

Statistisen fysiikan satunnaismalleja

1. Boltzmannin jakauma

1.1. Historia

Statistisen fysiikan juuret ovat kaasujen kineettisessä teoriassa, jonka mukaan lämpö on kaasumolekyylien kineettistä energiaa ja paine on kaasumolekyylien seinään törmäämisten aiheuttama voima. Tämä teoria voidaan laskea alkaneeksi Daniel Bernoullin (1700–1782, sveitsiläinen matemaatikko ja fyysikko) työstä, jossa hän selitti Boylen lain $PV = \text{vakio}$ kineettisen teorian oletuksia käyttäen.

Kaasujen kineettistä teoriaa tutkiessaan James Clerk Maxwell (1831–1879, skottilainen matemaattinen fyysikko) ja häntä seuraten Ludwig Boltzmann (1844–1906, itävaltalainen fyysikko ja filosofi) muotoilivat Maxwell–Boltzmann-jakauman, jonka mukaan kaasun molekyylit kulkevat todennäköisyysjakauman mukaista satunnaista nopeutta. Ideaalikaasussa molekyylit eivät vuorovaikuta keskenään ja eri molekyylien nopeudet ovat riippumattomia. Boltzmann yhdisti entropian todennäköisyysjakaumaan ja osoitti, että ns. Boltzmannin yhtälön ratkaisuille entropia kasvaa.

J. Willard Gibbs (1839–1903, yhdysvaltalainen matemaattinen fyysikko) oli varsinaisen statistisen fysiikan isä. Hän määritteli statistinen ensemble ikäänkuin irrallaan ylläolevan kaltaisista kaasu molekyylien törmäysprosesseista. Boltzmannin lähestymisessä energia säilyi aina isommassa mittakaavassa, joskin Maxwell–Boltzmann-jakauma päti yksittäisille atomeille äärettömän isoissa systeemeissä. Gibbsin kanononisessa ensemblessä energia saa vaihdella. Häneltä on peräisin mikrokanoninen, kanoninen ja suurkanoninen jakauma.

Mainitaan vielä Claude Shannon (1916–2001, yhdysvaltalainen matemaatikko, sähköinsinööri ja kryptografi), jota voidaan pitää informaatioteorian isänä. Hän muotoili, miten informaation välityskanavan informaation sisältö voidaan määrittellä kvantitatiivisesti entropiana.

1.2. Optimoinnista

Luvun 1 varsinaisen asia alkaa Luvusta 1.3. Tässä luvussa perustelemme lyhyesti ns. Lagrangen kertojen menetelmän, kts. Esimerkin II.1 loppu. Tätä menetelmää voi soveltaa yleisempiin funktioihin kuin vain konvekseihin.

Joukko $U \subset \mathbb{R}^n$ on *konvekksi*, jos kaikilla $\underline{x}, \underline{y} \in U$ ja kaikilla $\lambda \in [0, 1]$ piste $\lambda \underline{x} + (1 - \lambda)\underline{y}$ kuuluu myös joukkoon U . Kun U on konvekssi alue, sanomme funktiota $f : U \rightarrow \mathbb{R}$, joka toteuttaa $\underline{x}, \underline{y} \in U$ ja kaikilla $\lambda \in [0, 1]$

$$f(\lambda \underline{x} + (1 - \lambda)\underline{y}) \leq \lambda f(\underline{x}) + (1 - \lambda)f(\underline{y})$$

konveksiksi funktioksi. Melko helposti voitaisiin osoittaa, että jos funktio on konvekxi jonkun pisteen ympäristössä, on se myös jatkuva tässä pisteessä. Sanomme funktiota f *konkaaviksi funktioksi*, jos funktio $-f$ on konvekxi. Funktio f on aidosti konvekxi, jos yllä oleva epäyhtälö on aito, kun $\underline{x} \neq \underline{y}$ ja $\lambda \in (0, 1)$.

Kun f on kahdesti jatkuvasti derivoituva ja matriisi $A(\underline{x}) = (\partial_{ij} f(\underline{x}))_{i,j=1}^d$, on konveksisuus yhtäpitävää sen kanssa, että $\underline{v}^\top A(\underline{x}) \underline{v} \geq 0$ kaikilla $\underline{x} \in U$ ja $\underline{v} \in \mathbb{R}^n$. Jos tämä epäyhtälö pätee aitona epäyhtälönä kaikille $\underline{v} \neq \underline{0}$, niin funktio aidosti konvekxi. Voimme alla olettaa turvallisesti, että f on kahdesti derivoituva.

