kauliuko šone arba dalelės energija elektronvoltais. Jei žinome laukiamą bendrą rezultatą:

$$E = \sum_{i=1}^{r} p(C_i)E_i, \tag{13}$$

tai atskirų rezultatų tikimybės $p(C_i)$, kurios turi tenkinti sąlygą:

$$1 = \sum_{i=1}^{r} p(C_i) \tag{14}$$

yra apibrėžtos santykinio atsitiktinumo prielaida: negalima pastebėti jokio dėsningumo eksperimento rezultatuose, išskyrus bendro rezultato E išlaikymą. Tai ekvivalentu entropijos $H(\xi)$ didinimui priklausančiam nuo lygybių (13) ir (14). Jei $g_1 = E - \sum_{i=1}^{r} p(C_i)E_i$, $g_2 = 1 - \sum_{i=1}^{r} p(C_i)$, o $f = H(\xi) + \beta g_1 + \lambda g_2$, kur β ir λ — Lagranžo daugikliai tai priimant stacionarumo sąlygą $\dfrac{\partial f}{\partial p(C_i)} = 0$, gausime: $\ln p(C_i) = -\beta E_i - (1+\lambda)$ arba $p(c_i) = \dfrac{e^{-\beta E_i}}{e^{1+\lambda}}$. Įstatant reikšmes į (14) ir turint omenyje, kad λ nepriklauso nuo i, gauname:

$$p(C_i) = Z(\beta)^{-1} e^{-\beta E}, \tag{15}$$

kur Z (pasiskirstymo funkcija):

$$Z(\beta) = \sum_{i=1}^{r} e^{-\beta E}. \tag{16}$$

Pastovi reikšmė E išreiškiama kaip:

$$E = -\frac{d \ln Z(\beta)}{d\beta}. \tag{17}$$

Ši funkcija grąžins β, jei pasiskirstymo funkcija yra tolydinė ir griežtai logaritmiškai išgaubta (*strictly logarithmically convex*). Entropiją (11) galime rasti:

$$-\sum_{i=1}^{r} p(C_i) \ln p(C_i) = \sum_{i=1}^{r} p(C_i)[\beta E_i + (1+\lambda)] \tag{18}$$

arba:

$$H(\xi) = \beta E + \ln Z(\beta). \tag{19}$$

Čia reikia turėti omenyje, jog (15)–(19) lygybės yra tik paseka ξ rezultatų santykinio atsitiktinumo (13) ir reikšmės E_i priskyrimo kiekvienai būsenai C_i. Nėra jokios būdingos priežasties dėl kurios toks eksperimentas ξ netiktų kaip modelis "ne-pusiausvyros" sąlygose. (Kai kuriais atvejais papildomos reikšmės F_i, G_i ir tt. gali būti priskirtos konkrečiai būsenai, su žinomomis bendromis reikšmėmis F, G ir t.t.. Lagranžo daugiklių metodas šiam atvejui taip pat tinka, naudojant vektorius vietoje skaliarinių reikšmių). Eksperimentas su ξ atsitiktinėmis tikimybėmis apibrėžiamas (13) lygybe yra vadinamas *kanoniniu statistiniu ansambliu*. Sąryšis su termodinamika atsiranda, kai E_i ir E yra energijos, matuojamos atitinkamais vienetais. Įvedant *bolcmano konstantą* k, *termodinaminė entropija* yra:

$$S = kH. \tag{20}$$

Temperatūra:

$$T = \frac{1}{k\beta}. \tag{21}$$

Termodinaminis potencialas:

$$\Psi = \ln Z(\beta). \tag{22}$$

Helmholtco laisvoji energija:

$$F = -kT\Psi. \tag{23}$$

Lygybė (19) tuomet susiprastina iki:

$$F = E - TS. \tag{24}$$

Turima omenyje, kad (15) ir (21)–(23) duoda būsenos C_i tikimybę kaip:

$$p(C_i) = \frac{e^{-E_i/kT}}{e^{-F/kT}}. \tag{25}$$

Iš šių lygybių galima nesunkiai išvesti kinetinės teorijos lygtis.

