VYTAUTO DIDŽIOJO UNIVERSITETAS

Gamtos mokslų fakultetas

Emilis DAMBAUSKAS

Fizikinė informacinės entropijos lygties interpretacija

Kursinis darbas

Vadovas: prof. Liudvikas Pranevičius

KAUNAS 2002

Turinys

1 Įžanga

Šio darbo tikslas — aprašyti kai kuriuos informacinės entropijos ir atsitiktinumo savokų aspektus, ypač jų taikymo fizikoje srityse.

Referate pateikiu fizikinės bei informacinės entropijų apibrėžimus, informacinės entropijos lygties išvedimą, pabrėžiant atsitiktinumo priklausomybę nuo turimos bandymų įrangos. Aptariamas šios priklausomybės makroskopinis atitikmuo, turint omenyje *Prigogine ir George* (1983 m.) pasiūlytą idėją taikyti antrąjį termodinamikos dėsnį kaip "atrankos principą" laiko simetrijai diferencialinėse lygtyse suardyti. Panašios prielaidos dėl turimos įrangos padeda pasiūlyti sprendimo variantą, taip vadinamam "Gibso paradoksui".

2 Fizikinė bei informacinė entropijos

2.1 Fizikinė entropija.

Fizikoje 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]

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, tai jį dar galima vadinti energijos U mikrobūsenos termodinamine tikimybe[2].

2.2 Informacinė entropija.

Informacijos teorijoje entropija — tai dydis, parodantis kiek informacijos gaunama iš bandymo rezultato. Sistemai, turinčiai baigtinę makrobūsenų $\xi = \{C_1, \ldots, C_r\}$ aibę, informacinė entropija išreiškiama lygtimi:

$$H(\xi) = \sum_{i=1}^{r} p_i \ln p_i \tag{3}$$

---

[1] *Lietuviškoji tarybinė enciklopedija*, t.3, (Leidykla “Mokslas”, Vilnius 1978 m.), p. 359

[2] A. Tamašauskas ir kiti, *Fizika*, t.3, (Vilnius, “Mokslas” 1992. ISBN 5-4200-00793-2), p.112, 109.

Čia $H(\xi)$ — informacinė entropija, kur p_i — konkretaus bandymo rezultato tikimybė.

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 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. Informacinė entropija yra adityvi: dviejų nepriklausomų bandymų bendra entropija yra lygi jų atskirų entropijų sumai.

3 Atsitiktinumas, tikimybė ir entropija

Didžioji dalis nesusipratimų su informacinės entropijos sąvoka kyla dėl ilgo rimto atsitiktinumo supratimo trūkumo. Matų teorija tikimybių teorijai suteikė tvirtą aksiomatinį pagrindą iš grynai matematinės pusės, bet jai tiesiog trūko aksiomų taikymo ribų nagrinėjimo. Ir nors klausimas jokiu būdu nėra galutinai išspręstas, *Martin-Löf* (1966) priėjimas prie problemos suponuoja naudingą apibrėžimą klasikiniam atvejui, kurį ir nagrinėju toliau.

Įsivaizduokime mokslinį eksperimentą su N galimais rezultato variantais. Modernaus mokslo susitarimu, “mokslinis” eksperimentas turi būti įmanomas pakartoti ir atgaminti. 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. 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 (kartais naudojami logaritmai, kurių pagrindas 10 ir entropija matuojama Hartliais). 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 (4)):

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

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 Sąryšis su statistine fizika

Eksperimentui $\xi = \{C_1, \ldots, C_r\}$, aptartam praėjusiame skyriuje, 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 (*stationarity*) sąlygą $\partial f/\partial p(C_i) = 0$ gausime: $\ln p(C_i) = -\beta E_i - (1 + \lambda)$ arba $p(c_i) = e^{-\beta E_i}/e^{1+\lambda}$. 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 tam tikrose "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 ansambliu*, nes jis apibendrintai aprašo šio pavadinimo modelius statistinėje fizikoje. 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 lengvai išvesti kinetinės teorijos lygtis.