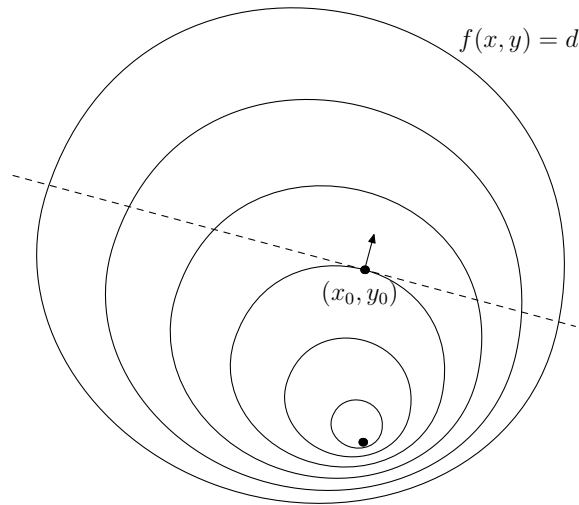
Esimerkki II.1 (Konvekxi rajoitteellinen optimointitehtävä). Olkoon $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ konvekxi funktio ja $a, b, c \in \mathbb{R}$. Tarkastellaan seuraavaa optimointitehtävää

$$\begin{cases} \min f(x, y) \\ \text{ehdolla } ax + by = c, \end{cases} \quad (\text{II.1})$$

jonka ratraisu on siis mikä tahansa pari (x_0, y_0) siten, että $ax_0 + by_0 = c$ ja että $f(x, y) \geq f(x_0, y_0)$ kaikilla $(x, y) \in \mathbb{R}^2$, jotka toteuttavat rajoitteen $ax + by = c$. Ratkaisuja voi olla nolla, yksi tai äärettömän monta. Esitämme seuraavaksi ehdon, joka on riittävä siihen, että (x_0, y_0) on yksikäsitteinen minimi.

Rajoitusehto määrää suoran ja tälle suoralle rajoitettuna f on edelleen konvekxi. Siispä mikä tahansa lokaali minimi on myös globaali minimi.

Joukot, jotka ovat muotoa $\{(x, y) \in \mathbb{R}^2 : f(x, y) \leq d\}$, ovat konvekseja. Niinpä mikäli suora $ax + by = c$ sivuaa tasa-arvopintaa $\{(x, y) \in \mathbb{R}^2 : f(x, y) = d\}$ niiden leikkaus on joko piste tai jana. Kuvasta II.1 havaitsemme, että tällainen sivuamispiste on välttämättä lokaali minimi: jos liikumme sivuamispisteestä hieman sivuun suoraa pitkin, ei funktio voi vähetä.



KUVA II.1. Funktion $f(x, y)$ tasa-arvopintoja $f(x, y) = d$ sekä katkoviivalla lineaarisen rajoitteen määräämä suora. Pisteessä (x_0, y_0) , joka on (usein) yksikäsitteinen, suora sivuaa erästä (yksikäsitteistä) tasa-arvopintaa. Tässä pisteessä funktion f gradientti on suoran normaalin suuntainen.

Kirjoitetaan tämä hieman toisella tavalla. Funktion f gradienttia merkitsemme ∇f . Suoran $ax + by = c$ normaali on $\underline{n} = (a, b) = \nabla(ax + by - c)$. Siispä *riittävä ehto* sille, että suora sivuaa f :n tasa-arvopintaa pisteessä (x_0, y_0) on, että f on jatkuvasti derivoituva pisteessä (x_0, y_0) ja

$$0 = \nabla f(x_0, y_0) + \lambda \underline{n} = \nabla(f + \lambda(g - c))(x_0, y_0)$$

jollain $\lambda \in \mathbb{R}$, missä $g(x, y) = ax + by$. Huomaa, että mikäli $\nabla f(x_0, y_0) = 0$ on (x_0, y_0) funktion f globaali minimi koko tasossa.

Tuntematonta vakiota $\lambda \in \mathbb{R}$ sanotaan *Lagrangen kertojaksi* ja optimointitehtävän ratkaiseminen voidaan pelkistää seuraavanlaiseksi menetelmäksi.

- (1) Ottamalla ylimääräinen muuttuja $\lambda \in \mathbb{R}$ muodostamme funktion

$$f(x, y) + \lambda(g(x, y) - c).$$

- (2) Etsimme tämän funktion gradientin nollakohtan, joka riippuu tietysti λ :sta.
 (3) Sitten yritämme etsiä sellaisen $\lambda \in \mathbb{R}$, että nollakohta (x_0, y_0) toteuttaa rajoitusehdon $g(x_0, y_0) = c$. Mikäli tällainen ratkaisu löytyy, on se tehtävän (II.1) ratkaisu.