5 Tolydiniai pasisikirstymai

Lygybė (11) aprašo informacinę entropiją diskrečiam tikimybių pasisikirstymui $\xi = (p_1, \ldots, p_r)$. Literatūroje galima rasti tokios formos entropijos H išraišką vienmačiam atsitiktiniam kintamajam X, su tankio funkcija $p(x)$:

$$H(X) = -\int_{-\infty}^{\infty} p(x) \ln p(x) dx \tag{26}$$

[3, p. 35]. Nėra nieko klaidingo (26) lygtyje, kaip grynai matematinėje formulėje (įvertinant integralo konvergavimą ir tankio p tolydumą). Visgi fizikiniuose taikymuose koordinatė x šioje lygtyje simbolizuoja abscisių ašį — atstumą nuo fiksuoto atskaitos taško. Šis dydis x išreiškiamas ilgio matavimo vienetais. Tankio funkcija $p(x)$ yra apibrėžta taip, kad

$$P(a \leq X \leq b) = \int_a^b p(x) dx \tag{27}$$

yra tikimybė, kad kintamojo X reikšmė paklius į intervalą $[a, b)$. Kaip tikimybė, kairioji lygties (27) pusė yra bematė. Kadangi nykstamai mažas dydis dx turi ilgio dimensiją, vadinasi tankio $p(x)$ dimensija yra (ilgis)$^{-1}$. Kai $0 \leq z \leq 1$ logaritmą galime išplėsti:

$$-\ln(1 - z) = z + \frac{1}{2}z^2 + \frac{1}{3}z^3 + \ldots . \tag{28}$$

Tam, kad (28) lygybėje išlaikytume nuoseklumą, yra būtina, kad logaritmo argumentas būtų bedimensis. Taigi, matome, kad (26) lygybė yra neteisinga, nes jos dešinėje pusėje esantis logaritmo argumentas yra matuojamas tikimybės tankiu

Kad pataisytume entropijos apibrėžimą tolydianiam atvejui (26), turime normalizuoti logaritmo argumentą, padalydami $p(x)$ iš kito tankio. Pavadinkime šį tankį standartiniu tankiu (*reference density*) $q(x)$. Formulę (26) tuomet galime perrašyti taip:

$$-\int_{-\infty}^{\infty} p(x) \ln \frac{p(x)}{q(x)} dx \tag{29}$$

Jei η yra tikimybės pasiskirstymas $(q_1, \ldots, q_r)$, tuomet kryžminė entropija (*cross-entropy*) arba ξ *entropija* η *atžvilgiu* bagtiniam atvejui:

$$H(\xi|\eta) = -\sum_{i=1}^{r} p_i \ln \frac{p_i}{q_i} = \sum_{i=1}^{r} p_i (\ln q_i - \ln p_i). \tag{30}$$

Čia reikia atkreipti dėmesį, kad jei η yra entropiją didinantis tolyginis pasiskirstymas μ su $q_i = \frac{1}{r}$, kai $1 \leq i \leq r$, tai (30) susiprastina iki:

$$H(\eta|\mu) = -\log r + H(\xi), \tag{31}$$

dydžio mažesnio už nulį. $H(\xi)$ didinimas yra ekvivalentus (31) didinimui.

Grįžtant prie tolydinio atvejo, tarkime kad standartinis tankis apibrėžia atsitiktinio kintamojo Y pasiskirstymą analogiškai kaip (27):

$$P(a \leq Y \leq b) = \int_{a}^{b} q(x) dx. \tag{32}$$

Tuomet dimensiškai teisingą dydį (29) galime interpretuoti kaip X entropiją

$$H(X|Y) = -\int_{-\infty}^{\infty} p(x) \ln \frac{p(x)}{q(x)} dx \tag{33}$$

atitinkamą *standartiniam atsitiktiniam kintamajam* Y. Šiuo atveju lieka tinkamo standartinio atsitiktinio kintamojo parinkimo problema.