5 Priklausomybė nuo aparato

Trečiojoje dalyje duotas bendras supratimas apie entropijos, tikimybės ir atsitiktinumo sąvokas. Taip pat buvo parodyta, jog šios sąvokos priklauso nuo turimo aparato galios. *Prigogine ir George* 1983 m. pasiūlė naudoti *Antrąjį termodinamikos dėsnį kaip atrankos principą*, sulaužantį mikroskopinę laiko simetriją, fizikos dėsniuose, išreikštuose diferencialinėmis lygtimis. Tarkime, kad diferencialinė lygtis yra pastovi keičiant laiko kryptį, nes joje naudojamos tik lyginių laipsnių laiko išvestinės. Tuomet ji [lygtis] formaliai gali turėti sprendinius, kurie vyksta pagal ir prieš laiko tėkmę. Reikia kažkaip atmesti sprendinius, kurie vyksta atbulai. Antrojo termodinamikos desnio, kaip atrankos principo idėja leidžia atmesti sprendinius vykstančius prieš laiko tėkmę kaip "neįtikimus". Taip pat čia reikia dar kartą pabrėžti, kad atbulų sprendinių neįtikinamumas priklauso nuo eksperimentinio aparato sudėtingumo.

Antrojo termodinamikos dėsnio kaip atrankos principo panaudojimą iliustruosime pavyzdžiu: tarkime turime diferencialinio pavidalo banginę funkciją:

$$\frac{\partial^2\phi}{\partial r^2} + \frac{2\partial\phi}{r\partial r} + \frac{1\partial^2\phi}{c\partial t^2} \tag{26}$$

pritaikytą sferiškai simetriškam oro slėgiui $\phi(r,t)$ laiku t ir atstume r nuo simetrijos centro (c — garso greitis ore). Yra du paprasti sprendinių tipai: besiplečiančios bangos:

$$\phi = \frac{1}{r}f(r-ct) \tag{27}$$

tam tikrai funkcijai $f(x)$, ir susitraukiančios bangos:

$$\phi = \frac{1}{r}f(r+ct) \tag{28}$$

Pirmojo tipo sprendinys gali apibrėžti garso bangas atsirandančias dėl fizikinio reiškinio simetrijos centre. Iš kitos pusės antrojo tipo sprendiniai yra keblumas teorijai, nes apibrėžia bangas, kurios koherentiškai susieina simetrijos centre. Antrasis termodinamikos dėsnis kaip atrankos principas atmeta susitraukiančias bangas, kaip neįtikimas. Visgi, esant pakankamai sudėtingam aparatui, susitraukiančios bangos, tokios kaip apibrėžta (26) sprendinyje gali atsirasti. Vienas pavyzdys gali būti atominės bombos sprogimo lęšis (*detonation lens*) — labai aiškiai išskirtinis atvejis. Kitas pavyzdys atsiranda elipsinės arba pusiau elipsinės "šnabždesių galerijos" fokuse. Šiam atvejui galerijos lubos turi būti specialiai suprojektuotos ir tiksliai pastatytos.

Gibso paradoksas. Taip vadinamas "Gibso paradoksas" gali būti išsprendžiamas naudojant panašias prielaidas apie naudojamų aparatų galią. Tarkime turime indą su dviejų skirtingų rūšių dujomis. Būsenos, kuriose skirtingos dujos yra atsiskyrusios skirtinguose indo galuose yra aiškiai neįtikimos. Jei turimas aparatas gali atskirti skirtingų rūšių dujas, tai jis galėtų padėti įvertinti vykstantį darbą kartu su termodinaminės entropijos padidėjimu, kuris atsiranda pasekoje to, kad dujos susimaišė. Iš kitos pusės, jei aparatas būtų nepakankamai sudėtingas, jis negalėtų atskirti būsenų, kuriose dujos visiškai susimaišę. Tokiu atveju būsenos su atsiskyrusiomis dujomis termodinaminė entropija būtų tokia pati, kaip ir sumišusios būsenos.

6 Išvados

Pamatėme, jog entropija ar ji būtų informacinė, ar termodinaminė, priklauso nuo bandymų įrangos (aparato). Griežtai matematinėse formulėse, tokiose, kaip (3) ši priklausomybė atsiranda dėl tikimybės sąvokos apibrėžimo. Žinodami esant šią priklausomybę nuo aparato galime išvengti daugelio paradoksų susijusių su entropija.

7 Naudota literatūra

1. Jonathan D.H. Smith, "Some Observations on the Concepts of Information-Theoretic Entropy and Randomness" *Entropy* (2001, 3, 01-11. ISSN 1099-4300).

2. A. Tamašauskas ir kiti, *Fizika* t.3, red. Prof. L. Pranevičius, (Vilnius, "Mokslas" 1992. ISBN 5-4200-00793-2), p.106-113.

3. C.E. Shanon, "A Mathematical Theory of Communication", *The Bell System Technical Journal* (Vol. 27, pp. 379-423, 623-656, July, October, 1948), p. 10-15.

4. *Lietuviškoji tarybinė enciklopedija*, t.3, (Leidykla "Mokslas", Vilnius 1978 m.), p. 359

5. James Baugh, Gibbs' Paradox , aplankyta 2002-11-21