Lause II.2. *Olkoon $U \subset \mathbb{R}^n$ konvekksi alue, $f : U \rightarrow \mathbb{R}$ konvekksi funktio ja g_j , $j = 1, 2, \dots, k$, lineaarisia funktiota. Tarkastellaan optimointitehtävää*

$$\begin{cases} \min f(\underline{x}) \\ \text{ehdolla } g_j(\underline{x}) = c_j, \forall j. \end{cases}$$

Mikäli f on jatkuvasti derivoituva pisteessä \underline{x}_0 , piste \underline{x}_0 toteuttaa rajoitteet $g_j(\underline{x}_0) = c_j$ kaikilla j ja on olemassa $\lambda_j \in \mathbb{R}$, $j = 1, 2, \dots, k$ siten, että

$$\nabla \left(f + \sum_{j=1}^k \lambda_j (g_j - c_j) \right) (\underline{x}_0) = 0, \quad (\text{II.2})$$

niin \underline{x}_0 on tehtävän ratkaisu. Mikäli f on aidosti konvekksi on ratkaisu (mikäli se on olemassa) yksikäsitteinen.

Todistus. Kuten edellisessä esimerkissä, huomaamme, että rajoitteen toteuttavat pisteet muodostavat konveksin joukon ja f :n rajoittuma tähän joukkoon on konvekksi. Mikä tahansa lokaali minimi on globaali minimi. Niin ollen aidosti konveksille funktiolle tällainen globaali minimi on yksikäsitteinen.

Gradientti ∇g_j on vakiovektori ja se antaa pinnan $g_j = c_j$ normaalin $\underline{n}_j = \nabla g_j$. Mikäli \underline{x}_0 toteuttaa $g_j(\underline{x}_0) = c_j$ kaikilla j , niin kaikki rajoitteen toteuttavat pisteet ovat muotoa $\underline{x}_0 + \underline{v}$, missä \underline{v} toteuttaa $\underline{v} \cdot \underline{n}_j = 0$ kaikilla j . Jos \underline{x}_0 toteuttaa myös (II.2):n ja f on jatkuvasti derivoituva pisteessä \underline{x}_0 niin,

$$\begin{aligned} \frac{f(\underline{x}_0 + \varepsilon \underline{v}) - f(\underline{x}_0)}{\varepsilon} &= \underline{v} \cdot \nabla f(\underline{x}_0) + o(1) \\ &= \underline{v} \cdot \nabla \left(f + \sum_{j=1}^k \lambda_j (g_j - c_j) \right) (\underline{x}_0) - \underline{v} \cdot \sum_{j=1}^k \lambda_j \underline{n}_j + o(1) \\ &= o(1). \end{aligned}$$

Konveksisuuden perusteella, kun $0 < \varepsilon < 1$,

$$f(\underline{x}_0 + \underline{v}) - f(\underline{x}_0) \geq \frac{f(\underline{x}_0 + \varepsilon \underline{v}) - f(\underline{x}_0)}{\varepsilon}.$$

Ottamalla raja $\varepsilon \rightarrow 0$, näemme, että $f(\underline{x}) \geq f(\underline{x}_0)$ kaikilla rajoitteen toteuttavilla $\underline{x} = \underline{x}_0 + \underline{v}$. Siispä \underline{x}_0 on tehtävän ratkaisu. \square

Tämä tulos yleistyy tietysti monellakin tapaa yleisemmäksi epälineaarisen optimointitehtävän ratkaisumenetelmäksi.

1.3. Informaatioentropia

Haluaisimme määritellä suureen, joka mittaa todennäköisyysjakauman yksinkertaisuutta tai monimutkaisuutta. Claude Shannonia mukaillen *entropia* kertoo montako bittiä informaatiota tarvitaan todennäköisyysjakauman kuvaamiseen siinä mielessä,

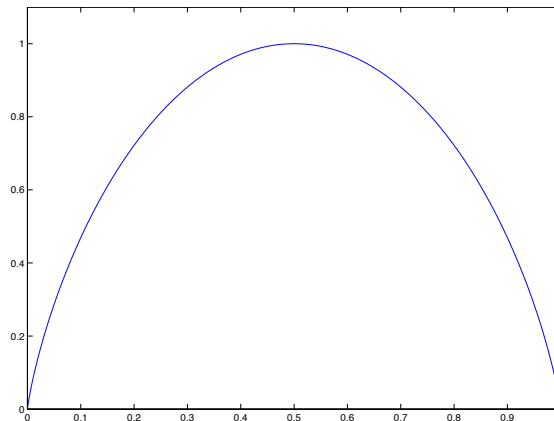
että jos tästä jakaumasta otetaan jono (riippumattomia) näytteitä, niin entropia kertoo kuinka monta bittiä suhteessa jonon pituuteen tarvitaan tyypillisen jonon kuvaamiseen optimaalisessa koodauksessa.