Patenkinamą standartinį tankį galima parinkti įvertinus fizikines sąlygas: kadangi turime baigtinį laiką t, per kurį atliekame savo eksperimentą, tai galime patikrinti ne daugiau, kaip $2ct$ ilgio intervalą (kur c — šviesos greitis). (Aukštesniuose matavimuose šis ilgis turėtų būti keičiamas atitinkamu tūriu, pvz.: $\frac{4}{3}\pi c^3 t^3$ trijuose matavimuose.) Žinome, kad eksperimento metu negalime gauti jokios informacijos už šio lango ribų. Taigi tinkamiausias tankis būtų toks, kuris būtų tolygus lango viduje ir lygus nuliui išorėje. Taip pasirinkus tankį ir jį atitinkantį standartinį atsitiktinį kintamąjį M (33) supaprastėja iki:

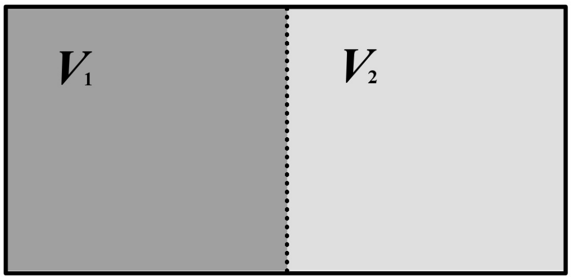
$$H(X|M) = -\int_{x_0-ct}^{x_0+ct} p(x) \ln 2ctp(x) dx, \tag{34}$$

kur x_0 yra mūsų tyrimo įrangos vieta.

6 Gibso paradoksas: priklausomybė nuo tyrimų įrangos

Čia parodysiu, kad "Gibso paradoksas" gali būti sprendžiamas atsižvelgiant į entropijos priklausomybę nuo stebėtojo turimos įrangos.

6.1 Problema



3: Indas su dujomis

Tarkime turime n_1 molių 1 tipo idealiųjų dujų ir n_2 molių 2 tipo idealiųjų dujų, viename inde, pertvara sudalintame į tūrius V_1 ir V_2. Pasirinkdami $\frac{V_1}{V_2} = \frac{n_1}{n_2}$, mes galime palaikyti dujose vienodą temperatūrą $T_1 = T_2$ bei slėgį $P_1 = P_2 = \frac{n_1 RT}{V_1}$. Po kurio laiko pertvarą išimame, leisdami dujoms susimaišyti bendrame tūryje V. Galų gale sistemoje nusistovės pusiausvyra su $n = n_1 + n_2$ moliais dujų mišinio, užimančio bendrą tūrį $V = V_1 + V_2$ su vienodu pasiskirstymu, temperatūra, slėgiu ir nepakitusia bendra sistemos energija.

Jei dujos skiriasi, entropijos pokytis, atsiradęs dėl maišymosi, pagal termodinamiką yra:

$$\Delta S = S_{galutin} - S_{pradin} = nR \ln V - (n_1 R \ln V_1 + n_2 R \ln V_2) \tag{35}$$

arba

$$\Delta S = -nR \left[f \ln f + (1-f) \ln(1-f) \right], \tag{36}$$

kur $f = \frac{n_1}{n} = \frac{V_1}{V}$ yra pirmųjų dujų medžiagos kiekio dalis bendrame mišinyje. Gibsas, nagrinėja konkretesnį atejį, kai $f = \frac{1}{2}$, taigi:

$$\Delta S = nR\ln 2 \tag{37}$$

Šioje lygybėje matosi, kad pats entropijos pokytis nuo dujų rūšies nepriklauso, išskyrus tai, kad dujos turi būti skirtingų rūšių. Tuo tarpu jei sumaišytume du tūrius vienos rūšies dujų, jos irgi susimaišytų, tačiau entropijos pokytis būtų lygus nuliui [7, p.3].

6.2 Sprendimas

Kai sakome, kad dviems skirtingų rūšių dujoms maišantis entropija didėja, turime omenyje, kad dujos gali būti vėl atskirtos į pradines būsenas, būdu kuris atliktų pakitimus išoriniuose kūnuose. Tie pakitimai galėtų būti svorio sumažėjimas arba šilumos perdavimas iš šiltesnio kūno šaltesniam.