Olkoon $\Omega = \{x_1, x_2, \dots, x_n\}$ äärellinen joukko ja \mathbf{P} todennäköisyysmitta Ω :llä. Kun kirjoitetaan $p_k = \mathbf{P}[\{x_k\}]$, niin huomaamme, että $(p_k)_{k=1, \dots, n}$ määrää \mathbf{P} :n ja päinvastoin. Mitä \mathbf{P} entropia määritellään kaavalla

$$S(\mathbf{P}) = K \sum_k -p_k \log p_k, \quad (\text{II.3})$$

missä K on vakio. Biteistä (esimerkiksi tiedonsiirron informaationa) puhuttaessa on järkevintä käyttää vakiota $K = (\log 2)^{-1}$ muutoin käytämme yksinkertaisesti arvoa $K = 1$. Termodynamiikassa K on Boltzmannin vakio ja entropian yksikkö on JK^{-1} (joulea kelviniä kohti). Käytämme aina sopimusta $0 \log 0 = 0$, joka on luonnollinen valinta jatkuvuuden perusteella.

Esimerkki II.3 (Vääristyneen kolikon entropia). Olkoon $X \sim \text{Bernoulli}(p)$ eli satunnaismuuttuja, jolle $\mathbf{P}[X = 1] = p$ ja $\mathbf{P}[X = 0] = 1 - p$, missä $0 \leq p \leq 1$. Tämän jakauman entropia bitteinä on $S = (-p \log p - (1 - p) \log(1 - p)) / (\log 2)$, jonka kuvaaja on hahmoteltu kuvassa II.2. Tämä funktio on konkaavi ja sen maksimi on pisteessä $p = 1/2$.



KUVA II.2. Funktion $y(x) = (-x \log x - (1 - x) \log(1 - x)) / (\log 2)$ kuvaaja.

Kun $p \approx 0$, niin edellisen esimerkin todennäköisyysjakaumasta otettu jono otoksia sisältää enimmäkseen nollia, joiden joukossa on silloin tällöin ykkösiä. Voimme ajatella, että $\log(1/\mathbf{P}(X = x))$ kuvaa arvon x havaitsemisen yllättävyyttä. Kun havaitsemme 0:n, emme ole yllättyneitä, mutta kun havaitsemme 1, olemme erittäin yllättyneitä. Valitettavasti tämä yllättävä tapahtuma tapahtuu niin harvoin, että entropia, joka on siis tässä tulkinnaissa *yllättävyyden odotusarvo*, on pieni. Jakauksen yllättävyys on maksimaalinen odotusarvon mielessä, kun jakauma on tasainen, kuten seuraavasta esimerkistä näemme.

Esimerkki II.4. Kun annettulla Ω maksimoimme S :ää kaikkien Ω :n todennäköisyysmittojen \mathbf{P} muodostamassa joukossa, päädyimme seuraavaan *rajoitteelliseen optimointitehtävään*

$$\left\{ \begin{array}{l} \min \sum_k p_k \log p_k \\ \text{ehdolla } \sum_k p_k = 1 \\ (p_k \geq 0 \text{ kaikilla } k) \end{array} \right.$$

joka ratkaistaan *Lagranjen kertojan menetelmällä*. Ratkaisu $\underline{p} = (p_k)_{k=1,\dots,n}$ on paikallinen minimi vain, jos on olemassa $\lambda \in \mathbb{R}$ siten, että lausekkeen

$$\sum_k p_k \log p_k + \lambda \left(\sum_k p_k - 1 \right) \quad (\text{II.4})$$

gradientti muuttujan \underline{p} suhteen on nolla ja \underline{p} toteuttaa rajoitteen $\sum_k p_k = 1$. Derivoimalla lauseketta (II.4) saamme ehdoksi

$$\log p_k + 1 + \lambda = 0$$

kaikilla k . Tästä seuraa, että kaikilla k

$$p_k = \frac{1}{n} \quad \text{ja} \quad \lambda = \log n - 1.$$

Huomasimme siis, että *tasainen todennäköisyysjakauma maksimoi entropian*, kun ainoa rajoite on vaatimus, että \underline{p} määrittää todennäköisyysmitan.

Se, että löydetty piste on todella maksimi, seuraa siitä, että maksimoitava funktio on konkaavi.

Seuraava propositio on tärkein perustelu entropialle epäjärjestyksen mittana.