Sakydami “pradinė būsena” mes neturime omenyje, kad kiekviena molekulė sugrįžo į savo pradinę padėtį, o tik būseną, kuri yra neatskiriama nuo pradinės pagal makroskopinius parametrus, kuriuos mes stebime. Tam mums užtenka, kad molekulė buvusi tūryje V_1 grįžtų į V_1. Kitais žodžiais tariant, mes galime atstatyti pradinę *termodinaminę būseną*, nusakytą pasirinktais parametrais (pvz.: chemine sudėtimi, bendra energija, tūriu ir medžiagos kiekiu — ir niekuo daugiau).

Tačiau, kai mes sakome, kad du vienos rūšies dujų kiekiai susimaišo be entropijos pokyčio, mes neturime omenyje, kad molekulės iš V_1 gali būti grąžinamos į V_1 be išorinių jėgų poveikio. Termodinamika tvirtina, kad kai entropijos pokytis yra lygus nuliui, tuomet pradinė *termodinaminė būsena* gali būti atstatyta be išorinių jėgų poveikio. Iš tiesų mums užtenka atgal įstatyti pertvarą — kadangi visi stebimi termodinaminiai parametrai sumaišytoms ir atskirtoms tos pačios rūšies dujoms yra identiški, tai termodinaminė būsena nepakito. Taigi nei entropija, nei kuri kita būsenos funkcija nepakito.

Bandymas interpretuoti šį reiškinį kaip netolydų pokytį fizikinėje dujų sandaroje (t.y. jų mikrobūsenų elgesyje), tuomet kai jos tampa vienodos yra klaidingas. Termodinamikos dėsniai aprašo ne hipotetinių mikrobūsenų savybes, bet stebimas makrobūsenų sąvybes — čia turimas omenyje visai ne pradinių mikrobūsenų atstatymas.

Tuomet kai skirtingas dujas pakeičiame vienos rūšies dujomis, skirtumas atsiranda tame, ką mes apibūdiname žodžiais "sugrąžinti" ir "grįžtamasis" — vienos rūšies dujas galima sugrąžinti į pradinę termodinaminę būseną tiesiog įstatant atgal pertvarą, skirtingų rūšių dujoms tektų pritaikyti kokį nors mechanizmą ir atlikti tam tikrą darbą.

Visgi, jei tokie svarstymai paaiškina kodėl skiriasi vienodų bei skirtingų dujų maišymas, jie nesumažina to fakto, kad entropijos pokytis maišantis skirtingos rūšies dujoms nepriklauso nuo tų dujų rūšies svarbos. Nenusižengdami termodinamikos principams galime įsivaizduoti dujas, apie kurias šiandien dar nežinome ir nėra ribos panašumui, kuris gali tarp jų egzistuoti. ΔS visuomet bus nuo jų nepriklausomas.

Mes netgi galime įsivaizduoti dviejų rūšių dujas, kurios yra identiškos visomis savo sąvybėmis, kurios yra svarbios, kol dujos yra maišymosi inde, bet kurių elgesys kokioje nors kitoje aplinkoje skirtųsi. Maišantis tokioms dujoms atsirastų entropijos pokytis, nors vykstantis procesas gali būti identiškas iki smulkiausių detalių (netgi iki kiekvieno atomo trajektorijos) procesui, kuriame sistemos entropija liktų nepakitusi. Šiuo atžvilgiu entropija smarkiai skiriasi nuo energijos.

Termodinaminę būseną apibūdina nedidelis kiekis makroskopinių dydžių, tokių kaip slėgis, tūris, temperatūra, įmagnetinimas, paviršiaus tempimo jėgos, ir t.t. — pažymėkime juos $\{X_1, X_2, \ldots, X_n\}$. Šie dydžiai yra stebimi ir/arba kontroliuojami eksperimentuotojo. n retai kada yra didesnis nei 4. Tai aiškiai kontrastuoja su mikrobūsena, kurią aprašo visų dalyvaujančių atomų (kurių gali būti apie 10^{24} eilės) koordinatės ir judesio kiekiai.