Propositio II.5 (*S:n ominaisuudet*). *Olkoon S ja P kuten yllä.*

- (1) $S(\mathbf{P}) \geq 0$ ja $S(\mathbf{P}) = 0$ jos ja vain jos $\mathbf{P}(\{x_k\}) = 1$ jollakin k .
- (2) *S:n maksimi Ω :n todennäköisyysmitoilla saavutetaan tasaisella todennäköisyysmitalla $\mathbf{P}_{tas.}(\{x_k\}) = 1/|\Omega|$ kaikilla k , jolloin*

$$S(\mathbf{P}_{tas.}) = \log |\Omega|. \quad (\text{II.5})$$

- (3) *Olkoon \mathbf{P}_{12} todennäköisyysmitta tuloavaruudella $\Omega = \Omega_1 \times \Omega_2$ ja \mathbf{P}_1 ja \mathbf{P}_2 sen marginaalijakaumat Ω_1 :llä ja Ω_2 :lla. Silloin*

$$S(\mathbf{P}_{12}) \leq S(\mathbf{P}_1) + S(\mathbf{P}_2)$$

ja yhtäsuuruus pätee vain kun $\mathbf{P}_{12} = \mathbf{P}_1 \times \mathbf{P}_2$

Todistus. Enimmäinen väite seuraa siitä, että $-x \log x > 0$, kun $0 < x < 1$. Toisen väitteen todistimme Esimerkissä II.4.

Kolmatta väitettä varten olkoon r_{jk} yksiidien todennäköisyydet avaruudella $\Omega = \Omega_1 \times \Omega_2$ ja marginaalijakaumat $p_j = \sum_k r_{jk}$ ja $q_k = \sum_j r_{jk}$.

$$\begin{aligned} S(\mathbf{P}_{12}) &= \sum_{jk} -r_{jk} \log r_{jk} = \sum_{j,k} -r_{jk} \log p_j + \sum_{jk} -r_{jk} \log \frac{r_{jk}}{p_j} \\ &= S(\mathbf{P}_1) + \sum_{j,k} r_{jk} \log \frac{p_j}{r_{jk}} \\ &= S(\mathbf{P}_1) + S(\mathbf{P}_2) + \sum_{j,k} r_{jk} \log \left(\frac{p_j q_k}{r_{jk}} \right) \\ &\leq S(\mathbf{P}_1) + S(\mathbf{P}_2) + \log \left(\sum_{j,k} r_{jk} \frac{p_j q_k}{r_{jk}} \right) = S(\mathbf{P}_1) + S(\mathbf{P}_2) \end{aligned} \quad (\text{II.6})$$

missä käytimme Jensenin epäyhtälöä¹ konkaaviin funktioon $x \mapsto \log x$. Yhtäsuuruus pätee jos ja vain jos $p_j q_k / r_{jk}$ on vakio kaikilla j, k eli kun $\mathbf{P}_{12} = \mathbf{P}_1 \times \mathbf{P}_2$. \square

¹Jensenin epäyhtälö: jos $f : [a, b] \rightarrow \mathbb{R}$ on konvekksi ja $X \in [a, b]$ on satunnaismuuttuja, niin $f(\mathbf{E}[X]) \leq \mathbf{E}[f(X)]$. Konkaaville funktiolle epäyhtälön suunta on vastakkainen. Tässä käytämme epäyhtälöä muodossa $\sum_\alpha \lambda_\alpha f(x_\alpha) \leq f(\sum_\alpha \lambda_\alpha x_\alpha)$, kun $\lambda_\alpha = r_{jk}$ ja $x_\alpha = \frac{p_j q_k}{r_{jk}}$

1.4. Boltzmannin jakauma

Jatkamme nyt Esimerkkiä II.4 ja perustelemme Boltzmannin jakauman entropian avulla. Tämä argumentti voitaisiin nähdä jonkinlaisena termodynamiikan 2. pääsäännön muotoiluna: tasapainossa systeemin entropia on maksimaalinen. Otamme sen kuitenkin lähinnä todennäköisyyslaskennan motivaationa tälle jakaumalle: se on mahdollisimman tasainen todennäköisyysjakauma annetulla rajoitteella.

Esimerkki II.6 (Boltzmannin jakauma perustelu entropialla). Olkoon $\phi : \Omega \rightarrow \mathbb{R}$ jokin satunnaismuuttuja. Ajatteleme, että ϕ on joku suure, josta olemme saaneet tietoa mittaamalla. Koska ϕ saa satunnaisesti vaihtelevia arvoja mittauksesta toiseen, ajatellaamme, että olemme mitanneet tälle suurelle otoskeskiarvon $N^{-1} \sum_{j=1}^N \phi(X_j) = F_0$, missä X_j :t ovat riippumattomia ja jakautuneet \mathbb{P} :n mukaisesti. Jotta mittausasetelma on järkevä, voimme olettaa, että N on suuri. Niinpä päättelemme, että \mathbb{P} on todennäköisyysmitta, jolle $\mathbb{E}_{\mathbb{P}}[\phi] = F_0$.