Visos termodinaminės funkcijos — taip pat ir entropija — yra pagal apibrėžimą termodinaminę būseną aprašančios sąvybės:

$$S = S(X_1, \ldots, X_n). \tag{38}$$

Termodinaminius kintamuosius gali atitikti (arba ne) tam tikros mikrobūsenos. Mes laikome bendrą masę ir bendrą energiją "realiomis" fizikinėmis mikrosistemos sąvybėmis, tačiau aukščiau aprašytos prielaidos rodo, kad entropija yra kitokia.

Kad suprastume tai, svarbu pabrėžti, kad "termodinaminė būsena" pažymėta

$$X \equiv \{X_1, \ldots, X_n)\} \tag{39}$$

apibūdina didelę mikrobūsenų aibę $C(X)$. Boltzmanas parodė, kad makrobūsenos entropiją galime interpretuoti kaip

$$S(X) = k \ln W(C), \tag{40}$$

kur $W(C)$ yra mikrobūsenų esančių C užimamas tūris fazinėje erdvėje.

Iš to išplaukia kai kurios svarbios išvados: Termodinaminė entropija $S(X)$ pagal apibrėžimą yra ne vienos mikrobūsenos sąvybė, o tam tikros mikrobūsenų aibės $C(X)$ sąvybė — tai yra tos aibės dydis. Tuomet jei dvi skirtingos mikrobūsenos priklauso C, mes priskiriame tą pačią entropiją sistemoms esančioms šiose skirtingose mikrobūsenose. Bet yra įmanoma ir tai, kad du eksperimentuotojai priskiria skirtingas entropijas S ir S' toms pačioms mikrobūsenoms (kiekvieno atomo padėties ir judesio kiekio prasme) ir abu yra teisūs. Taip yra dėl to, kad šie eksperimentuotojai kiekvienas aprašo sistemą skirtingu makroskopinių parametru rinkiniu, priskirdami šią mikrobūseną skirtingoms aibėms C ir C'.

Iš to galime padaryti dvi svarbias išvadas. Pirmiausia yra svarbu nuspresti keliant problemą kuriuos makroskopinius parametrus (arba laisvės laipsnius) mes matuosime ir/arba kontroliuosime. Tuomet šiame kontekste, taip apibūdintos termodinaminės sistemos entropija bus tam tikra funkcija $S(X_1, \ldots, X_n)$, priklausanti nuo mūsų pasirinktų parametrų. Mes galime tikėtis, kad ji elgsis pagal antrąjį termodinamikos dėsnį $TdS \geq dQ$ tol, kol sistemoje kis tik mūsų numatyti parametrai. Jei kažkas be mūsų žinios pakeis parametrą X_{n+1}, neesantį mūsų stebimų parametrų rinkinyje, tai galės iššaukti tai, kas mums atrodytų kaip antrojo termodinamikos dėsnio pažeidimas, nes mūsų entropijos funkcija $S(X_1, \ldots, X_n)$ gali staiga sumažėti, tuo tarpu, kai jo $S(X_1, \ldots, X_n, X_{n+1})$ padidės.

Tai galima paaiškinti ir kitaip: jei mes neturime informacijos apie tam tikrą būsenos parametrą X_{n+1}, tai negalime ir priversti sistemą atlikti naudingą darbą išnaudojant šio parametro pokytį. Kitaip sakant nepakankamai tiksliai įvertintume ir laisvosios energijos pokytį: $-\Delta F = T\Delta S$, kur ΔF - laisvosios energijos pokytis. Kitaip sakant entropija gailma apibūdinti mūsų turimą informaciją apie darbą, kurį gali atlikti sistema.

Antra, netgi mūsų pasirinktame parametrų rinkinyje nukrypimai nuo antrojo termodinamikos dėsnio yra bent jau įsivaizduojami. Grįžkime prie identiškų (bet skirtingų rūšių) dujų maišymosi. Iš to, kad dujos maišosi be entropijos pokyčio neturėtume daryti išvados, kad jos gali būti vėl atskirtos be išorinių jėgų poveikio. Priešingai identiškų dujų "atskyrimas" yra visiškai *neįmanomas su ar be išorinių jėgų pagalba*. "Identiškos" reiškia, kad nėra paprastesnio būdo kuriuo "atmaišymo" prietaisas galėtų atskirti kuri molekulė pradžioje buvo tūryje V_1, o kuri V_2, išskyrus atsekant kiekvienos molekulės trajektoriją nuo pradinio taško.

Trečiojoje (3) dalyje pamatėme, kad atsitiktinumo, tikimybės ir entropijos (informacinės) priklausomybė nuo turimos bandymų įrangos (tuo atveju nuo galimybės įvertinti duomenų atsitiktinumą). Makroskopinis šios priklausomybės atitikmuo yra termodinaminės entropijos priklausomybė nuo pasirinkto termodinaminės sistemos parametrų rinkinio.

Įdomu, kad 1992 m. E.T. Jaynes rinkdamas medžiagą vadovėliui ir stengdamasis sutikrinti jos atitikimą originaliems darbams, pastebėjo, kad J.W. Gibbs savo ankstyvuosiuose darbuose (o būtent *Heterogeneous Equilibrium* (1875–78 m.) pilnai suprato šį iškilusį paradoksą ir be didesnio vargo, bei dėmesio šiai problemai buvo jį išsprendęs [7, p. 2].

Problema iškilo vėliau, įvairiems autoriams interpretuojant sudėtingai parašytą Gibbso paaiškinimą.

7 Išvados

1. Pamatėme, jog iš informacinės entropijos lygčių galima nesunkiai išvesti lygtis analogiškas statistinės fizikos lygtims aprašančioms kanoninį ansamblį.

2. Galima pastebėti, kad entropija ar ji būtų informacinė, ar termodinaminė, priklauso nuo bandymų įrangos (aparato), tuo pačiu ir nuo stebėtojo pasirinktos aprašymo sistemos. Griežtai matematinėse formulėse, tokiose, kaip (3) ši priklausomybė atsiranda dėl tikimybės sąvokos apibrėžimo. Statistinės fizikos lygtyse, tokiose kaip (40) ši priklausomybė atsiranda dėl pasirinktų termodinaminių parametrų rinkinio.

Literatūra

[1] *Lietuviškoji tarybinė enciklopedija*, t.3, (Leidykla "Mokslas", Vilnius 1978 m.)

[2] A. Tamašauskas ir kiti, *Fizika* t.3, red. Prof. L. Pranevičius, (Vilnius, "Mokslas" 1992 m., ISBN 5-4200-00793-2)

[3] C.E. Shanon, "A Mathematical Theory of Communication",
http://cm.bell-labs.com/cm/ms/what/shannonday/paper.html
The Bell System Technical Journal (Vol. 27, p. 379-423, 623-656, July, October, 1948 m.)

[4] Paul M.B. Vitányi, "Randomness",
http://www.arxiv.org/abs/math.PR/0110086,
Storia del XX Secolo, Vol. 12, (Instituto della Enciclopedia Italiana, 1997 m.).

[5] Jonathan D.H. Smith, "Some Observations on the Concepts of Information-Theoretic Entropy and Randomness",
http://www.mdpi.org/entropy/papers/e3010001.pdf,
Entropy (2001 m., Nr. 3, p. 01-11, ISSN 1099-4300).

[6] Edwin T. Jaynes, "Information Theory and Statistical Mechanics"
http://bayes.wustl.edu/etj/articles/brandeis.pdf,
The Physical Review, (1957 m., Vol. 106, No. 4), p. 620-630

[7] Edwin T. Jaynes, "The Gibbs Paradox",
http://bayes.wustl.edu/etj/articles/gibbs.paradox.pdf,
Maximum Entropy and Bayesian Methods, red. C.R. Smith, G.J. Erickson & P.O. Neudorfer, , (Kluwer Academic Publishers, Dordrecht, Holland, 1992 m., p. 1-22)

[8] James Baugh, "Gibbs' Paradox",
http://www.maxwellian.demon.co.uk/art/lnG/Gibbs1.html,
aplankyta 2002-11-21