Nyt maksimoimme S :ää kaikkien jakaumien \mathbb{P} joukossa, jotka toteuttavat sekä ehdon $\underline{p} = (p_k)_{k=1, \dots, n}$ normalisaatiosta todennäköisyysjakaumana että ehdon ϕ :n odotusarvosta. Päädyimme seuraavaan rajoitteelliseen optimointitehtävään

$$\left\{ \begin{array}{l} \min \sum_k p_k \log p_k \\ \text{ehdoilla } \sum_k p_k = 1 \\ \sum_k p_k \phi(x_k) = F_0 \\ (p_k \geq 0 \text{ kaikilla } k) \end{array} \right.$$

Nyt Lagrangen kertojia tarvitaan kaksi λ ja β ja vaadimme että lausekkeen

$$\sum_k p_k \log p_k + \lambda \left(\sum_k p_k - 1 \right) + \beta \sum_k (p_k \phi(x_k) - F_0)$$

gradientti muuttujan \underline{p} suhteen on nolla. Derivoimalla lauseketta saamme ehdoksi

$$\log p_k + 1 + \lambda + \beta \phi(x_k) = 0$$

kaikilla k . Tästä seuraa, että kaikilla k

$$p_k = \frac{1}{Z(\beta)} \exp(-\beta \phi(x_k)) \quad \text{ja} \quad \lambda = \log Z(\beta) - 1,$$

missä $Z(\beta) = \sum_k \exp(-\beta \phi(x_k))$ on nk. *tilasumma* tai *partitiofunktio*. Merkitään tätä todennäköisyysmittaa \mathbb{P}_{β} :lla (ja vastaavia odotusarvoa \mathbb{E}_{β} ja varianssia Var_{β} jne.). Kun huomaamme, että

$$\mathbb{E}_{\beta}[\phi] = -\partial_{\beta} \log Z(\beta),$$

niin vakio β määräytyy vaatimalla, että

$$-\partial_{\beta} \log Z(\beta) = F_0.$$

On melko helppo huomata, että tämä ratkaisu on yksikäsitteinen. Nimittäin

$$\partial_{\beta} (-\partial_{\beta} \log Z(\beta)) = -\text{Var}_{\beta}[\phi] < 0 \tag{II.7}$$

eli $\beta \mapsto \mathbb{E}_{\beta}[\phi]$ on aidosti vähenevä funktio.

Muotoilemme seuraavan *entropian maksimaalisuuden periaatteen*, johon Boltzmannin jakauman perustamme: oletetaan, että systeemi voi vaihtaa energiaa ympäristönsä kanssa, ja että

jos systeemi on tasapainossa ympäristön kanssa, sen entropia on maksimaalinen

Edellisen esimerkin perusteella voimme tehdä seuraavan määritelmän.

Määritelmä II.7. Olkoon Ω äärellinen joukko ja $H : \Omega \rightarrow \mathbb{R}$ funktio, jota kutsumme *energiaksi* tai systeemin *Hamiltonin funktioksi*. Olkoon $\beta \geq 0$, määritellään *tilasumma* tai *partitiofunktio*

$$Z(\beta) = \sum_{x \in \Omega} \exp(-\beta H(x)) \quad (\text{II.8})$$

ja *Boltzmannin jakauma* (tai *Gibbsin jakauma*)

$$P_\beta(x) = \frac{\exp(-\beta H(x))}{Z(\beta)}. \quad (\text{II.9})$$

Boltzmannin jakaumalle on erilaisia perusteluita, joista edellä on esitetty yksi. Samoin entropian määritelmää voidaan perustella esimerkiksi seuraavan Proposition II.5 tai vaikka Tehtävän II.3 avulla. Otamme nämä kuitenkin ensisijaisesti määritelmänä — tosin varsin keskeisinä määritelmänä, koska Boltzmannin jakauma on aina (tasapaino-) statistisen fysiikan mallien taustalla.

Huomautus II.8. Yhtälön (II.7) perusteella $\beta \mapsto E_\beta[H]$ on vähenevä. Jos haluamme tulkita sisäisen energian $E_\beta[H]$ kertovan lämpötilasta T täytyy $T \mapsto \beta(T)$:n olla vähenevä funktio.

Huomataan myös, että kun $\beta = 0$, kaikki tilat ovat yhtä todennäköisiä. Tämä tulkintaan niin, että $T = \infty$. Vastaavasti, jos x^* on tila, jolle $H(x) > H(x^*)$ kaikilla $x \neq x^*$, niin silloin rajalla $\beta \rightarrow \infty$ todennäköisyyksimmä $P_\infty[x] = \delta_{x,x^*}$. Eli $\beta = \infty$ vastaa $T = 0$:aa.

Havaintomme β ja lämpötilan T yhteydestä on tähän mennessä seuraava

Mikäli $\beta = \beta(T)$, täytyy $T \mapsto \beta(T)$:n olla vähenevä ja toteuttaa $\beta(0) = \infty$ ja $\beta(\infty) = 0$.

1.5. Kahden systeemin tasapaino ja lämpötila

Tarkastellaan kahta systeemiä $x \in \Omega_1$ ja $y \in \Omega_2$, jotka ovat aluksi erillään ja noudattavat jakaumaa, joka koostuu kahdesta Boltzmannin jakaumasta, eli

$$P_{\beta_1, \beta_2}(x, y) = \frac{\exp[-\beta_1 H_1(x) - \beta_2 H_2(y)]}{Z_1(\beta_1) Z_2(\beta_2)},$$

missä $\beta_j \geq 0$. Jos $\beta_1 \neq \beta_2$, tämän jakauman entropia ei ole maksimaalinen, vaan voimme eräs Boltzmannin jakauma $P_{\beta_{12}}$ maksimoi entropian niiden todennäköisyysjakaumien joukossa $\Omega_1 \times \Omega_2$:lla, joille $E[H_1 + H_2]$ on annettu vakio.

Kun nämä kaksi syteemiä ovat tasapainossa keskenään, ovat ne erityisesti tasapainossa ympäristön kanssa ja entropian maksimaalisuuden periaatteen mukaan ne noudattavat yhteistä Boltzmannin jakaumaa

$$P_{\beta_{12}}(x, y) = \frac{\exp[-\beta_{12}(H_1(x) + H_2(y))]}{Z_{12}(\beta_{12})}.$$

Lämpötilan idea pohjautuu siihen, että kaksi systeemisä voivat olla tasapainossa keskenään vain, kun niiden lämpötilat ovat samat. Siispä havaitsemme, että

funktio $T \mapsto \beta(T)$ on kaikille systeemeille sama.

On järkevää määritellä lämpötila $T(\beta)$ niin, että jollekin yksinkertaiselle systeemille sisäinen energia on $\mathbf{E}_\beta[H] = cNT(\beta)$, missä N on systeemin koko. Osoittautuu, että ideaalikaasulle $\mathbf{E}_\beta[H] \sim \beta^{-1}$. Siispä määrittelemme lämpötilan T siten, että

$$\beta = \frac{1}{T}. \quad (\text{II.10})$$

Käytämme yleensä merkintää β Boltzmannin jakauman parametrille, joka tarvittaessa ymmärretään käänteislämpötilana, mutta jota voidaan pitää vain parametri-
na matemaattisesti. Korostamme kuitenkin, että aina

Boltzmannin jakaumasta puhuttaessa voimme tulkita $\beta = 1/T$.

Esimerkki II.9 (Ideaalikaasun tilasumma ja sisäinen energia.). Tarkastellaan ideaalikaasua, jossa on N hiukkasta, joiden massa on m ja jotka ovat d -dimensiosessa avaruudessa. Tätä systeemiä kuvaa Hamiltonin funktio

$$H(\underline{x}, \underline{u}) = \sum_{k=1}^{dN} \frac{u_k^2}{2m}$$

missä $\underline{x} \in \mathbb{R}^{dN}$ ja $\underline{u} \in \mathbb{R}^{dN}$ ovat hiukkasten paikat ja liikemäärät koottuna yhteen vektoriin. Edellä olevasta poiketen todennäköisyysavaruus ei ole diskreetti vaan jatkuva. Tästä huolimatta määrittelemme seuraavan jatkuvan jakauman Boltzmannin jakauman tapaan

$$P_\beta(\underline{x} \in A, \underline{u} \in B) = \frac{1}{Z(\beta)} \left(\int_A dx_1 \dots dx_{dN} \right) \left(\int_B \exp(-\beta H(\underline{u})) du_1 \dots du_{dN} \right)$$

missä A ja B ovat Borel-joukkoja ja vakio $Z(\beta)$ on tilasumma, joka jatkuvalle jakaumalle on

$$Z(\beta) = \left(\int dx_1 \dots dx_{dN} \right) \left(\int_{\mathbb{R}^{dN}} \exp(-\beta H(\underline{u})) du_1 \dots du_{dN} \right).$$

Kutsumme tätä jakaumaa *Maxwell-Boltzmann-jakaumaksi*. Tekemällä muuttujan vaihto $u_k \sqrt{\beta/m} = v_k$ saamme

$$Z(\beta) = \left(\frac{\beta}{m} \right)^{-dN/2} I^{dN} V^N,$$

missä $I := \int_{\mathbb{R}} \exp(-v^2/2) dv = \sqrt{2\pi}$ ja V on astian tilavuus. Siispä sisäiseksi energiaksi saadaan

$$\mathbf{E}_\beta[H] = -\partial_\beta \log Z(\beta) = \frac{d}{2} N \beta^{-1} = \frac{3}{2} N \beta^{-1},$$

kun $d = 3$.

Viimeinen huomiomme β :sta ja T :stä:

kaasujen kineettisessä teoriassa lämpö on kaasumolekyylien kineettistä energiaa siten, että $\mathbf{E}_\beta[H] = \text{vakio} \cdot \beta^{-1}$.

1.6. Lisää esimerkkejä ideaalikaasusta

Jatkamme vielä muutamalla laskulla koskien ideaalikaasua ja kaasujen kineettistä teoriaa.

Esimerkki II.10 (Ideaalikaasun paine ja tilanyhtälö). Ideaalikaasun (kts. Esimerkki II.9) paine voidaan laskea seuraavalla tavalla lähtien paineen luonnollisesta määritelmästä: P saadaan kertomalla seinään törmäysten lukumäärä niissä muuttuneella liikemäärällä ja jakamalla tämä aikayksiköllä ja pinta-alalla. Selitämme tämän tarkemmin alla.

Oletetaan, että astia on $[0, L]^d \subset \mathbb{R}^d$ laatikko, jonka sivut ovat koordinaattiakselien suuntaiset. Tarkastellaan seinää, jonka määrää $x_1 = 0$, ja hiukkasta, jonka liikemäärä x_1 -suuntaan

on u_1 . Kun osuu mihin tahansa seinään, sen seinää kohtisuoraan vasten oleva nopeuskomponentti vaihtaa merkkiä (kimmainen törmäys). Tarkastelavaan seinään osuessa tämän hiukkasen liikemäärä x_1 -suunnassa muuttuu $-|u_1|$:stä $|u_1|$:teen. Siltä kuluu aika $t_0 := 2Lm/|u_1|$ kulkea laatikon päästä toiseen ja takaisin. Jos tämän hiukkasen hetkellinen liikemäärä x_1 -suuntaan on $u_1(t)$ ja seinään kohdistama hetkellinen voima $F_1(t)$, niin voima keskimäärin ajassa on suuruudeltaan

$$\bar{F}_1 = \frac{1}{t_0} \int_0^{t_0} F_1(t) dt \stackrel{*}{=} \frac{1}{t_0} \int_0^{t_0} \dot{u}_1(t) dt = \frac{2|u_1|}{2Lm/|u_1|} = \frac{u_1^2}{mL}.$$

Kohdassa * käytimme Newtonin mekaniikan lakeja. Seinään kohdistuva paine, kun $A = L^{d-1}$ on seinän $x_1 = 0$ pinta-ala, voidaan kirjoittaa muotoon

$$P = N \frac{\mathbb{E}[\bar{F}_1]}{A} = \frac{N\mathbb{E}[u_1^2]}{mV} = \frac{2\mathbb{E}[H]}{dV} = \frac{N\beta^{-1}}{V}$$

joka yhdessä β lämpötilatulkinnan kanssa antaa ideaalikaasun kokeellisesti todetun, tutun tilanyhtälön $PV = nRT$.

Esimerkki II.11 (Ideaalikaasun entropia ja vapaaenergia). Jatkamme Esimerkkejä II.9 ja II.10.

Käytämme entropialle määritelmää

$$S = S(\mathbf{P}_\beta) = -\mathbb{E}[\log(\exp(-\beta H)/Z(\beta))] = \beta\mathbb{E}_\beta[H] + \log Z(\beta)$$

Sijoittamalla Esimerkin II.9 tulokset saamme

$$S = \frac{d}{2}N + N \log \left[\left(\frac{\beta_0}{\beta} \right)^{d/2} \frac{V}{V_0} \right]$$

missä vakiot β_0 ja V_0 toteuttavat

$$\beta_0^{d/2}/V_0 = (2\pi m)^{d/2}.$$

Jos käytämme sisäisen energian ja lämpötilan määritelmiä $U = \mathbb{E}_\beta[H]$ ja $T = \beta^{-1}$, niin termodynaaminen vapaa energia on

$$F = U - TS = -\beta^{-1} \log Z = -\beta^{-1} N \log \left[\left(\frac{\beta_0}{\beta} \right)^{d/2} \frac{V}{V_0} \right].$$